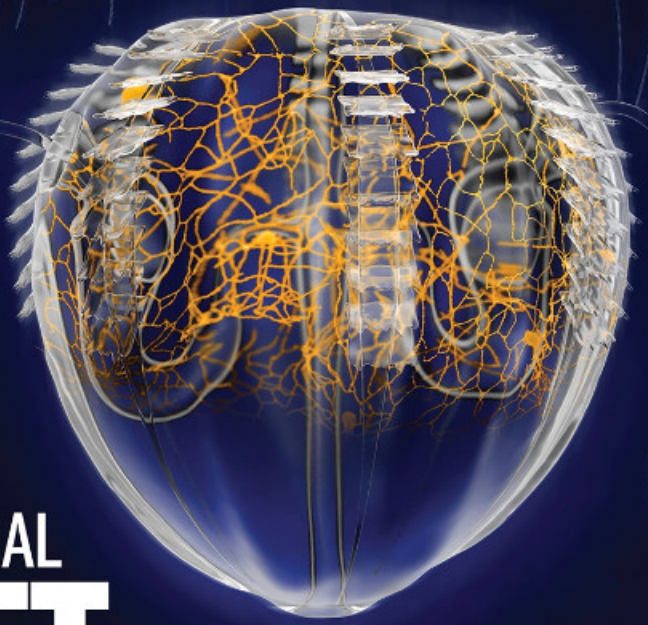
Can nuclear weapons be revamped without tests? p. 234

Membrane innovation for water treatment needs a rethink p. 242

Quantum kittens, cats, and lions p. 274

Science

$15
21 APRIL 2023
science.org
AAAS



NEURAL NET

Clues to brain evolution from comb jellies

pp. 241 & 293

EDITORIAL

Judicial interference with mifepristone

In the days since Texas federal judge Matthew J. Kacsmaryk invalidated the approval by the US Food and Drug Administration (FDA) of mifepristone, a medication used to terminate pregnancy, a shock wave of concern has swept through many people, organizations, and companies that work closely with the agency. The strong opposition reflects the high stakes not only for pregnant persons and for the FDA, but also for the scientific process of drug development and public access to safe and effective medications. Twists and turns in the case are already happening. A federal appeals court stayed the full suspension of mifepristone, but permitted multiple restrictions on its availability. Then the Supreme Court, which recently overturned the constitutional right to abortion, kept the status quo in place for a few days while considering the government's appeal. The results of the legal battle will be enormously consequential for reproductive health care—and far beyond, for innovation, science, and health.

> "The results of the legal battle will be enormously consequential..."

The FDA plays such an important role in the health of Americans that it is easy to take its functions for granted. More than 15,000 agency employees regulate an estimated $2.7 trillion in consumer goods, including all medical products. Over more than a century, the FDA has developed extensive processes that govern the collection and review of preclinical and clinical data on biologics and drugs with defined scientific and legal-regulatory standards, earning high levels of trust from the public in the process.

The agency's review of mifepristone in 2000 was thorough and fair. The drug's manufacturer submitted a large dataset for the agency's experts to review. An external advisory panel supported its approval. After a 6-month review, FDA's scientific staff concluded that mifepristone is safe and effective. Over the past two decades, the medication's safety record has grown stronger, with major medical professional associations in full support of access. Over time, the FDA, after thorough safety reviews, loosened restrictions on distribution.

The FDA's expertise and diligence, however, barely seemed to matter to Kacsmaryk in his unprecedented decision last week. The judge's use of extreme rhetoric, reliance on noncredible sources, and tendentious reasoning may have raised the hopes of the plaintiffs, who have a strong ideological opposition to abortion, but the decision also shredded any pretense of judicial objectivity and lowered the bar for efforts to overturn well-considered and justified determinations by the FDA.

It's no surprise that a broad diversity of patient advocacy groups have expressed alarm. "The implications of this ruling go far beyond mifepristone," read a statement of opposition signed by 30 organizations representing those with serious health conditions. They warned, "If this judge's ruling is allowed to stand, patients may no longer have the security of knowing that determinations about drug safety ultimately lie with the experts." Also objecting are pharmaceutical companies that recognize the threat to innovation posed by arbitrary judicial orders. The interim director of the biotech organization BIO described the decision as "an assault on science" while the pharmaceutical organization PhRMA expressed "serious concerns with any court substituting its opinion for the FDA's expert approval decision-making."

In an amicus brief to the US Court of Appeals for the Fifth Circuit, many leading companies and trade organizations declared that if upheld, the decision could "empower any plaintiff to grind drug approvals to a halt, disrupting patients' access to critical medicines." The brief also noted the potential to "wreak havoc on drug development and approval generally, causing widespread harm to patients, providers, and the entire pharmaceutical industry."

As the former commissioner and principal deputy commissioner of the FDA, we couldn't have said it better ourselves. The FDA is a unique institution, bringing together intellectual resources from inside and outside government to make decisions on thousands of products each year. Once courts dismiss core scientific judgments by the agency, there is no reason to believe they will limit themselves to this one medication. There is already political pressure against vaccines, antidepressants and other psychotropic medication, and certain cell-derived therapies. If judges begin to dictate the terms of medication access, then others will seek to use ideology and influence to advance their agendas.

Respect for the integrity of the FDA underlies decades of progress in using science to save lives. Cracks in this foundation are as dangerous as they are unwarranted.

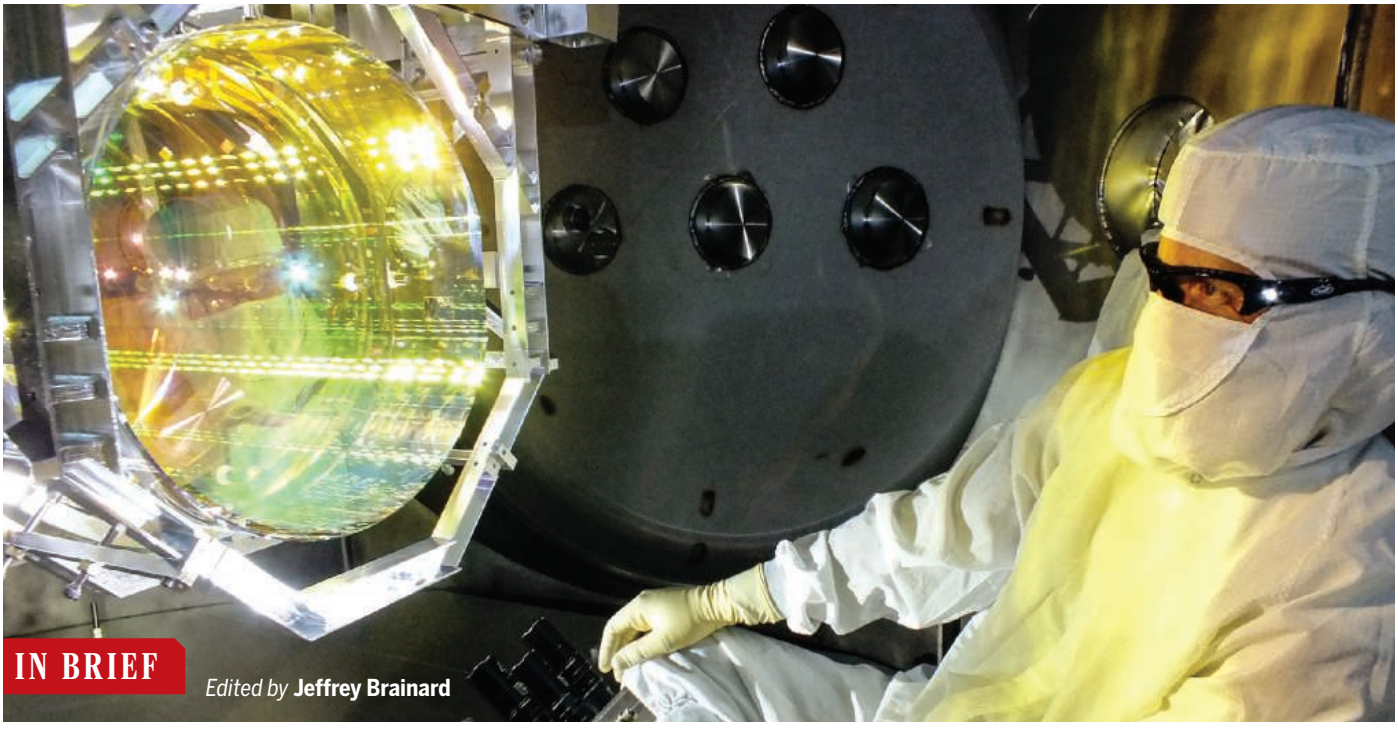
–Margaret Hamburg and Joshua Sharfstein

Margaret Hamburg is copresident of The InterAcademy Partnership for Science Health and Policy, Washington DC, USA. She served as commissioner of the FDA from 2009 to 2015. peggy@hbfam.net

Joshua Sharfstein is a professor of the practice at the Johns Hopkins Bloomberg School of Public Health, Balitmore, MD, USA. He served as principal deputy commissioner of the FDA from 2009 to 2011. joshua.sharfstein@jhu.edu

Published online 13 April 2023; 10.1126/science.adi2587

> "It put kids at risk, there was false hope for families."
>
> **University of California, Davis, cell biologist Paul Knoepfler,** in *Vice*, about a recently ended Duke University program that gave autistic children unproven stem cell treatments. Allowed under FDA's "compassionate use" rules, it charged parents up to $15,000.



IN BRIEF

Edited by **Jeffrey Brainard**

ASTROPHYSICS

India to build gravitational wave detector

After more than a decade of planning, India this week approved building a gargantuan detector to sense ripples in space and time called gravitational waves, extending an international network of such devices. Expected to be operational by the decade's end, the new detector, near Aundha in western India, will be nearly identical to those in the Laser Interferometric Gravitational-Wave Observatory (LIGO) in Louisiana and Washington state. Each of those detectors is an L-shaped optical device called an interferometer with arms 4 kilometers long that make ultraprecise measurements of space in perpendicular directions. LIGO-India will use a spare set of LIGO mirrors and lasers. India will spend $320 million to construct the vacuum chamber and buildings to house the device. Since 2015, LIGO and Italy's Virgo detector have detected the fleeting signals of gravitational waves from 90 high-energy celestial events in which two massive objects such as black holes collided. Adding another detector should enable scientists to more precisely pinpoint sources in the sky.

LIGO-India will deploy a set of ultraprecise, 40-kilogram mirrors left over from the U.S. LIGO project.

Ghana OKs new malaria vaccine

INFECTIOUS DISEASES | Although results from clinical trials are still pending, Ghana has become the first country to approve a new malaria vaccine, which will likely be less expensive than an existing vaccine. The new vaccine, called R21/Matrix-M, is unlikely to be used widely before the World Health Organization (WHO) endorses it, which would enable international funders to purchase the vaccine. R21/Matrix-M showed promising results in a preliminary trial involving 450 children, but researchers have yet to report final results from a trial with 4800 children in four countries. Ghana is already using the first malaria vaccine, endorsed by WHO and called Mosquirix; more than 450,000 children had received at least one dose by December 2022. R21/Matrix-M will likely require at least four doses per child at $3 a dose, and global health experts caution that vaccines should be weighed against cheaper malaria-control tools, such as insecticide-treated bed nets.

U.S. wields light touch on fusion

ENERGY | The U.S. Nuclear Regulatory Commission (NRC) decided last week to regulate future fusion power plants under standards used for particle accelerators and radioactive medical technologies, not the more stringent rules it requires for nuclear fission power plants. The Fusion Industry Association (FIA) had sought the decision, contending that fusion power generation will be less dangerous than fission plants and shouldn't be regulated as tightly. Operators of fusion plants will need to handle small amounts of tritium—a short-lived radioactive isotope of hydrogen—and safely dispose of low-level waste consisting of reactor parts made radioactive by fusion reactions. The hazards are similar to those of particle accelerators, according to FIA. Although a commercial fusion reactor remains possibly decades away, the NRC decision "will give fusion developers the regulatory certainty they need to innovate," FIA says.

Early test for Parkinson's disease

BIOMEDICINE | A large study has shown that a test can indicate a person has Parkinson's disease before they start having symptoms. The chronic degenerative disease currently lacks a definitive biochemical test. A research team recruited 1123 participants, some of whom had symptoms of Parkinson's, and used spinal taps to measure their levels of a protein called alpha-synuclein, which clumps and damages brain cells in people with the disease. Study participants whose alpha-synuclein clumped to an extent above a threshold level were deemed to have Parkinson's. In 88% of the subjects, the test accurately indicated whether they had the disease, researchers reported last week in *The Lancet Neurology*. The new approach is likely too costly and invasive to be used widely for screening but may inform research into treatments, scientists say.

Roll-up space telescope mirrors

ASTRONOMY | Researchers have developed a technique to produce flexible, high-quality mirrors for space telescopes. The mirrors could be rolled up and stored compactly in rockets and enable more-powerful orbiting observatories. A team at the Max Planck Institute for Extraterrestrial Physics created prototypes up to 30 centimeters wide and plans to make larger ones. The mirrors are made by rotating liquid in a vacuum chamber as a parabolic "mold," onto which chemical vapors are deposited to form a polymer layer. A reflective metal layer is then added. They described the work in *Adaptive Optics*.

PHOTO: S. CASTELLANI; FROM DODD, FRONTONI, GALLI 2023

ARCHAEOLOGY

Roman emperors played winemaker for a day

In ancient Rome, wine mattered so much that the emperor himself opened each year's wine grape harvest by cutting a symbolic cluster of grapes. Now, archaeologists have discovered that an elaborate facility just outside Rome probably was built specially for this occasion. They excavated the unusual, 1000-square-meter space in 2017 and 2018 and dated it to about 240 C.E. The facility was like a dude ranch, where the nobility played winemaker for a day. It featured high-end imported red marble on which workers squashed grapes; juice flowed through marble-lined channels into fermentation vats beneath a decorated floor. The slippery floor material, unusual in a Roman winery, suggested the facility was designed for showing off, not efficiency, the research team writes this week in *Antiquity*.

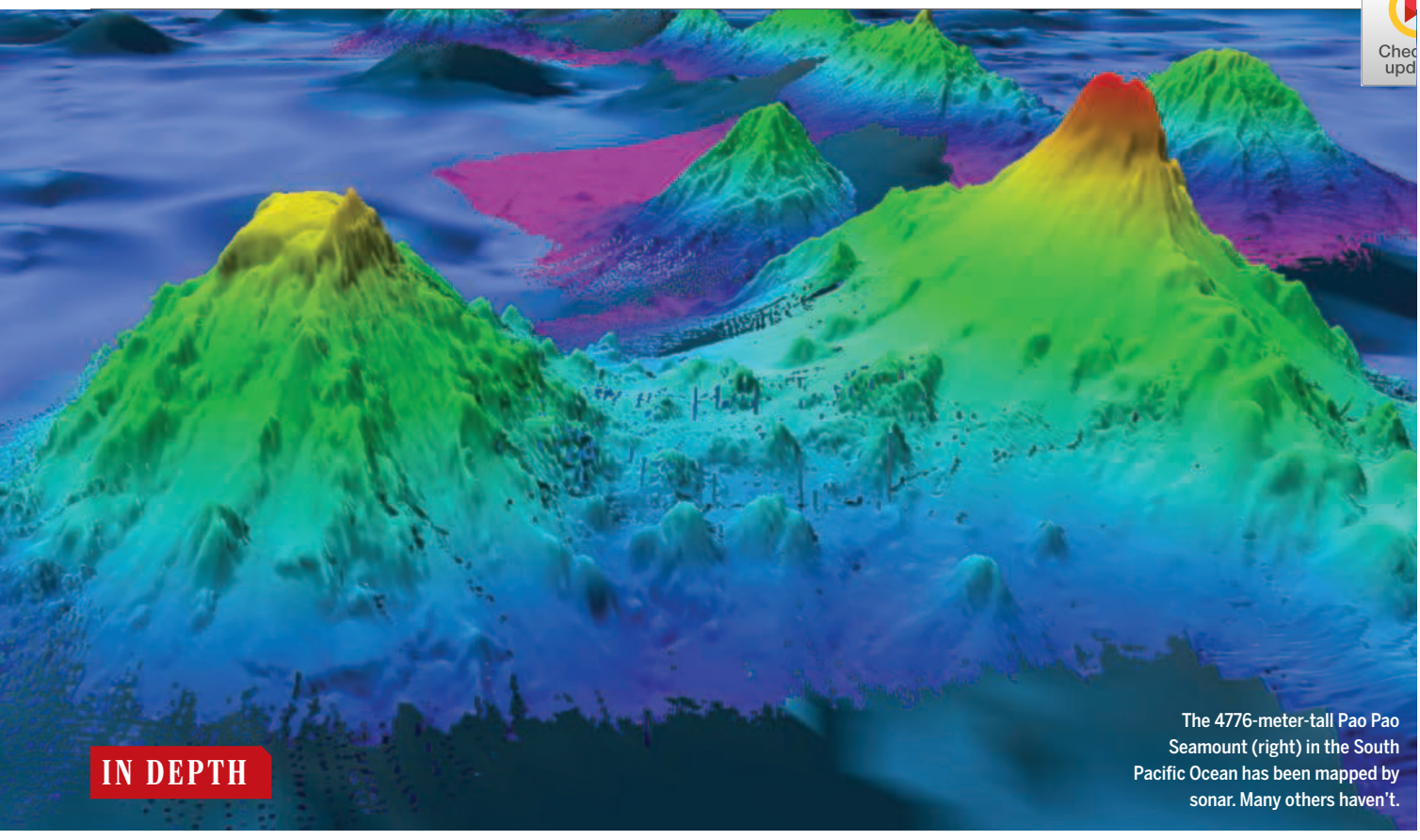
A ceremonial winery for Rome's nobility featured wine fountains and private rooms, perhaps for dining.

Black doctors are Rx for longevity

EQUITY | Black people who reside in U.S. counties with at least one practicing Black physician live longer than those elsewhere—and their longevity increases with the number of Black doctors, a study has found. The analysis is among the first to show that increasing diversity among physicians can help lessen persistent racial disparities in mortality. But researchers from the U.S. Health Resources and Services Administration found that only half of the more than 3000 U.S. counties had even one Black physician in at least one of 3 years included in the study—2009, 2014, and 2019. The findings about longevity may reflect that some Black people seek care from Black doctors based on shared culture, the team wrote last week in *JAMA Network Open*. But medical care should not be segregated, and all physicians should increase their "cultural competency," tailoring care to patients' cultural needs, the authors wrote.

Editors bolt over author fees

PUBLISHING | More than 40 academic editors at *NeuroImage* and an affiliated journal resigned this week to protest the fee that the owner, publishing giant Elsevier, charges to make papers open access. In a statement, they decried as "unethical and unsustainable" *NeuroImage*'s $3450 sum, billed to authors to provide papers free to read upon publication. *NeuroImage* has published all articles open access since 2020. A companion journal, *NeuroImage: Reports*, was introduced in 2021. The departing editors say they asked Elsevier to reduce *NeuroImage*'s fee, and it refused. Now they plan to launch a new, open-access journal—which they hope will "replace *NeuroImage* as the top journal in our field"—to be published by the nonprofit MIT Press. An Elsevier spokesperson says the publisher set *NeuroImage*'s fee below the market average for publications of similar quality.



IN DEPTH

The 4776-meter-tall Pao Pao Seamount (right) in the South Pacific Ocean has been mapped by sonar. Many others haven't.

OCEANOGRAPHY

A 'mind boggling' 19,000 seamounts discovered

New maps of undersea volcanoes could aid in studies of plate tectonics and ocean mixing

By **Paul Voosen**

The U.S. submarine fleet's biggest adversary lately hasn't been Red October. In 2005, the nuclear-powered USS San Francisco collided with an underwater volcano, or seamount, at top speed, killing a crew member and injuring most aboard. It happened again in 2021 when the USS Connecticut struck a seamount in the South China Sea, damaging its sonar array.

With only one-quarter of the sea floor mapped with sonar, it is impossible to know how many seamounts exist. But radar satellites that measure ocean height can also find them, by looking for subtle signs of seawater mounding above a hidden seamount, tugged by its gravity. A 2011 census using the method found more than 24,000. High-resolution radar data have now added more than 19,000 new ones. The vast majority—more than 27,000—remain uncharted by sonar. "It's just mind boggling," says David Sandwell, a marine geophysicist at the Scripps Institution of Oceanography, who helped lead the work.

Published this month in *Earth and Space Science*, the new seamount catalog is "a great step forward," says Larry Mayer, director of the University of New Hampshire's Center for Coastal and Ocean Mapping. Besides posing navigational hazards, the mountains harbor rare-earth minerals that make them commercial targets for deep-sea miners. Their size and distribution hold clues to plate tectonics and magmatism. They are crucial oases for marine life. And they are pot-stirrers that help control the large-scale ocean flows responsible for sequestering vast amounts of heat and carbon dioxide, says John Lowell, chief hydrographer of the National Geospatial-Intelligence Agency (NGA), which runs the U.S. military's satellite mapping efforts. "The better we understand the shape of the sea floor, the better we can prepare [for climate change]."

After the USS San Francisco accident, Sandwell and his colleagues secured funding from the Navy and NGA to hunt for seamounts with satellites. They identified thousands, including 700 particularly shallow ones that posed hazards to submarines. But the team knew its first catalog was far from complete. Now, armed with data from high-resolution radar satellites, including the European Space Agency's CryoSat-2 and SARAL from the Indian and French space agencies, the team can detect seamounts just 1100 meters tall—close to the lower limit of what defines a seamount, Sandwell says.

Seamounts often occur in chains formed as tectonic plates ride over stationary plumes of hot rock rising from the mantle. As a result, the catalog will pay immediate dividends for studies of Earth's interior, says Carmen Gaina, a geophysicist at the Queensland University of Technology. It has already identified new seamounts in the northeast Atlantic Ocean that could help track the evolution of the mantle plume that feeds Iceland's volcanoes. The survey also spotted seamounts near a ridge

in the Indian Ocean where fresh crust is made as tectonic plates spread apart. They suggest a surprising amount of volcanism in a region once thought to be magma starved, Gaina says.

To biologists, seamounts' steep slopes resemble crowded, boisterous skyscrapers for corals and other marine life. "They're oases for biodiversity and biomass," says Amy Baco-Taylor, a deep-sea biologist at Florida State University. Whales use them as waypoints. But biologists debate the role seamounts play in marine biodiversity: Are they home to genetically distinct species, like remote islands? Or do they serve as stepping stones for life to hopscotch through the oceans? By pushing up the density of seamounts, the new maps could strengthen the argument for the latter, Baco-Taylor says.

They will also boost efforts to protect biodiversity in international waters under a new marine protection treaty (*Science*, 10 March, p. 971). "We can't protect the things if we don't know they're there," says Chris Yesson, a marine biologist at the Zoological Society of London's Institute of Zoology. The maps will provide a practical payoff, Yesson adds: "We won't waste our time as much." Some of his colleagues, he says, once traveled to the Indian Ocean to study a seamount that turned out to be a phantom created by an error in presonar depth records.

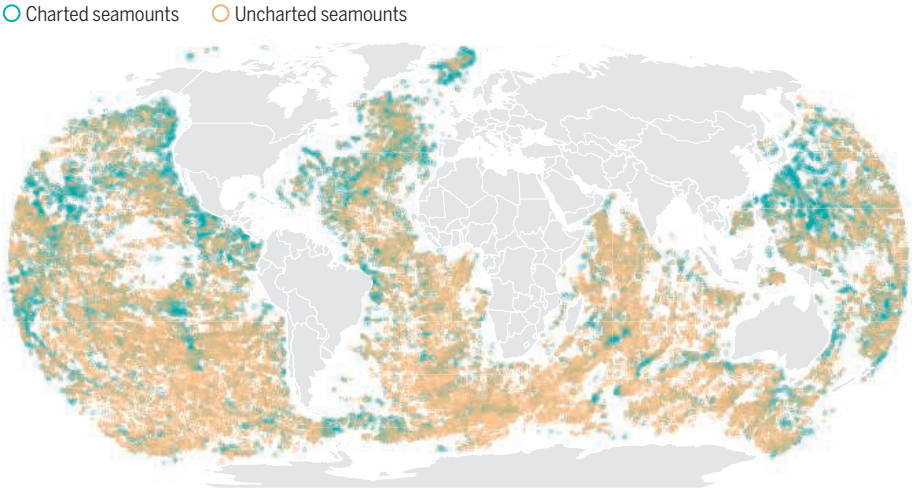
Nowhere will the new maps be as important as in understanding the ocean's globe-girdling conveyor belt of currents. The currents ferry heat from the equator to the poles, where the water cools and gains density until it plunges downward, carrying heat and carbon dioxide into the abyss. But the flip side of this perpetual motion machine—deep ocean waters defying gravity and rising upward—has long been a mystery. The "upwelling" was once thought to happen evenly across the ocean, driven by turbulent waves at boundaries between deep ocean layers of different densities. Now, researchers believe it is concentrated at seamounts and ridges (*Science*, 25 March 2022, p. 1324). "There's a zoo of interesting things that happen when you have topography," says Brian Arbic, a physical oceanographer at the University of Michigan, Ann Arbor.

When ocean currents curl around seamounts, they create turbulent "wake vortices" that can provide the energy to push cold water up, says Jonathan Gula, a physical oceanographer at the University of Western Brittany. In unpublished research, Gula and co-authors have found that these wake vortices make seamounts the leading contributor to upward ocean mixing, and a central player in climate. Since the team relied on the old Scripps catalog, not the new one, the effect of the seamounts is probably even larger, Gula adds.

The seamount catalog is sure to expand further with Seabed 2030, an international project to accelerate high-resolution sonar mapping that Mayer is helping lead. But space surveys will improve too. NASA's Surface Water and Ocean Topography satellite, launched in December 2022, can measure the height of a water surface to within a couple of centimeters. Better remote sensing would be welcome, given the cost of sonar mapping voyages, Mayer says. "I would love to see it threaten what I do." ■

A bumpy ocean bottom

Satellites have detected more than 43,000 seamounts. But only 16,000 have been charted in detail by sonar from ships and submarines.



CREDITS: (GRAPHIC) D. AN-PHAM/*SCIENCE*; (DATA) DAVID SANDWELL

SCIENTIFIC MISCONDUCT

Acclaimed physicist accused of copying thesis

After superconductivity claims, Ranga Dias dogged by plagiarism allegations

By **Daniel Garisto**

In March, University of Rochester (U of R) physicist Ranga Dias made a blockbuster announcement: His team had detected superconductivity at room temperature, in a material that did not need to be squeezed to incredibly high pressures. Many physicists regarded the claim warily because 6 months earlier, *Nature* had retracted a separate room-temperature superconductivity claim from Dias's group, amid allegations of data manipulation.

Now come accusations that Dias plagiarized much of his Ph.D. thesis, completed in 2013 at Washington State University (WSU). *Undark*, *The New York Times*, and *Physics* magazine previously reported that his thesis contains many passages identical to those from a 2007 thesis written by James Hamlin at Washington University in St. Louis. Hamlin, now a high-pressure experimentalist at the University of Florida, and Simon Kimber, a physicist most recently at the University Burgundy Franche-Comté, have gone through the thesis by hand and say they have discovered more widespread examples of copying. In an analysis shared with *Science*, they find Dias's thesis contains at least 6300 words—some 21% of the thesis—that are identical to passages from 17 other sources. Dias's website at U of R also contains text that appears to have been copied without attribution from other sources, Hamlin and Kimber say.

Experts who examined the analysis agree the thesis is heavily plagiarized. "It's obvious," says Lisa Rasmussen, a research ethicist at the University of North Carolina, Charlotte. A U of R spokesperson noted that the plagiarism concerns are largely confined to the methodology and background section. But that doesn't absolve Dias, says Vanja Pupovac, an expert in research integrity at the University of Ri-

in the Indian Ocean where fresh crust is made as tectonic plates spread apart. They suggest a surprising amount of volcanism in a region once thought to be magma starved, Gaina says.

To biologists, seamounts' steep slopes resemble crowded, boisterous skyscrapers for corals and other marine life. "They're oases for biodiversity and biomass," says Amy Baco-Taylor, a deep-sea biologist at Florida State University. Whales use them as waypoints. But biologists debate the role seamounts play in marine biodiversity: Are they home to genetically distinct species, like remote islands? Or do they serve as stepping stones for life to hopscotch through the oceans? By pushing up the density of seamounts, the new maps could strengthen the argument for the latter, Baco-Taylor says.

They will also boost efforts to protect biodiversity in international waters under a new marine protection treaty (*Science*, 10 March, p. 971). "We can't protect the things if we don't know they're there," says Chris Yesson, a marine biologist at the Zoological Society of London's Institute of Zoology. The maps will provide a practical payoff, Yesson adds: "We won't waste our time as much." Some of his colleagues, he says, once traveled to the Indian Ocean to study a seamount that turned out to be a phantom created by an error in presonar depth records.

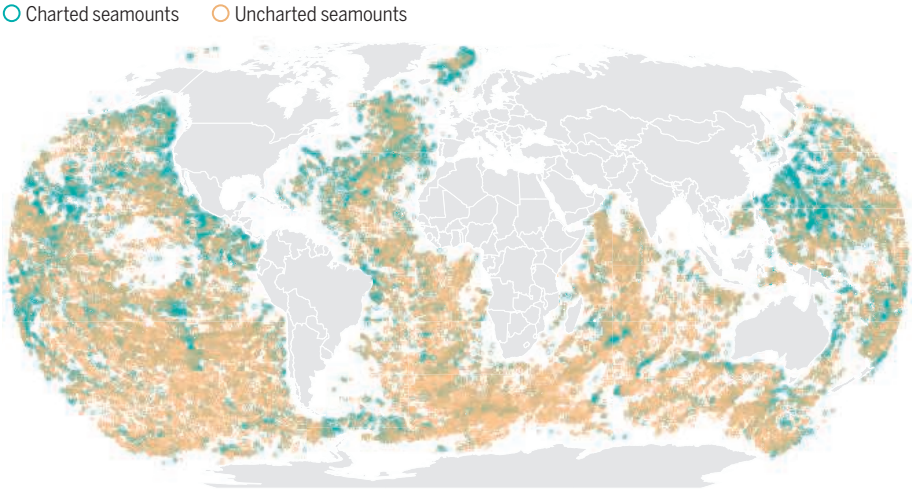
Nowhere will the new maps be as important as in understanding the ocean's globe-girdling conveyor belt of currents. The currents ferry heat from the equator to the poles, where the water cools and gains density until it plunges downward, carrying heat and carbon dioxide into the abyss. But the flip side of this perpetual motion machine—deep ocean waters defying gravity and rising upward—has long been a mystery. The "upwelling" was once thought to happen evenly across the ocean, driven by turbulent waves at boundaries between deep ocean layers of different densities. Now, researchers believe it is concentrated at seamounts and ridges (*Science*, 25 March 2022, p. 1324). "There's a zoo of interesting things that happen when you have topography," says Brian Arbic, a physical oceanographer at the University of Michigan, Ann Arbor.

When ocean currents curl around seamounts, they create turbulent "wake vortices" that can provide the energy to push cold water up, says Jonathan Gula, a physical oceanographer at the University of Western Brittany. In unpublished research, Gula and co-authors have found that these wake vortices make seamounts the leading contributor to upward ocean mixing, and a central player in climate. Since the team relied on the old Scripps catalog, not the new one, the effect of the seamounts is probably even larger, Gula adds.

The seamount catalog is sure to expand further with Seabed 2030, an international project to accelerate high-resolution sonar mapping that Mayer is helping lead. But space surveys will improve too. NASA's Surface Water and Ocean Topography satellite, launched in December 2022, can measure the height of a water surface to within a couple of centimeters. Better remote sensing would be welcome, given the cost of sonar mapping voyages, Mayer says. "I would love to see it threaten what I do." ■

A bumpy ocean bottom

Satellites have detected more than 43,000 seamounts. But only 16,000 have been charted in detail by sonar from ships and submarines.



CREDITS: (GRAPHIC) D. AN-PHAM/*SCIENCE*; (DATA) DAVID SANDWELL

SCIENTIFIC MISCONDUCT

Acclaimed physicist accused of copying thesis

After superconductivity claims, Ranga Dias dogged by plagiarism allegations

By **Daniel Garisto**

In March, University of Rochester (U of R) physicist Ranga Dias made a blockbuster announcement: His team had detected superconductivity at room temperature, in a material that did not need to be squeezed to incredibly high pressures. Many physicists regarded the claim warily because 6 months earlier, *Nature* had retracted a separate room-temperature superconductivity claim from Dias's group, amid allegations of data manipulation.

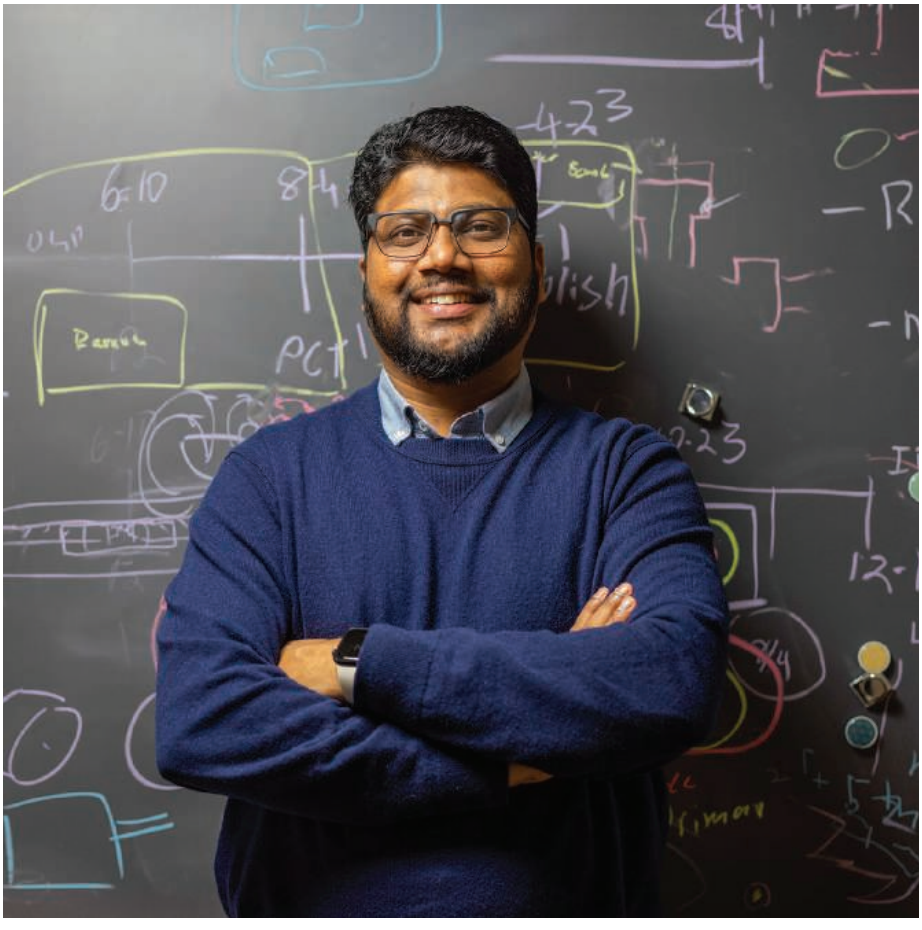
Now come accusations that Dias plagiarized much of his Ph.D. thesis, completed in 2013 at Washington State University (WSU). *Undark*, *The New York Times*, and *Physics* magazine previously reported that his thesis contains many passages identical to those from a 2007 thesis written by James Hamlin at Washington University in St. Louis. Hamlin, now a high-pressure experimentalist at the University of Florida, and Simon Kimber, a physicist most recently at the University Burgundy Franche-Comté, have gone through the thesis by hand and say they have discovered more widespread examples of copying. In an analysis shared with *Science*, they find Dias's thesis contains at least 6300 words—some 21% of the thesis—that are identical to passages from 17 other sources. Dias's website at U of R also contains text that appears to have been copied without attribution from other sources, Hamlin and Kimber say.

Experts who examined the analysis agree the thesis is heavily plagiarized. "It's obvious," says Lisa Rasmussen, a research ethicist at the University of North Carolina, Charlotte. A U of R spokesperson noted that the plagiarism concerns are largely confined to the methodology and background section. But that doesn't absolve Dias, says Vanja Pupovac, an expert in research integrity at the University of Ri-

jeka, in Croatia. "[It] demonstrates, at best, the [Ph.D.] candidate's gross negligence and, at worst, their lack of understanding of the topic," she said in an email.

Through a spokesperson, Dias declined to answer questions and stated he is "addressing these issues directly with his thesis adviser." The U of R spokesperson also declined to comment further until an academic integrity process at WSU was complete. WSU did not respond to a request for comment.

Kimber became suspicious about Dias's work when Hamlin pointed out a detail in a 2021 *Physical Review Letters* (PRL) paper on which Kimber and Dias were co-authors: A plot about manganese sulfide was nearly identical to a plot about a different material, germanium selenide, in Dias's thesis. On 20 March, PRL attached an expression of concern to the paper, which is now under investigation because of questions about the "integrity of the data."



Accusations of scientific misconduct are mounting for University of Rochester physicist Ranga Dias.

Kimber had heard about the similarities between Hamlin's and Dias's theses and wanted to see for himself. He was soon struck by a line about a complex topic that seemed tangential to Dias's thesis:

> *In systems of itinerant fermions the relevant excitations were identified as 'quasiparticles' which are in a one-to-one correspondence with the single-particle momentum eigenstates of a non-interacting system, but which have a modified mass and a finite lifetime due to the quasiparticle interaction.*

"It just stands out so incredibly," Kimber says. Plugging it into Google Scholar, he found it was identical to a sentence from a 1998 paper.

Wondering how much more there was to find, Kimber joined up with Hamlin to take on the arduous task of searching through the rest of the thesis. They checked sentences by hand for similarities to other papers in Google Scholar. The manual effort was required, Kimber says, because automated plagiarism detectors are sometimes inaccurate.

One of the 17 sources Hamlin and Kimber identified is a 1999 paper by Dias' thesis adviser, WSU materials scientist Choong-Shik Yoo. Yoo says he had spotted the apparent duplications himself while reviewing the thesis. "I thought that that was just a simple mistake, so I didn't think of that as a big deal," says Yoo, who adds that his paper is referenced elsewhere in the thesis.

The university's academic honesty policy, however, "makes no distinction between intentional and unintentional plagiarism." According to Yoo, Dias submitted a request to correct his thesis on 30 March and it is now under review with the university's Academic Integrity Hearing Board. Although WSU does not specify the sanctions the board could impose, Rasmussen says plagiarism cases this serious could include revocation of the thesis.

Yoo did not review the files Dias submitted, but says, "I'm pretty sure his corrected thesis has implemented the correct information." But Pupovac says this kind of revision creates serious problems. "Allowing correction of the thesis sends the message that dishonesty can be rectified without significant consequences, which undermines the trust of scientists in the integrity of the academic system, and also the trust of the public in science," Pupovac says.

Asked how the plagiarism allegations, on top of other charges of data manipulation, affect her view of Dias's superconductivity claims, Rasmussen says, "It helps to paint a picture of someone as not really caring about standards. It suggests that they don't mind cutting corners; that they may not have as many original ideas as they're presenting themselves to have."

The apparent plagiarism did not stop with the 2013 thesis. Hamlin and Kimber also found that descriptions of research on Dias's websites, for both U of R and Harvard University, where he completed a postdoc, contain several passages identical to at least three other sources. One page on the site, about research in 2D materials, has sentences that match a description on the website of University of Washington researcher Matthew Yankowitz. "I am essentially certain that the text … was plagiarized from my own," Yankowitz said in an email.

In the meantime, several research groups have failed to replicate Dias's latest superconductivity claim. Faith in the result may be diminishing, but interest in room-temperature superconductivity remains hot. As Dias's Harvard website puts it, "[e]fforts to identify and develop new superconducting materials continue to increase rapidly, motivated by both fundamental science and the prospects for applications."

A 2010 *Nature* paper by a group in Japan started with almost identical words. ■

Daniel Garisto is a science journalist in New York City.

Ukrainian tanks take cover in a pine plantation near Kreminna in February.

PHOTO: SCOTT PETERSON/GETTY IMAGES

ECOLOGY

Scientists plan a comeback for Ukraine's war-ravaged forests

Destruction could open the door to management reforms

By **April Reese**

In addition to its horrific human toll, the war in Ukraine has inflicted widespread damage on the nation's forests. Bombs and missiles have sparked thousands of fires, and "artillery breaks trees in half—it basically mows the forest," says Brian Milakovsky, a U.S.-born forest ecologist who lived in eastern Ukraine before fleeing the country.

Ironically, some forestry experts say the destruction could lead to a major overhaul of how Ukraine manages its forests, changes they say will help ensure these landscapes can better cope with climate change, support biodiversity, and protect water quality. Optimistic that Ukraine will prevail in the war, the researchers are already planning for this greener postwar future. Milakovsky and Sergiy Zibtsev, a forest scientist at the National University of Life and Environmental Sciences of Ukraine, shared their vision during a webinar held last week by the Yale School of the Environment.

"We need to look at solutions that lead to different forest landscapes," says Milakovsky, who continues to work on Ukraine forest issues from his new home in Latvia. "Because the status quo just is really struggling under climate change and war."

Even before the current war, Ukraine's forests were considered some of the world's most damaged. The expansion of agriculture in this major food exporter had vastly reduced forest cover; nearly half of Ukraine is now cropland. In many of the forests that remained, open stands of fire-adapted Scots pine had been replaced by crowded, more fire-prone plantations. The dense stands were encouraged by Soviet-era policies that aimed to "pack as much wood as you can on every hectare," Milakovsky says. But if "fire gets in, it just is death."

Plantations are common in eastern Ukraine, where much of the fiercest fighting is now taking place. Since the Russian invasion began in February 2022, almost 20,000 fires have burned across 755,638 hectares, according to remote sensing data. Farmers accidentally started some fires when they burned fields to clear them, but weaponry ignited many others. Forests also have been damaged by the construction of trenches, bunkers, and roads. Some of the worst damage is along rivers, such as the Siverskyi Donets, that have become crucial lines of defense, Milakovsky says.

Ukrainian law encourages foresters to replant plantations whenever an area is logged or burned. To try a different management regime, they have to get special permission, and few seek it. "Economics, legislation, and habit" enable plantations to persist, Milakovsky said, despite increasing concerns that the monocultures do relatively little to support native species and can suck up scarce water.

Researchers say the war damage presents an opportunity for a long overdue policy shift. The blazes and military activities are breaking up some plantations, for example, opening the door to creating more diverse mosaics of forest types, managed for a mix of restoration and logging, Milakovsky said. That will require political will but could result in more resilient woodlands. Forests with mixed species and well-spaced trees of varying ages would be less susceptible to intense fires and the droughts that are expected to become more common as the region's climate warms, Milakovsky says. "Ukraine is pretty dry, and it's getting drier."

This year, a rainy winter has encouraged growth, priming forests for big burns, says Petro Testov, an ecologist with the Ukrainian Nature Conservation Group. "If next year will be dry, we could see very huge forest fires like in 2020," he says, when blazes tore through pine forests around Luhansk, killing 17 people.

The researchers also hope to build groundwater protection into forest management. Sandy landscapes like those in southeastern Ukraine allow water to soak into the ground and replenish aquifers, but plantations can interfere. When fires thin them, wetlands often reappear and groundwater levels rebound. Shrinking plantations could avoid the "repeated depletion" of these water resources, Milakovsky said.

Zibtsev is planning to soon convene Ukrainian forest scientists—many of whom have relocated to other nations—to discuss how to advance these and other reforms and improve collaboration. Once the conflict ends, he and Milakovsky also hope to resume work the two began before the war with local foresters around the city of Kreminna—now one of the hottest combat zones. The local partners agreed to test alternative management methods, such as allowing low-value areas to naturally regenerate. "They liked those [sustainable forestry] ideas," Zibtsev says, but worried about running afoul of government regulations. "All of them were just like: 'Help us get some kind of [legal] protection.'"

Even as Ukraine's forest scientists look to the future, the war's continuing toll on the forests they have worked in for years is never far from their minds. Copernicus, the European Union's satellite monitoring system, shows constellations of fires across eastern Ukraine, particularly between the cities of Kharkiv and Luhansk.

Still, Zibtsev has little doubt that his vision of a conflict-free Ukraine, rich in sustainably managed forests, will be realized. "We expect within the next 20 years, there will be quite radical changes," he says. "We're in a position to push this agenda. And we already have some progress." ■

April Reese is a journalist in Aveiro, Portugal.

LBNL's fully automated A-Lab can churn out new materials 24/7 without human intervention.

MATERIALS SCIENCE

AI-driven robotics lab joins the hunt for materials breakthroughs

Setup is the first fully automated effort seeking novel inorganic materials for emerging technologies

By **Robert F. Service**, *in San Francisco*

Imagine a cookbook with 150,000 tempting dishes—but few recipes for making them. That's the challenge facing an effort at the Lawrence Berkeley National Laboratory (LBNL) known as the Materials Project. It has used computers to predict some 150,000 new materials that could improve devices such as battery electrodes and catalysts. But the database's users around the globe have managed to make just a fraction of these for testing, leaving thousands untried. "Synthesis has become the bottleneck," says Gerbrand Ceder, a materials scientist at LBNL.

Now, Ceder and his colleagues have married artificial intelligence (AI) and robotics to eliminate that bottleneck. The AI system makes a best guess at a recipe for a desired material and then iterates the reaction conditions as robots try to create physical samples. The new setup, known as the A-Lab, is already synthesizing about 100 times more new materials per day than humans in the lab can manage. "This is the way to go," says Ali Coskun, a chemist at the University of Freiburg who isn't involved with the A-Lab, but attended the Materials Research Society meeting here last week, where the new AI approach was announced.

AI-driven robotics labs are becoming commonplace among pharmaceutical companies searching for new drugs and even some academic materials labs (*Science*, 13 December 2019, p. 1295) But those efforts primarily use liquid precursor compounds that are relatively straightforward to mix and process. "It's a lot more difficult to do this with solid materials," Coskun says. Synthesizing these materials typically requires mixing solid powders together and then adding different combinations of solvents, and experimenting with heat, drying time, and other inputs to try to get them to crystallize into the predicted material.

The number of recipes is essentially infinite, Ceder says. Although computers can predict which final compounds should lead to better devices, "there is no theory for synthesis that tells us what can and cannot be made," says Kristin Persson, who heads LBNL's Materials Project and announced the new A-Lab.

Previous automation efforts randomly mixed compounds in search of new materials, Ceder says, but the new AI-driven approach is more akin to the way traditional chemists do their jobs. The AI starts by coming up with a plausible way to synthesize a material, using its understanding of chemistry. It guides robotic arms to select among nearly 200 different powdery starting materials, containing elements such as lithium, nickel, copper, iron, and manganese. After mixing the precursors, another robot parcels out the mix into a set of crucibles, which are loaded into furnaces where they can be mixed with gases such as nitrogen, oxygen, and hydrogen. The AI then determines how long to bake the different mixes, the temperatures, drying times, and so on.

After the baking, a gumball-like dispenser adds a ball bearing to each crucible and shakes it to grind the new substance into a fine powder that's loaded onto a slide. A robot arm then grabs each sample and slides it into an x-ray machine or other equipment for analysis. Results are fed back into the Materials Project database of materials structures and properties, and if the outcome isn't what was predicted, the AI setup iterates the reaction conditions and starts anew.

LBNL researchers have spent the past several months working out the kinks in their system and testing it. In the process, the A-Lab has produced more than 40 target materials—about 70% of the compounds it has set out to produce. "I have made more new compounds in the last 6 weeks than my whole career," Ceder says.

LBNL's AI materials lab may not be alone for long. In a 3 April preprint, researchers from the Samsung Advanced Institute of Technology reported that they, too, have set up a computer-driven robotics lab to search for new electronic materials. Results from that report show their setup performed more than 200 reactions to make 35 inorganic compounds, including certain oxides commonly used in battery electrodes, solid oxide fuel cells, and superconductors. In each stage of their robotic experiments "AI is used to some degree," says Samsung's Jeong-Ju Cho.

Ceder notes that despite the move to fully automated synthesis and analysis, researchers are just as likely as ever to make unexpected discoveries. "That's no different with the A-Lab." Except now, the hits and the surprises will likely come faster. ■

SYNTHETIC BIOLOGY

Yeast are engineered to thrive on light

Experiment shows ease by which organisms could harness sunlight to produce energy

By **Elizabeth Pennisi**

Yeast are carb lovers, sustaining themselves by fermenting sugars and starches from sources such as dough, grapes, and grains, with bread, wine, and beer as happy byproducts. Now, researchers have made one type of yeast a little less dependent on carbs by enabling it to use light as energy.

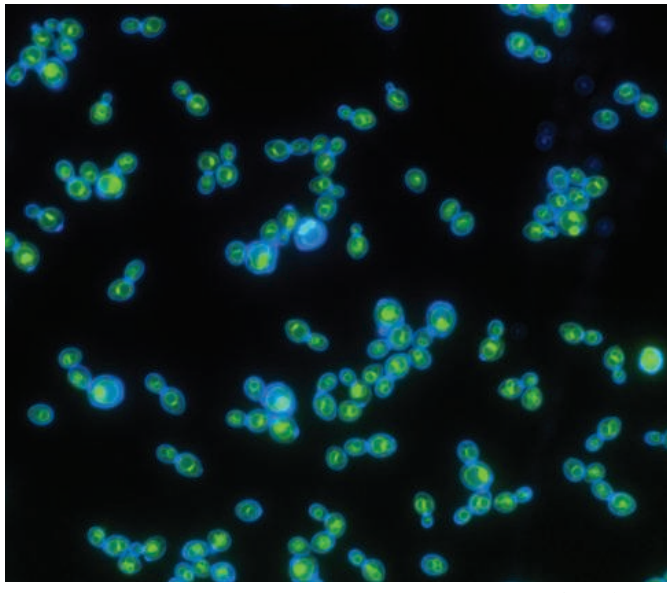
The work, reported last week on the preprint server bioRxiv, is "the first step in more complex modes of engineering artificial photosynthesis," says Magdalena Rose Osburn, a geobiologist at Northwestern University who was not involved in the research. It also recapitulates a key evolutionary transition—the harnessing of light. "It is extraordinary," says Felipe Santiago-Tirado, a fungal cell biologist at the University of Notre Dame. "To some extent, it's like turning an animal into a plant."

Well, not quite. To convert carbon dioxide into sugars that fuel life on Earth, plants rely on a protein complex that includes chlorophyll to shuttle both electrons and protons, which perform chemical reactions and transfer energy. Researchers have been working for years to recreate photosynthesis to explore how to use light more efficiently as an energy source for solar panels and other applications and to breed plants—and other organisms—to be more productive.

But the chlorophyll complex requires many other molecules to do its job. So Anthony Burnetti, a geneticist at the Georgia Institute of Technology, and Georgia Tech evolutionary biologist William Ratcliff sought a simpler solution. They homed in on a protein known as rhodopsin, which doesn't require a large molecular entourage. It's a solution nature has settled on as well: Bacteria, some protists, marine algae, and even algal viruses use rhodopsin to convert light into usable energy, often to pump protons for cellular functions.

The researchers began by inserting a rhodopsin gene that belonged to a marine bacterium into brewer's yeast (*Saccharomyces cerevisiae*) in a petri dish. Burnetti hoped the rhodopsin would find its way into the yeast's vacuole, an enzyme-laden sac that degrades unneeded proteins. An energy molecule called adenosine triphosphate (ATP) fuels the process by pumping protons into the vacuole to make its interior acidic—optimal for degradation.

Burnetti wondered whether light energy could do that job instead. But the team's first effort misfired when the rhodopsin protein made by the gene went to a different compartment known not for protein degradation, but for protein synthesis. So Burnetti looked instead for rhodopsin already known to exist in vacuoles. He settled on using one from corn smut, a fungal pathogen. By attaching a green fluorescent tag to the protein, he and his colleagues verified that it had localized to the yeast's vacuole, as they hoped.



The blue cell walls of individual yeast cells surround rhodopsin (green), which helps those cells grow faster.

PHOTO: ANTHONY BURNETTI

Graduate student Autumn Peterson, a member of Burnetti's team, went a step further to prove this engineered yeast was indeed using light. She grew the new strain in the same dish as the original, unaltered yeast and exposed it to green light, the wavelength rhodopsin is most sensitive to. The cells in the light-sensing strain had shorter lives but reproduced fast enough to outgrow the nonlight sensing yeast by 0.8%, the team found. That's a "massive advantage," says Santiago-Tirado. Over time, in the light, Peterson expects the light-using cells to eventually replace the unaltered ones just as early light users might have replaced their competitors in nature eons ago.

Burnetti and his colleagues think light induces the rhodopsin to pump more protons into the vacuole, relieving the cells' need to expend ATP for this task and instead freeing up that energy to help the cell grow in other ways. Increasing the acidity inside the vacuole may decrease it outside the vacuole, causing enzymes there to work faster and wear out sooner, which may also help explain the higher death rate among these altered cells. Whichever way it's working, "It is clearly of benefit to the yeast cells," says Michael McMurray, a molecular biologist at the University of Colorado Anschutz Medical Campus.

But the experiment may not reveal much about how rhodopsin use evolved in nature. "I think the authors overemphasize the evolutionary significance of their work," says Robert Blankenship, an emeritus biochemist at Washington University in St. Louis. "This is an artificial construct and is not the product of natural evolution."

Others think the work can have industrial, medical, and basic research applications. Alaattin Kaya, a biologist who studies aging at Virginia Commonwealth University, says these yeast cells can help clarify why vacuole acidification over the life of a cell sometimes seems to cause mitochondria to malfunction and in turn accelerate aging. He would love to add rhodopsin to mitochondria themselves to observe its impact.

Burnetti would like to target mitochondria as well, but for a different reason. "Even though it seems to have never happened in nature, we definitely plan to eventually put rhodopsin into the mitochondrion." Because mitochondria can make ATP efficiently, adding rhodopsin could provide a lot of energy directly from the Sun, just as photosynthesis does. In that regard, yeast would then be a little more like plants. ■

CANCER IMMUNOTHERAPY

Tailored melanoma vaccine may stave off cancer

Small trial offers first clinical evidence that supports a personalized vaccine approach

By **Jocelyn Kaiser**

A novel cancer vaccine tailored to genetic changes in a person's tumor is showing promise in the clinic. In a study of about 150 people who had surgery for melanoma, a type of skin cancer, those given a personalized vaccine along with an immunotherapy drug were more likely to remain free of cancer 18 months later than patients who did not receive the vaccine.

The results, reported this week at the annual meeting of the American Association for Cancer Research (AACR), offer the first clear evidence that a vaccine designed to target mutations within a patient's tumor can prevent its regrowth. That would be a milestone for the cancer vaccine field, which has struggled for decades to show results. It could also add to a growing arsenal of drugs, known as immunotherapies, that harness the immune system to fight cancer. "I was really, really excited to see these data," says Patrick Ott of the Dana-Farber Cancer Institute, who works on similar vaccines. Although small, the new study is "a very exciting first step," says cancer vaccine researcher Nina Bhardwaj of the Icahn School of Medicine at Mount Sinai.

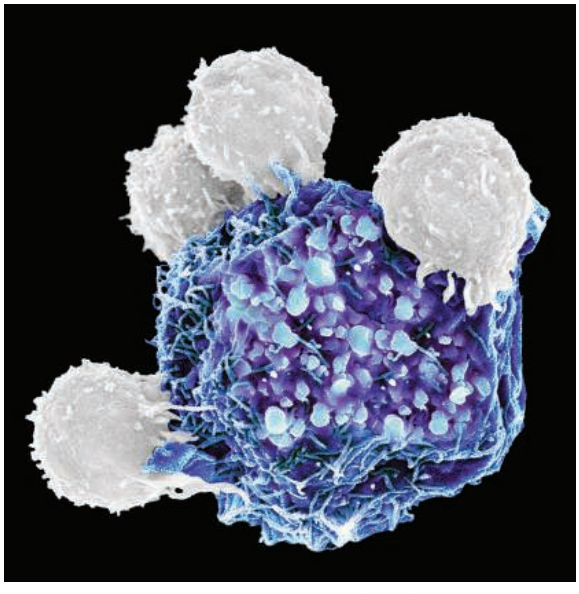
Cancer vaccines aim to teach the immune system's T cells to attack a tumor by exposing them to a protein, or antigen, that pokes out from a cancer cell. But most vaccines so far haven't worked well because the same antigens found on tumors also appear on normal cells.

In the early 2010s, as DNA sequencing costs dropped, some scientists turned instead to sequencing the mutations in a patient's tumor, then creating a vaccine to deliver a few of the corresponding mutated proteins, known as neoantigens, which are found only on the tumor cells. Several small trials published since 2015 by Ott's team and others have shown that neoantigen vaccines can stimulate vaccine-specific T cells in patients with solid tumors such as melanoma, colon, lung, and brain cancer, and at least in melanoma, may curb cancer growth.

To show this more definitively, Merck and Moderna conducted a randomized trial for patients who had advanced melanoma that had spread to lymph nodes and sometimes other sites, but that had been surgically removed. All got a type of drug, known as a checkpoint inhibitor, that blocks a crucial protein from enabling tumors to evade T cells. Two-thirds also got vaccine infusions every 3 weeks for about 4 months. Like Moderna's COVID-19 vaccine, the cancer vaccine delivered messenger RNA (mRNA) wrapped in lipid nanoparticles into cells, instructing them to make proteins—in this case, up to 34 tumor neoantigens per patient.

In December 2022, the companies made a splash when they reported that patients receiving the vaccine were 44% less likely to die or have a recurrence of their cancer. At the AACR meeting, academic collaborators shared more details: Eighty-four of the 107, or 79%, were still in remission after 18 months, compared with only 31 of 50 (62%) patients who got the checkpoint inhibitor alone. "These data give a very, very encouraging signal," says Jeffrey Weber of NYU Langone's Perlmutter Cancer Center, the trial's principal investigator.

Also encouraging is that the vaccine worked regardless of how many mutations the person's melanoma tumor had, suggesting it could work for cancer types with fewer mutations. With less to distinguish them from normal cells, such cancers tend to resist immunotherapy drugs. A larger study starting later this year aims to confirm these results and reveal whether the vaccine extends patients' lives, measures that could encourage regulators to approve it. For now, "these [are] intriguing early findings," says immunotherapy researcher Suzanne Topalian of Johns Hopkins University. Like other researchers, she hopes to see more details, including evidence that patients who did well made T cells specific to the neoantigens and didn't just get an immune boost from the vaccine's nanoparticles. Weber says those data will be reported in papers the team is submitting to journals.



Tailored cancer vaccines help train T cells (white) to attack tumors.

Other companies are also testing neoantigen vaccines in randomized trials. BioNTech and Genentech expect to report early results this year for a neoantigen mRNA vaccine for metastatic melanoma that can't be surgically removed—a tougher challenge partly because the patients have weakened immune systems. And Gritstone bio is testing a neoantigen mRNA vaccine against metastatic colon cancer; to boost the immune response, it is combined with a modified virus carrying the neoantigens. The Gritstone team reported in *Nature Medicine* in August 2022 that in several cancer patients, this resulted in "very significant numbers of T cells," according to Bhardwaj, a promising sign of efficacy.

One of the most intriguing studies so far tested a BioNTech and Genentech neoantigen vaccine for pancreatic cancer. Investigators reported last summer that eight of 16 patients in a trial had T cell responses to the vaccine and were still cancer-free up to 2.5 years later. The other eight did not show an immune response and six had relapsed by 18 months. The companies plan to launch a randomized trial of that vaccine for pancreatic cancer this year.

Because pancreatic cancers have few mutations, "you might think this is the last tumor type" that a neoantigen vaccine would work for, says principal investigator Vinod Balachandran of Memorial Sloan Kettering Cancer Center, who presented full details at AACR and whose team has a paper in press. "If you can even do this in pancreatic cancer, this is very encouraging for testing personalized vaccines" for other cancers. ■

FEATURES

TRUST BUT VERIFY

U.S. labs are overhauling the nuclear stockpile. Can they validate the weapons without bomb tests?

A Cold War stalwart, the Titan II missile (seen in an Arizona museum) carried warheads that had been tested. Weapons must now be certified without tests.

By **Sarah Scoles,**
in Los Alamos, New Mexico

Behind a guard shack and warning signs on the sprawling campus of Los Alamos National Laboratory is a forested spot where scientists mimic the first moments of a nuclear detonation. Here, in the Dual-Axis Radiographic Hydrodynamic Test (DARHT) facility, they blow up models of the bowling ball–size spheres of plutonium, or "pits," at the heart of bombs—and take x-ray pictures of the results.

In a real weapon, conventional explosives ringing an actual pit would implode the plutonium to a critical density, triggering an explosive fissile chain reaction. Its energy would drive the fusion of hydrogen isotopes in the weapon's second stage, generating yet more neutrons that would split additional fission fuel.

This fission-fusion-fission process eats up some of the atoms' mass and, according to $E=mc^2$, Albert Einstein's famous equation, releases ferocious amounts of energy. That's

why a warhead about 1 meter long can explode with the force of a megaton of TNT. If dropped on a city like Washington, D.C., it would instantly vaporize an area more than 2.5 kilometers across while crumpling buildings much farther out with its radioactive blast. It would kill nearly half a million people and injure or sicken almost as many.

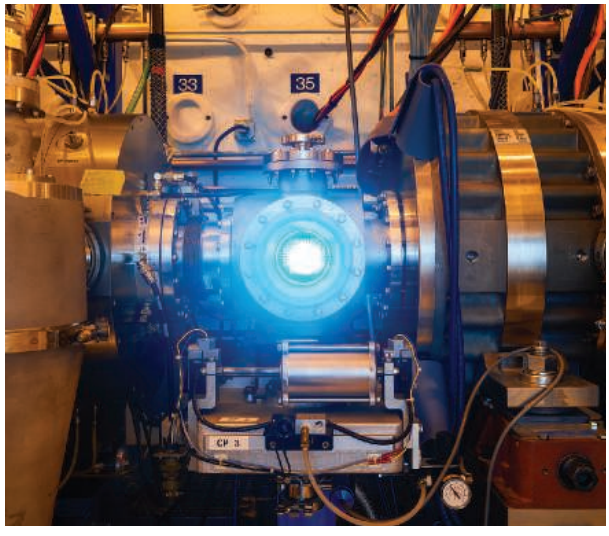
DARHT's experiments take place within a steel vessel shaped like a diving bell. The mock pits, made of dense metals such as lead, tantalum, or depleted uranium, have properties similar to plutonium—minus its tendency to fission. As the explosive charges are detonated, two perpendicular beams of x-rays document the pit's implosion like high-speed cameras. Weapons scientists compare those pictures with classified supercomputer simulations of the bomb blasts to see how well the real and digital worlds match.

Facilities like DARHT have been important since 1992, when the Department of Energy's (DOE's) three weapons labs—Los Alamos, Lawrence Livermore National Laboratory, and Sandia National Laboratory—stopped full-fledged tests of nuclear weapons. By 1996, the United States had signed the Comprehensive Nuclear-Test-Ban Treaty—credited not only with stopping the environmental damage of nuclear testing, but also with disincentivizing new weapons designs.

Without tests, however, the only things ensuring that warheads work are facilities like DARHT, computer simulations from "weapons codes," and a cache of data from the old days of nuclear testing. For relatively minor changes to old weapons—new fuses, fresh top-ups of the hydrogen isotope tritium—that has been enough. Every year, DOE's National Nuclear Security Administration (NNSA) and the Department of Defense have certified the stockpile, an assessment that means they are convinced the weapons will work when they're supposed to, as they're supposed to—and not do anything when they're not supposed to. "Because we've blown up so many of them, these things are incredibly reliable," says Geoff Wilson, director of the Center for Defense Information at the Project on Government Oversight, which argues nuclear weapons spending should be reduced.

But now the stockpile is getting an overhaul, the biggest in decades. This fiscal year, NNSA has a record $22.2 billion budget. Much of the money will go to producing new plutonium pits to replace those in the arsenal and to modernizing four warheads. A fifth weapon, dubbed the W93—a submarine-launched warhead—is a new design program. "It's really the first warhead program we've had since the end of the Cold War" that isn't a life extension or modernization of an existing weapon, says Marvin Adams, NNSA's deputy administrator for defense programs.

The work has become more urgent, with the post-Cold War calm turning stormy again. Russia has backed out of its only remaining major arms-control treaty with the United States, while making regular nuclear threats during its invasion of Ukraine. China is thought to be expanding its stockpile, while Iran and North Korea continue to bolster nuclear programs. "Everybody went to sleep for 25 years," says Charlie Nakhleh, Los Alamos's head of weapons physics. "I think we're awake now."



At one Los Alamos facility, x-ray beams are used to image imploding mock "pits," the spheres of plutonium at the heart of nuclear weapons.

Wilson worries that the international dynamics and the U.S. overhaul could ultimately lead to a revival of bomb tests, bringing back their hazards and stoking a new arms race. "It is not unfathomable to me, which is scary to say." It's one thing to tweak weapons with a deep heritage. It's another to infer functionality for modified weapons that have never been fully tested, he says.

Weapons physicists at the labs are confident they can improve existing weapons and design new ones without tests. Their computer simulations are vastly superior to those of the past, and experiments like DAHRT's are more powerful. "Would you design a new Formula One car without taking it on the track? Or would you design a new Boeing jetliner without flying at first?" asks Rob Neely, Livermore's program director for weapons simulation and computing. In the case of nuclear weapons and their plutonium pits, he says, the answer appears to be, "Actually, yes."

As the simulations and experiments have improved, they've also revealed gaps in nuclear knowledge, and approximations in the codes that haven't been updated in decades. Despite the doubts, Neely brims with confidence. "Not only will these things work, but they're going to work better."

SIMPLY REPLACING the bombs' plutonium pits poses a science challenge: understanding how subtle changes affect their behavior. They aren't easy to make, in part because plutonium, a metal only in existence since 1940, is mysterious and hard to handle. The last time anyone made pits at scale—in the 1980s at Colorado's Rocky Flats plant—DOE's contractor was shut down for environmental violations and forced to pay an $18.5 million fine.

This time, NNSA is splitting production between Los Alamos and the Savannah River Site in South Carolina. It has tasked them with making 80 new pits per year by 2030, a deadline NNSA admits it will not meet.

Los Alamos's pits will be made at a facility called PF-4, a set of high-security buildings surrounded by cyclone fences with razor wire. Inside PF-4 are glovebox enclosures—radiation-shielded workstations where workers use thick gloves and peer through glass windows to manipulate the exotic metal. The lab is hiring thousands of workers, and its first pit is likely to be ready for the stockpile next year.

The gargantuan effort is motivated by a simple fact: many current pits are more than 40 years old, and plutonium behaves in confounding ways as it ages and radioactively decays. A green, fuzzy coating grows on it as its surface oxidizes. Atoms in its metallic lattice are knocked out of place as it spits out uranium isotopes. Its dimensions shift when it slips between six different solid phases. And the pits do not necessarily degrade smoothly. "We know at some point there will be a nonlinear piece," says David Clark, director of Los Alamos's National Security Education Center and editor of the *Plutonium Handbook*. "We just haven't seen it."

So far, the silvery spheres seem to be holding up. Internal and external assessments have vouched for their integrity, suggesting the pits could have decades of viability left. "We haven't seen any issues," Clark says.

But Jason, a secretive group of physicists who advise the government on national security matters, raised concerns that galvanized DOE. In a 2019 report, the group urged the agency to reestablish pit production "as expeditiously as possible" to "mitigate

PHOTO: MICHAEL PIERCE/ LOS ALAMOS NATIONAL LABORATORY

against potential risks posed by Pu aging."

One might think the new pits would make it easier to certify the stockpile, by avoiding the uncertainties of aging plutonium. But they come with uncertainties of their own. The new pits won't be twins of their predecessors, so weapons scientists will have to understand how the alterations change pit behavior. They are being manufactured using recycled and purified plutonium from old pits, not fresh material, unlike the originals. Moreover, they will be made with different processes, and in some cases designed to slightly different specifications. "If you look at a new requirement," Adams says, "you often will find that the old pits we have available to us are really, really suboptimal."

IN SOME WAYS, understanding the behavior of nuclear weapons has grown harder as scientists have gotten better at their jobs. The higher quality simulations enabled by ever more powerful supercomputers, for instance, have sometimes revealed new problems. This was the case with "boost physics," or the processes at work in the first stage of a thermonuclear bomb, where fissioning plutonium triggers fusion reactions in a deuterium-tritium booster, which releases neutrons that spark more fission in the weapon's pit.

For a long time, the simulations couldn't reproduce what physicists saw in data from underground nuclear tests without the application of digital fudge factors. In 2006, scientists increased the simulations' resolution. "And, lo and behold, we found obviously some interesting things that got a bunch of people scratching their heads," Neely says. That helped spawn years of research in a program called the National Boost Initiative, which aimed to understand the fundamental physics of thermonuclear burn and to incorporate more basic physics into simulations, rather than relying on calibrations and approximations.

Pesky approximations rear their fuzzy heads throughout the weapons codes, Neely says. One is inherent to the nature of the simulations. They are all "meshed"—simulated in gridlike parts, like pixels in a digital image. Within each mesh element, physical properties are assumed to be the same. The mesh is getting more refined, but it's still not a precise representation of reality. "You're just able to capture better approximations," Neely says, "but still an approximation."

There's also fuzziness in the physics that governs the simulations. To make the simulations run more efficiently, scientists often rely on math tricks and approximations rather than explicit, first-principles solutions.

Christopher Fryer, head of Los Alamos's Center for Nonlinear Studies, has found

Weapon of mass destruction

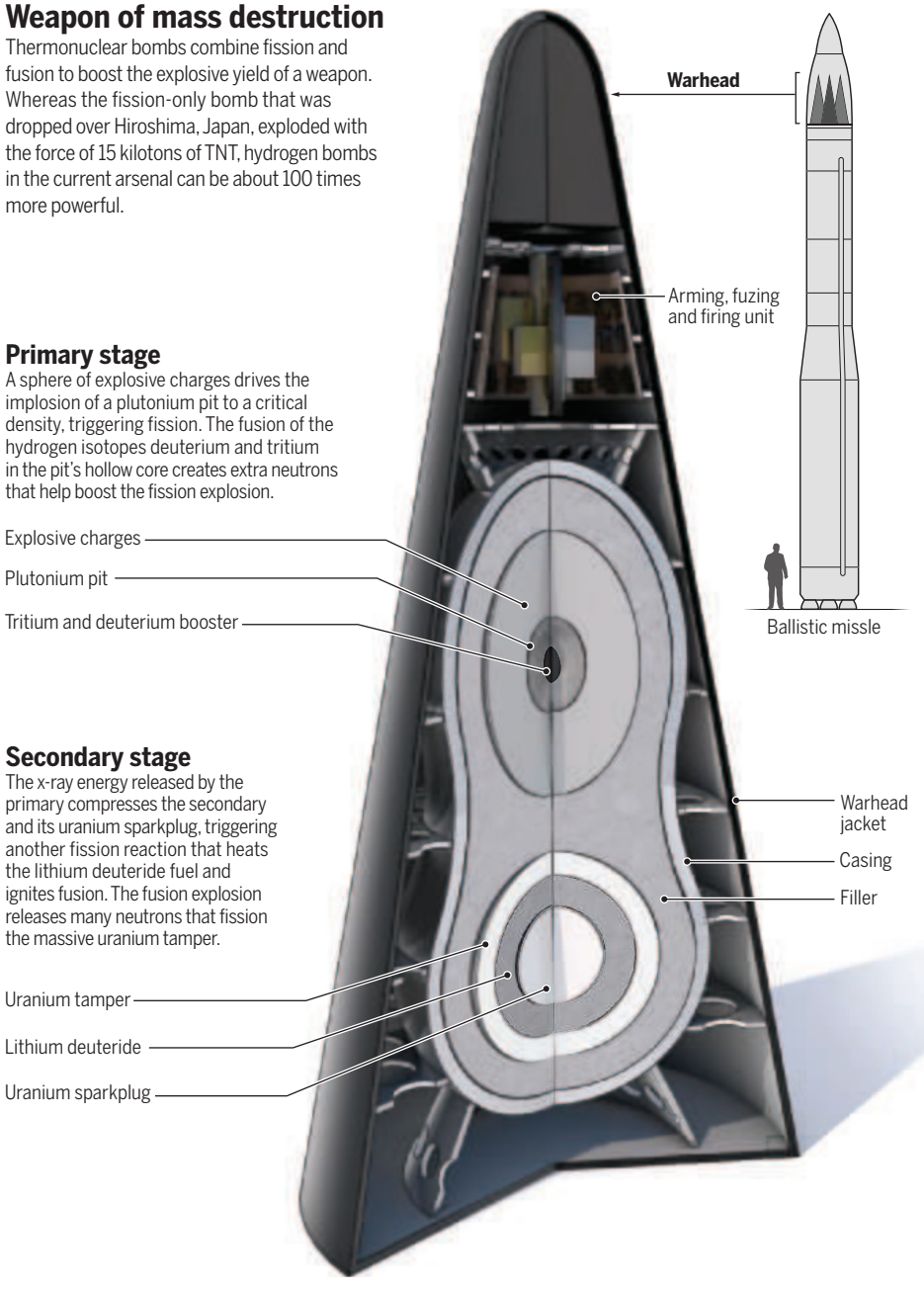
Thermonuclear bombs combine fission and fusion to boost the explosive yield of a weapon. Whereas the fission-only bomb that was dropped over Hiroshima, Japan, exploded with the force of 15 kilotons of TNT, hydrogen bombs in the current arsenal can be about 100 times more powerful.

Primary stage

A sphere of explosive charges drives the implosion of a plutonium pit to a critical density, triggering fission. The fusion of the hydrogen isotopes deuterium and tritium in the pit's hollow core creates extra neutrons that help boost the fission explosion.

Secondary stage

The x-ray energy released by the primary compresses the secondary and its uranium sparkplug, triggering another fission reaction that heats the lithium deuteride fuel and ignites fusion. The fusion explosion releases many neutrons that fission the massive uranium tamper.



Boom times

The three U.S. Department of Energy weapons laboratories are getting billions of dollars to upgrade four weapons. A new design program, the W-93, could end up fielding the first new weapon since 1988.

NAME	UPGRADE
W93	Will be put in service by 2040 and launched from submarines
W88-Alt-370	An alteration to submarine-launched W88 weapons will replace fuze assemblies, add a lightning protector, and replace the conventional explosives.
W87-1	A replacement to the land-launched W78, the W87-1 will have enhanced safety features and use insensitive explosives.
W80-4	This weapon will extend the life of the air-launched W80-1. It was engineered with the Air Force, which designs the delivery systems.
B61-12	A replacement for all four variants of the air-dropped B61 will have maneuverable fins, enabling better targeting that will allow designers to reduce the yield.

that the weapons codes still contain computational tricks conjured up decades ago by Manhattan Project luminaries such as Hans Bethe and Richard Feynman. "Instead of relying on them to be the clever people, we're going to have to be clever again," he says.

One of Bethe's recipes, still stirred into some fusion simulations, involves the movement of charged particles. The recipe assumes that if a particle travels X distance, it loses Y energy—a kind of average scattering that isn't always realistic, particularly in reactions that happen quickly. "It fits the data so well until you find out it doesn't," Fryer says. Replacing it could mean simulating each particle and its particulars—too tough a task even for the latest supercomputers. "This is why we haven't done it," Fryer says.

But other approximating physics could be replaced by better or truer formulas. Los Alamos theoretical physicist Mark Paris is working on a numerical approach to solving the nonlinear differential equations that pulse throughout the weapons codes. "You're actually solving the system of equations that govern the system," he says, "not the pastiche of physical mechanisms that are approximately derived from the system of equations."

Simulation can't be the only tool used to understand the bombs, however. All humans, even weapons physicists, are storytellers, Nakhleh says, and the simulations help them create confident narratives. But that can only go so far. "At some point," Nakhleh says, "you have to step into the unknown—walk into the dark room, and see, 'What did the experiment have to say?'"

thimble-size target containing hydrogen isotopes to spark tiny fusion explosions. NIF creates temperatures and pressures that don't exist anywhere else on Earth, says Laura Berzak Hopkins, associate program director for integrated weapons science at Livermore. "These conditions are those of astrophysical bodies—the center of Jupiter, the core of the Sun," she says.

NIF drew headlines in 2022 when it produced more energy from the thimble than the lasers put in, a milestone relevant to civilian efforts to generate fusion power. But the achievement came about a decade and billions of dollars later than scientists first expected. And they still can't accurately predict how much energy they'll get out of a given fusion shot. "We don't know the physics," Fryer says. That physics is important for understanding the fusion components of the left "no doubt in the minds of our adversaries or allies," Adams says. Without tests, the United States has to signal confidence in a quieter way. "If you can prove to them you understand this physics well enough, you're not bluffing," Fryer says.

That's one reason why the national labs also work on unclassified, fundamental science that overlaps with weapons science, subjects like star formation and supernovae. Lab scientists can publish that work, talk about it, stick it on a poster at an international conference. In that sense, Fryer says, Los Alamos's Center for Theoretical Astrophysics "is a deterrent."

Some still doubt, though, whether that physics-based storytelling will continue to adequately substitute for testing as the weapons overhaul progresses. "It's all well and good for the engineers to go, 'Boom, here's this new

The Trinity supercomputer at Los Alamos National Laboratory simulates weapons explosions at high resolution.

THAT IS THE POINT of expensive, high-powered efforts like DARHT. To help illuminate the inner workings of bomb primaries, Los Alamos wants to increase the number of DARHT tests per year, currently seven, and improve its imaging abilities so it can take more x-ray pictures during any given test.

A second facility, an underground complex in Nevada called U1A, is also being revamped. It will soon be home to the Enhanced Capabilities for Subcritical Experiments (ECSE), a setup in which scientists *will* implode real plutonium, tiptoeing toward a chain reaction without actually triggering one. In ECSE, scientists will take x-ray pictures of scale-model pits as they collapse and investigate how neutrons behave during those crucial instants before the nuclear detonation. Because there is no nuclear explosive yield, such experiments technically adhere to the test-ban treaty.

Perhaps the most famous experimental site is Livermore's National Ignition Facility (NIF), which focuses 192 laser beams onto a

weapons—and also for how they would themselves hold up to a nuclear blast, a branch of research dubbed "weapon survivability."

NO MATTER HOW GOOD the combination of theory, simulation, and experiment gets, it will probably never fully represent what happens in a weapon, Nakhleh says. "Omniscience is going to be a ways away," he says. "The idea is to push that boundary of knowledge as far as possible."

That knowledge isn't only important for maintaining an arsenal. It's also important for broadcasting to the world that the country *knows* the weapons will work. Nuclear deterrence—the idea that one country can prevent attacks by threatening an attack of similar magnitude—only holds up if the other country actually finds your threat credible.

In the era of explosive nuclear testing, conveying that message was simple. Other countries could pick up the seismic signal from an Earth-shaking blast half a world away. That warhead. We're super sure it works,'" Wilson says. That might not be enough certainty for the military. At a certain point, Wilson says, someone might say that "a cheaper way to do this would be 'Let's just blow one up.'"

And there's some political appetite for testing: Senator Tom Cotton (R-AR), for instance, has suggested the country withdraw from the test-ban treaty. In 2020, he proposed an amendment to the National Defense Authorization Act that would provide funding to prepare for potential nuclear tests. It passed the Senate, but the House of Representatives's version of the bill prohibited such spending.

Holding off any push for testing motivates Fryer to dig deeper into the physics, he says. "For me, it comes down to 'I don't want to resume testing,'" he says. If the alternative is understanding the physics better, so be it. ■

Sarah Scoles is a journalist in southern Colorado and author of a forthcoming book about the 21st century nuclear complex.

PHOTO: LOS ALAMOS NATIONAL LABORATORY

ECOLOGY

Could you spare an acre for conservation?

Private landowners must engage in biodiversity conservation

By **Ricardo B. Machado** and **Ludmilla M. S. Aguiar**

Biodiversity, including species, ecosystems, and ecosystem services, is in rapid decline worldwide (*1*), with severe consequences for human populations (*2*). Several countries have established protected areas, the backbone for conserving species and habitats (*3*). However, who should oversee the conservation of biodiversity? The government? The private sector? Local communities? All of them? On page 298 of this issue, De Marco *et al.* (*4*) present an analysis of the Brazilian Cerrado—a vast tropical savanna in central Brazil—that suggests that sharing this responsibility with the private sector could rapidly increase international commitments to avoid biodiversity loss.

Protected areas across the world belong to different management categories (such as national parks, wilderness areas, and managed resource protected areas) that cover ~15.8% of all terrestrial and inland waters and ~8.16% of marine environments (*5*). These numbers are far from the 30×30 protection target (30% of the planet's surface protected by 2030) that was proposed during the 15th meeting of the Conference of the Parties to the United Nations Convention on Biological Diversity held in Montreal, Canada, in December 2022. To achieve this target, the 193 member countries of the Convention should add ~2.685 billion ha to the protected-area system. It means that public and private areas must be considered (*6*). The role of private lands can be substantial in regions where private ownership is dominant, such as the Brazilian Cerrado.

De Marco *et al.* analyzed data on natural vegetation remnants and the potential distribution of 290 threatened vertebrates (except fish) across the Brazilian Cerrado, as well

Department of Zoology, Institute of Biological Sciences, University of Brasilia, Brasilia, 70910-900 DF, Brazil. Email: rbmac@unb.br

An agricultural field is positioned next to a native Cerrado region in Formosa do Rio Preto, Brazil (2019). Changes to Brazilian laws about the cultivation of private land could help reduce the loss of natural vegetation in the Cerrado.

as the legal requirements in Brazil regarding the protection of native vegetation. They conclude that private areas could hold 15.5 to 25% of the species present in the region. The role of private land in protecting the Cerrado's biodiversity could be made more substantial, especially by restoring degraded areas. The study suggests that private landowners could keep 15 million to 20 million ha of degraded areas located in private areas covered by native vegetation. Such areas were excessively occupied by human populations, despite the legal obligation of private landowners to maintain 20 to 35% of their property covered by native vegetation (*7*–*9*). Therefore, restoring degraded areas in private lands and halting devegetation are two necessary measures to promote biodiversity conservation outside public protected areas.

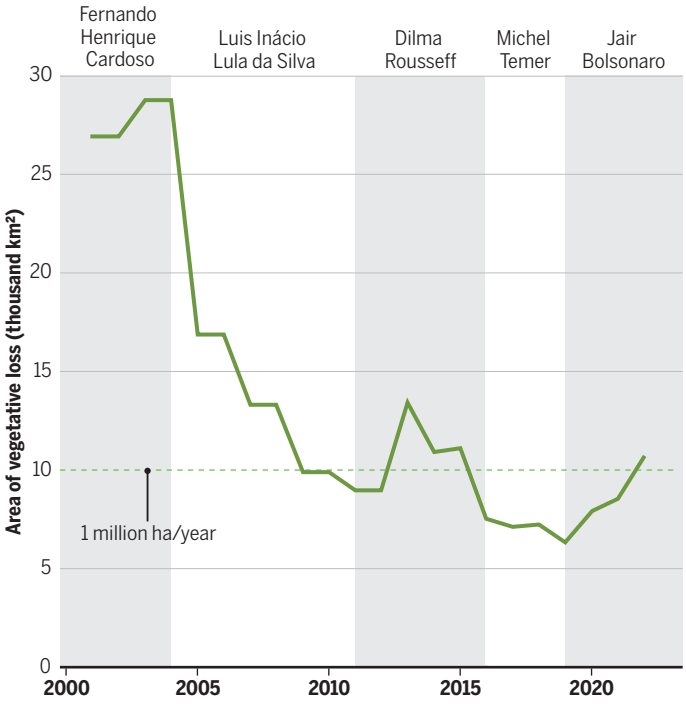
The natural vegetation removal rate is very high in the Brazilian Cerrado. Between 2002 and 2011, devegetation was nearly 1% per year, a value 2.5 times higher than that for deforestation in the Amazon (*9*). This amounts to nearly 1 million ha of loss per year. In 2022, the National Institute for Space Research (a research unit of the Brazilian Ministry of Science, Technology and Innovations) released a report indicating a reduction in devegetation in the Cerrado after 2013 (*10*). However, over the past 4 years, devegetation increased to ~1 million ha per year in 2022 (see the figure), which included privately owned regions. This decrease in vegetaion has reduced connectivity of the Cerrado landscape and, consequently, has compromised the availability of natural habitat for the Cerrado's species and the local dynamic of wild populations (*11*).

What is the contribution of private lands to biodiversity conservation worldwide? Data on protected areas indicate that it is about 1.41% of the area under protection (*5*). Thus, the participation of private lands, including those from nongovernmental institutions, could be more robust. About 75% of the Cerrado's 2 million km^2 is privately owned, constituting only 0.06% of the protected area, which leaves 74.94% of private areas without protection (*12*). This is in contrast with the 3.09% of the biome's public area that is for biodiversity conservation.

Convincing landowners to engage in biodiversity conservation is challenging because agribusiness profit is much more attractive than the few incentives offered by the government. The latter include an exception from rural territory tax, payment for water supply, and priority in the analysis of rural credit applications in official banks. Agribusiness activity expansion is continuing to replace native areas (*13*). Although clearing native vegetation areas is cheaper than restoring degraded areas, it is not sustainable.

Loss of vegetation in the Brazilian Cerrado

Devegetation across the Brazilian Cerrado over the terms of five consecutive presidents shows that vegetative loss decreased from about 2004 to 2011, and then again from 2013 to about 2019. Devegetation has increased since 2019.



New strategies that could prevent further degradation by keeping natural vegetation inside private areas are needed. This will require solving conflicts between divergent Brazilian laws. The Native Vegetation Protection Law (federal law no. 12.651/2012, known as the Brazilian Forest Code) allows, but does not force, private landowners to clear up to 80% of their property to implement agricultural activities. Theoretically, a landowner could pursue agriculture in a smaller portion of the property, leaving the rest covered by native vegetation. However, a law known as the Agrarian Reform Law (federal law no. 8629/1993) indicates that unproductive rural private properties that are not economically exploited "properly" are subject to expropriation—that is, the rural owner must economically explore at least 80% of the property. Thus, the landowners choose to remove natural vegetation to reduce the risk of losing property to the government. Another substantial change that is needed is developing a new economic model for the Cerrado based on the sustainable exploitation of the biodiversity components without extinguishing them. Payment for ecosystem services on private lands in the Cerrado is another action that should be expanded and regulated throughout the biome. Certification of rural properties that balance agricultural production and biodiversity conservation is a valid incentive to encourage landowners to conserve. In addition, private landowners of degraded areas in the Cerrado should have financing and technical support to recover these areas.

The most important message that the decision-makers should take home from the study of De Marco *et al.* is that the role of private rural properties in protecting native species in the Cerrado is vital to complement the conservation activities promoted within public protected areas. Only a combination of public and private efforts will achieve the international commitments to prevent biodiversity loss. ■

REFERENCES AND NOTES

1. S. H. M. Butchart *et al.*, *Science* **328**, 1164 (2010).
2. S. S. Hasan, L. Zhen, M. G. Miah, T. Ahamed, A. Samie, *Environ. Dev.* **34**, 100527 (2020).
3. J. Geldmann *et al.*, *Biol. Conserv.* **161**, 230 (2013).
4. P. De Marco Jr. *et al.*, *Science* **380**, 298 (2023).
5. UN Environment Programme World Conservation Monitoring Centre (UNEP-WCMC), International Union for Conservation of Nature (IUCN), Protected Planet: The World Database on Protected Areas (WDPA) and World Database on Other Effective Area-based Conservation Measures (WD-OECM) (2023); www.protectedplanet.net.
6. Y. Zeng, L. P. Koh, D. S. Wilcove, *Sci. Adv.* **8**, eabl9885 (2022).
7. J. Schüler, M. M. C. Bustamante, *J. Appl. Ecol.* **59**, 2616 (2022).
8. B. Soares-Filho *et al.*, *Science* **344**, 363 (2014).
9. B. B. N. Strassburg *et al.*, *Nat. Ecol. Evol.* **1**, 0099 (2017).
10. National Institute for Space Research (INPE), TerraBrasilis (2022); http://terrabrasilis.dpi.inpe.br/app/dashboard/deforestation/biomes/cerrado/increments.
11. T. O. Grande *et al.*, *Landsc. Ecol.* **35**, 639 (2020).
12. Ministry of Environment, Painel Unidades de Conservação Brasileiras (2023); https://cnuc.mma.gov.br/powerbi.
13. S. P. Polizel *et al.*, *Land Use Policy* **109**, 105713 (2021).

10.1126/science.adh1840

CREDITS: (GRAPHIC) D. AN-PHAM/*SCIENCE*; (DATA) NATIONAL INSTITUTE FOR SPACE RESEARCH (INPE), HTTP://TERRABRASILIS.DPI.INPE.BR/DOWNLOADS/

CANCER

Tumor suppression by RNA surveillance

A cyclin-dependent kinase triggers degradation of prematurely terminated RNAs

By **Robert P. Fisher**

Cyclin-dependent kinases (CDKs) form heterodimers with cyclins to control cell division and RNA polymerase II (RNAPII)–dependent transcription (*1*). The transcriptional CDK family (CDK7, -8, -9, -11, -12, -13, and -19) is the focus of recent drug discovery efforts, which are premised on the heightened transcriptional dependencies of cancer cells (*2*, *3*). On page 258 of this issue, Insco *et al.* (*4*) report uncovering a tumor-suppressive role of CDK13. Their findings challenge two long-standing generalizations—that transcriptional CDKs pair monogamously with dedicated cyclin partners and that they execute their functions exclusively during transcription. They also suggest a seemingly general role for nuclear RNA surveillance in restraining oncogenesis.

Insco *et al.* analyzed cancer genomic data and identified *CDK13* as a melanoma driver gene, mutation or down-regulation of which is oncogenic and associated with decreased survival. Recurrent, cancer-associated mutations were enriched within the kinase domain, suggesting that CDK13 mutants, which retained cyclin-binding ability but had little or no kinase activity in vitro, might act in a dominant-negative manner to impede wild-type CDK13 function in tumors, which were uniformly heterozygous at the *CDK13* locus. In a zebrafish melanoma model, expression of human melanoma–associated *CDK13* variants (collectively called $CDK13^{mut}$) generated more aggressive melanomas in the presence of wild-type zebrafish *cdk13*, consistent with dominant-negative action. Unexpectedly, whereas CRISPR-mediated loss of the canonical CDK13 partner cyclin K (*5*) accelerated oncogenesis by $CDK13^{mut}$, deletion of the gene encoding cyclin T1—thought to be the dedicated partner of CDK9 (*2*)—delayed $CDK13^{mut}$ melanoma progression in zebrafish, indicating a requirement for cyclin T1 in $CDK13^{mut}$-driven tumors.

Department of Oncological Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
Email: robert.fisher@mssm.edu

Analysis of steady-state RNA levels in $CDK13^{mut}$ human cells revealed accumulation of prematurely terminated RNA (ptRNA) ending at intronic polyadenylation (IPA) sites. The distribution of nascent transcripts was largely unchanged, suggesting that increased stability of truncated transcripts, rather than increased frequency of premature termination, explained the ptRNA accumulation. Consistent with this mechanism, in human melanoma cells CDK13 was associated with polyadenylate-binding nuclear protein 1 (PABPN1) and zinc finger CCCH domain–containing protein 14 (ZC3H14), which interact with the poly(A) tail exosome targeting (PAXT) complex (*6*, *7*) that normally ensures rapid degradation of ptRNA. CDK13–cyclin T1 phosphorylated ZC3H14 on Ser^{475} in vitro; this phosphorylation was diminished in $CDK13^{mut}$ cells. Moreover, expression of $ZC3H14^{S475A}$, in which Ser^{475} is replaced by a nonphosphorylatable alanine residue, reduced the interaction of ZC3H14 with PAXT, increased ptRNA accumulation, and accelerated division in cells with wild-type *CDK13*.

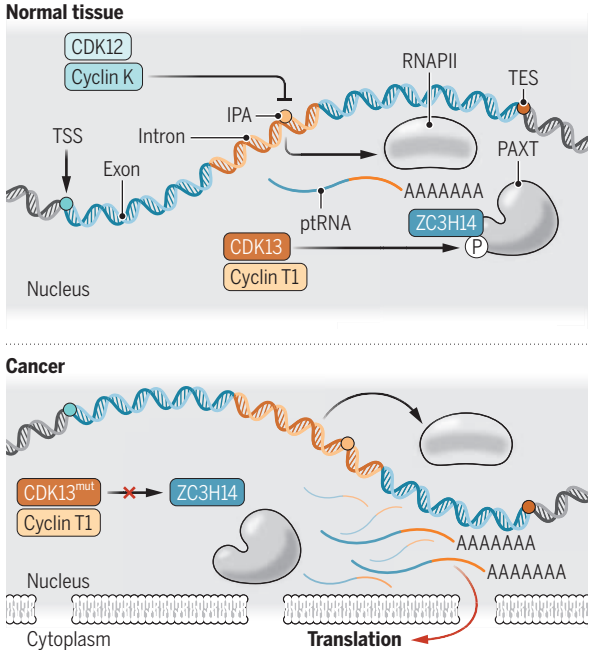
Is ptRNA accumulation oncogenic? Insco *et al.* mined RNA sequencing data to reveal increases in PAXT-regulated ptRNAs in $CDK13^{mut}$ tumors from multiple cancer types. Genes encoding PAXT subunits or adaptors are frequently deleted or mutated in melanomas and other cancers. Moreover, forced expression of ptRNAs up-regulated by $CDK13^{mut}$ accelerated melanoma onset in wild-type *CDK13* zebrafish. These findings suggest that ptRNA accumulation is intrinsically oncogenic and that a nuclear RNA surveillance pathway is tumor suppressive and vulnerable to genetic ablation.

It is less clear how ptRNA accumulation contributes to oncogenesis, which did not seem to depend on stabilization of specific truncated transcripts or loss of the corresponding, full-length mRNAs. Insco *et al.* provide evidence that ptRNA accumulation can perturb cellular metabolism. In $CDK13^{mut}$ human melanoma cells, ptRNAs were exported to the cytoplasm and translated, leading to the production of truncated proteins in amounts equaling those of many full-length proteins. The aberrant translation products contained sequences encoded in introns, which could generate neoantigens that elicit tumor-specific immune responses and might be leveraged for immunotherapy (see the figure). Whether these truncated proteins account for the proliferative advantage or increased aggressiveness of cancers deficient in PAXT-enforced RNA surveillance requires further investigation. In a previous study, impairment of this pathway led to cytoplasmic transport of normally unstable transcripts and global repression of protein synthesis (*7*).

There is precedent for increased accumulation of ptRNA arising from IPA events in cancer (*8*). Mutations in the gene encoding CDK12—a CDK13

Loss of RNA quality control drives cancer

In normal cells, ptRNAs initiating from the TSS and terminating at IPA sites are negatively regulated by CDK12–cyclin K and by a posttranscriptional RNA surveillance system whereby ptRNAs are targeted by PAXT for degradation, which is controlled by CDK13–cyclin T1. In melanomas and other cancers harboring CDK13 mutations ($CDK13^{mut}$), nuclear RNA surveillance is defective owing to loss of ZC3H14 phosphorylation (P) and stabilization of oncogenic, polyadenylated ptRNAs that are then translated in the cytoplasm.



CDK, cyclin-dependent kinase; IPA, intronic polyadenylation site; PAXT, poly(A) tail exosome targeting; ptRNA, prematurely terminated RNA; RNAPII, RNA polymerase II; TES, transcription end site; TSS, transcription start site; ZC3H14, zinc finger CCCH domain–containing protein 14.

paralog that partners with cyclin K (*5*)—promote ptRNA accumulation by increasing the frequency of premature termination (*9*, *10*) rather than acting posttranscriptionally. Mutations in *CDK12* and *CDK13* generate different ptRNA profiles, with *CDK12* mutations preferentially affecting genes involved in an effective DNA damage response (DDR) (*9*, *10*). The consequent reduction in full-length mRNAs encoding DDR factors is thought to underlie defects in homology-directed repair and sensitivity to poly(ADP-ribose) polymerase (PARP) inhibitors—a BRCA-like phenotype—of *CDK12* mutant cancers (*11*). The findings of Insco *et al.* raise the possibility that ptRNA translation might also contribute to oncogenesis in *CDK12* mutant tumors.

CDK12 and -13 share ~92% identity in their kinase domains, have similar structures in complexes with cyclin K (*5*, *12*, *13*), and are sensitive to the same small-molecule inhibitors (*14*). Selective inhibition of either CDK caused perturbations of RNAPII elongation, which were more widespread when both kinases were inactivated (*15*). Therefore, CDK12 and -13 are at least partially redundant. The requirement of CDK13 in nuclear RNA surveillance, as uncovered by Insco *et al.*, may be unique to CDK13 and specific to CDK13–cyclin T1. By associating with multiple cyclins, CDKs involved in cell cycle control acquire different substrate specificities or subcellular localizations (*1*); this may be an example of such promiscuity by a transcriptional CDK, but whether CDK13–cyclin T1 differs from CDK13–cyclin K in substrate specificity or functional targeting (e.g., recruitment to chromatin or activity in the nucleoplasm) remains to be tested. It is not necessarily the case, however, that ZC3H14 phosphorylation must occur posttranscriptionally; it might occur during transcript elongation but execute its function after transcription terminates. As drugs targeting transcriptional CDKs advance toward clinical applications, a better understanding of CDK function in posttranscriptional RNA metabolism, in both normal and cancer cells, is needed. ■

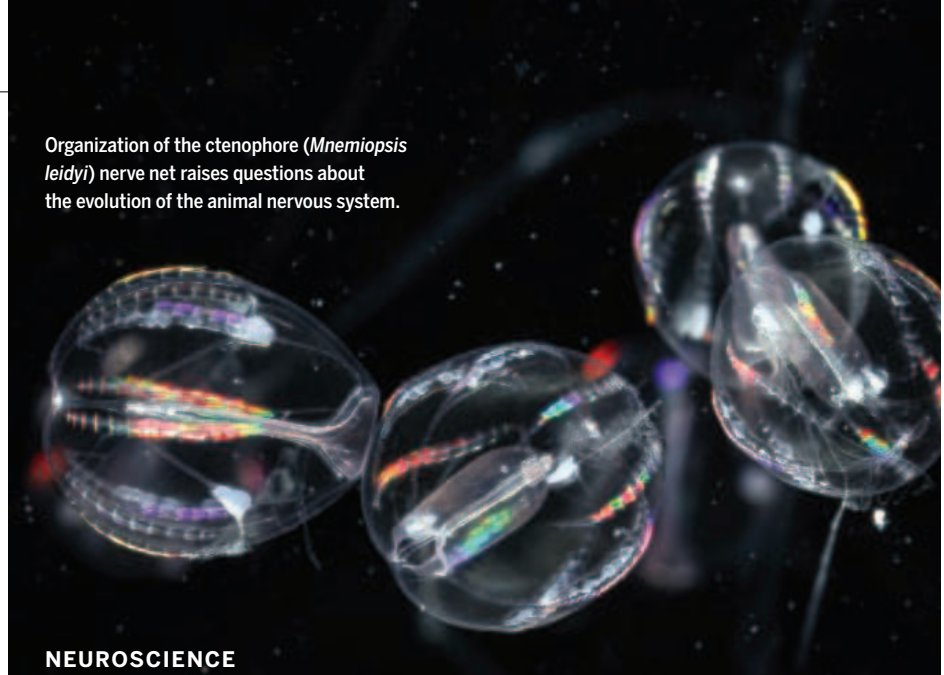
REFERENCES AND NOTES

1. M. Malumbres, *Genome Biol.* **15**, 122 (2014).
2. P. K. Parua, R. P. Fisher, *Nat. Chem. Biol.* **16**, 716 (2020).
3. S. J. Vervoort *et al.*, *Nat. Rev. Cancer* **22**, 5 (2022).
4. M. L. Insco *et al.*, *Science* **380**, eabn7625 (2023).
5. A. L. Greenleaf, *Transcription* **10**, 91 (2019).
6. N. Meola *et al.*, *Mol. Cell* **64**, 520 (2016).
7. K. Ogami *et al.*, *Genes Dev.* **31**, 1257 (2017).
8. S.-H. Lee *et al.*, *Nature* **561**, 127 (2018).
9. S. J. Dubbury *et al.*, *Nature* **564**, 141 (2018).
10. M. Krajewska *et al.*, *Nat. Commun.* **10**, 1757 (2019).
11. D. Blazek *et al.*, *Genes Dev.* **25**, 2158 (2011).
12. C. A. Bösken *et al.*, *Nat. Commun.* **5**, 3505 (2014).
13. A. K. Greifenberg *et al.*, *Cell Rep.* **14**, 320 (2016).
14. T. Zhang *et al.*, *Nat. Chem. Biol.* **12**, 876 (2016).
15. Z. Fan *et al.*, *Sci. Adv.* **6**, eaaz5041 (2020).

ACKNOWLEDGMENTS

R.P.F. is supported by NIH grant R35 GM127289.

10.1126/science.adh4051



Organization of the ctenophore (*Mnemiopsis leidyi*) nerve net raises questions about the evolution of the animal nervous system.

PHOTO: JOAN J. SOTO-ÀNGEL

NEUROSCIENCE

Neurons that connect without synapses

The ctenophore nerve net suggests a complex evolutionary history of the animal nervous system

By **Casey Dunn**

Long-standing wisdom about the evolution of the animal nervous system posits that all neurons connect to each other with synapses and that the nervous system arose once in evolutionary history and was never lost. But this tidy picture has seen surprising challenges in recent years. On page 293 of this issue, Burkhardt *et al.* (*1*) provide new information on the structure of the nervous system of ctenophores—marine invertebrates commonly known as comb jellies. These exciting findings further erode this traditional view and help build a fascinating and more complex understanding of nervous system evolution.

All living animals belong to one of five groups. Of these, Porifera (sponges) and Placozoa (small, disc-shaped animals) lack neurons. Ctenophora (comb jellies) and Cnidaria (corals, medusa jellyfish, siphonophores, and others) have nerve nets—nervous systems with neurons arranged into diffuse networks. Bilateria (the group that contains most animal species, including vertebrates, arthropods, and many other invertebrates) includes some animals with a nerve net, but most have a central nervous system.

Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA. Email: casey.dunn@yale.edu

The traditional explanation for this nervous system diversity is that these organisms represent ancestral steps in the increase of nervous system complexity. In this scenario, sponges diverged first from other animals, before the origin of the nervous system (*2*), and nervous system complexity increased in a ratchet-like manner in other animals.

It was unexpected, then, when the first Placozoa (*3*) and Porifera (*4*) genomes were sequenced and found to contain genes that were previously thought to be specific to nervous system function. A closer look in placozoans found that they have gland cells that secrete neurosecretory components (*5*). More recently, single-cell expression analyses revealed that some sponge cells communicate through structures that resemble synapses (*6*). This has made it clear that different nervous system features, such as neuron morphology and neuron signaling molecules, have different distributions across animals.

In parallel, traditional hypotheses about the earliest relationships in the animal phylogeny have been challenged. Some phylogenomic analyses support Porifera as the sister group to all other animals (*7*). There is growing evidence, however, that Ctenophora is the sister group to all other animals (*8, 9*). The latter indicates that some nervous system features arose independently in ctenophores or that some nervous system components were lost in

paralog that partners with cyclin K (*5*)—promote ptRNA accumulation by increasing the frequency of premature termination (*9*, *10*) rather than acting posttranscriptionally. Mutations in *CDK12* and *CDK13* generate different ptRNA profiles, with *CDK12* mutations preferentially affecting genes involved in an effective DNA damage response (DDR) (*9*, *10*). The consequent reduction in full-length mRNAs encoding DDR factors is thought to underlie defects in homology-directed repair and sensitivity to poly(ADP-ribose) polymerase (PARP) inhibitors—a BRCA-like phenotype—of *CDK12* mutant cancers (*11*). The findings of Insco *et al.* raise the possibility that ptRNA translation might also contribute to oncogenesis in *CDK12* mutant tumors.

CDK12 and -13 share ~92% identity in their kinase domains, have similar structures in complexes with cyclin K (*5*, *12*, *13*), and are sensitive to the same small-molecule inhibitors (*14*). Selective inhibition of either CDK caused perturbations of RNAPII elongation, which were more widespread when both kinases were inactivated (*15*). Therefore, CDK12 and -13 are at least partially redundant. The requirement of CDK13 in nuclear RNA surveillance, as uncovered by Insco *et al.*, may be unique to CDK13 and specific to CDK13–cyclin T1. By associating with multiple cyclins, CDKs involved in cell cycle control acquire different substrate specificities or subcellular localizations (*1*); this may be an example of such promiscuity by a transcriptional CDK, but whether CDK13–cyclin T1 differs from CDK13–cyclin K in substrate specificity or functional targeting (e.g., recruitment to chromatin or activity in the nucleoplasm) remains to be tested. It is not necessarily the case, however, that ZC3H14 phosphorylation must occur posttranscriptionally; it might occur during transcript elongation but execute its function after transcription terminates. As drugs targeting transcriptional CDKs advance toward clinical applications, a better understanding of CDK function in posttranscriptional RNA metabolism, in both normal and cancer cells, is needed. ■

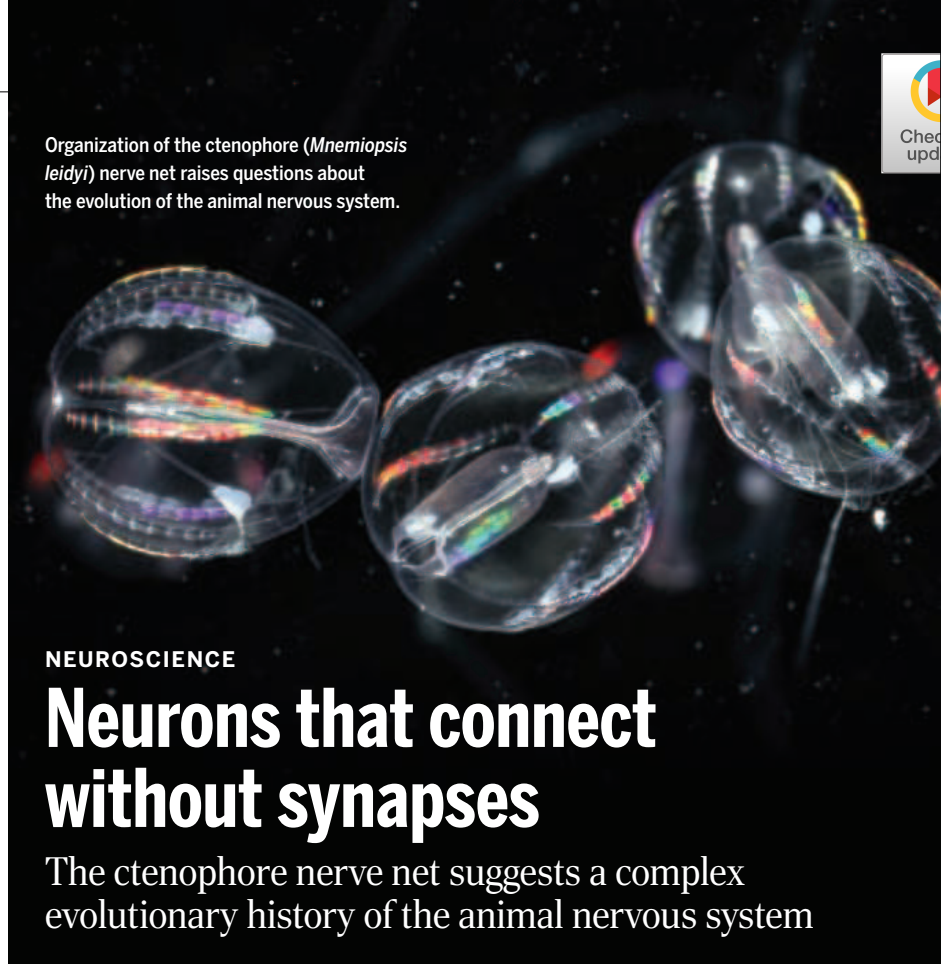
REFERENCES AND NOTES

1. M. Malumbres, *Genome Biol.* **15**, 122 (2014).
2. P. K. Parua, R. P. Fisher, *Nat. Chem. Biol.* **16**, 716 (2020).
3. S. J. Vervoort *et al.*, *Nat. Rev. Cancer* **22**, 5 (2022).
4. M. L. Insco *et al.*, *Science* **380**, eabn7625 (2023).
5. A. L. Greenleaf, *Transcription* **10**, 91 (2019).
6. N. Meola *et al.*, *Mol. Cell* **64**, 520 (2016).
7. K. Ogami *et al.*, *Genes Dev.* **31**, 1257 (2017).
8. S.-H. Lee *et al.*, *Nature* **561**, 127 (2018).
9. S. J. Dubbury *et al.*, *Nature* **564**, 141 (2018).
10. M. Krajewska *et al.*, *Nat. Commun.* **10**, 1757 (2019).
11. D. Blazek *et al.*, *Genes Dev.* **25**, 2158 (2011).
12. C. A. Bösken *et al.*, *Nat. Commun.* **5**, 3505 (2014).
13. A. K. Greifenberg *et al.*, *Cell Rep.* **14**, 320 (2016).
14. T. Zhang *et al.*, *Nat. Chem. Biol.* **12**, 876 (2016).
15. Z. Fan *et al.*, *Sci. Adv.* **6**, eaaz5041 (2020).

ACKNOWLEDGMENTS

R.P.F. is supported by NIH grant R35 GM127289.

10.1126/science.adh4051



Organization of the ctenophore (*Mnemiopsis leidyi*) nerve net raises questions about the evolution of the animal nervous system.

PHOTO: JOAN J. SOTO-ÀNGEL

NEUROSCIENCE

Neurons that connect without synapses

The ctenophore nerve net suggests a complex evolutionary history of the animal nervous system

By **Casey Dunn**

Long-standing wisdom about the evolution of the animal nervous system posits that all neurons connect to each other with synapses and that the nervous system arose once in evolutionary history and was never lost. But this tidy picture has seen surprising challenges in recent years. On page 293 of this issue, Burkhardt *et al.* (*1*) provide new information on the structure of the nervous system of ctenophores—marine invertebrates commonly known as comb jellies. These exciting findings further erode this traditional view and help build a fascinating and more complex understanding of nervous system evolution.

All living animals belong to one of five groups. Of these, Porifera (sponges) and Placozoa (small, disc-shaped animals) lack neurons. Ctenophora (comb jellies) and Cnidaria (corals, medusa jellyfish, siphonophores, and others) have nerve nets—nervous systems with neurons arranged into diffuse networks. Bilateria (the group that contains most animal species, including vertebrates, arthropods, and many other invertebrates) includes some animals with a nerve net, but most have a central nervous system.

Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA. Email: casey.dunn@yale.edu

The traditional explanation for this nervous system diversity is that these organisms represent ancestral steps in the increase of nervous system complexity. In this scenario, sponges diverged first from other animals, before the origin of the nervous system (*2*), and nervous system complexity increased in a ratchet-like manner in other animals.

It was unexpected, then, when the first Placozoa (*3*) and Porifera (*4*) genomes were sequenced and found to contain genes that were previously thought to be specific to nervous system function. A closer look in placozoans found that they have gland cells that secrete neurosecretory components (*5*). More recently, single-cell expression analyses revealed that some sponge cells communicate through structures that resemble synapses (*6*). This has made it clear that different nervous system features, such as neuron morphology and neuron signaling molecules, have different distributions across animals.

In parallel, traditional hypotheses about the earliest relationships in the animal phylogeny have been challenged. Some phylogenomic analyses support Porifera as the sister group to all other animals (*7*). There is growing evidence, however, that Ctenophora is the sister group to all other animals (*8, 9*). The latter indicates that some nervous system features arose independently in ctenophores or that some nervous system components were lost in

sponges. This further challenges the historically accepted notion that there has been a simple, stepwise increase in nervous system complexity through the course of animal evolution.

Burkhardt *et al.* provide critical understanding about the structure of the ctenophore nerve net that goes right to the heart of these questions. The authors report that the ctenophore nerve net is unlike the nervous systems of other animals. The difference is particularly relevant to debates at the dawn of neurobiology. In the late 19th century, Golgi proposed that the nervous system is a syncytial continuum, with the neurons directly connected with shared cell membranes and cytoplasm (*10*). This is known as the reticulate theory of nervous system structure. Ramón y Cajal proposed instead that neurons are distinct cells (*11*). This is known as the neuronal doctrine, and the discovery of synapses seemed to settle the debate in favor of this view.

Burkhardt *et al.* used serial block face scanning electron microscopy to make three-dimensional ultrastructural reconstructions of a ctenophore subepithelial nerve net. They observed that this nerve net is not formed by neurons connecting to each other with synapses. Instead, the processes of the neurons are directly fused to each other, forming a syncytial continuum. There are synapses elsewhere, including where the nerve net connects to effector cells, but the subepithelial nerve net itself is not formed with synaptic connections.

The findings of Burkhardt *et al.* suggest that Ramón y Cajal's neuronal doctrine and Golgi's reticulate theory are not universally exclusive hypotheses. Most animals with nervous systems (cnidarians and bilaterians) conform to the neuronal doctrine of separate cells that communicate through synapses. The subepithelial nerve net of this ctenophore species consists of fused neurons, as in Golgi's reticulate theory. It was never a question, then, of whether all animal nervous systems conform to the neuronal doctrine or the reticulate theory but rather of describing which animals conform to which theory.

There is much that remains unknown about the anatomy, physiology, genome biology, and natural history of nonbilaterian animals. This creates the dual illusions that nonbilaterians are simpler than they are (because studies tend to focus on bilaterian traits they lack rather than the many distinctive traits that they have) and that these animals are more similar to each other than they actually are (because superficial similarities are often prioritized over clear differences) (*12*). Studies such as that of Burkhardt *et al.* are important for dispelling such illusions. For example, ctenophores and cnidarians are both transparent and squishy (they are "jellyfish" in the broad sense), and both have nerve nets. Historically, they have therefore been placed together in the animal tree of life as Coelenterata. This grouping has been taken as evidence that Ctenophora cannot be a sister group to all other animals. But many animals that live suspended in the open ocean (as many ctenophores and cnidarians do), including some annelids and molluscs, have independently converged in soft transparent bodies, which indicates that so-called jellyfish traits are not good support for Coelenterata (*12*). This leaves the nerve net as one of the only traits uniting Ctenophora and Cnidaria into the single group Coelenterata. Burkhardt *et al.* show that resemblances of nerve nets between Ctenophora and Cnidaria are also superficial, and they remove some of the last remaining evidence for Coelenterata.

The findings of Burkhardt *et al.* help drive home the point that the gain of the nervous system should not be marked as a singular event in the history of animal evolution. Instead, the evolution of many constituent traits that can together make up a nervous system should be considered, including morphology, molecular inventory, and physiology (*13*). Some animals may have lost nervous system components, as may be the case in sponges. Sponges are filter feeders, which tend to have reduced nervous systems even within Bilateria. Other animals may have convergently evolved superficially similar nervous system features, such as the nerve nets of ctenophores and cnidarians, when faced with the same functional challenges.

It is exciting that such fundamental observations about animal anatomy, like those described by Burkhardt *et al.*, can still have such big implications for the study of animal evolution. The work of Burkhardt *et al.* shows how much potential lies at the growing intersection of comparative morphology, phylogenetics, physiology, and genomics. Answering the most important open questions about early animal evolution will require the integration of all these diverse approaches and perspectives. ■

REFERENCES AND NOTES

1. P. Burkhardt *et al.*, *Science* **380**, 293 (2023).
2. C. Nielsen, *R. Soc. Open Sci.* **6**, 190638 (2019).
3. M. Srivastava *et al.*, *Nature* **454**, 955 (2008).
4. M. Srivastava *et al.*, *Nature* **466**, 720 (2010).
5. C. L. Smith *et al.*, *Curr. Biol.* **24**, 1565 (2014).
6. J. M. Musser *et al.*, *Science* **374**, 717 (2021).
7. R. Feuda *et al.*, *Curr. Biol.* **27**, 3864 (2017).
8. C. W. Dunn *et al.*, *Nature* **452**, 745 (2008).
9. Y. Li *et al.*, *Mol. Biol. Evol.* **38**, 4322 (2021).
10. C. Golgi, *Sulla fina anatomia degli organi centrali del sistema nervosa* (Reggio Emilia Tipi di Stefano Calderini e Compagno, 1885).
11. S. Ramón y Cajal, *Rev. Trim. Histol. Norm. Patol* **1**, 1 (1888).
12. C. W. Dunn, S. P. Leys, S. H. D. Haddock, *Trends Ecol. Evol.* **30**, 282 (2015).
13. B. J. Liebeskind *et al.*, *Trends Ecol. Evol.* **31**, 127 (2016).

10.1126/science.adh0542

MATERIALS SCIENCE

Fixing the desalination membrane pipeline

Materials discovery alone has not translated into lower-cost water treatment

By **Jeffrey R. McCutcheon**[1] *and* **Meagan S. Mauter**[2]

Global water scarcity is motivating the expanded treatment of seawater, brackish water, and wastewater. Robust treatment trains typically include semipermeable reverse osmosis (RO) membrane barriers that allow the passage of clean water while retaining the majority (>99%) of salts, dissolved organics, and pathogens. Despite considerable research effort to optimize membrane chemistry, morphology, and module designs for diverse source-water and end-use applications, most treatment trains deploy RO membrane modules that closely resemble those developed for seawater desalination over 50 years ago. The enduring dominance of these traditional RO membranes reveals a broader need within the water treatment community to reassess the innovation pipeline for membranes for desalination and water treatment.

Past breakthroughs in membrane-based processes for desalination and water treatment were enabled by the joint discovery of new materials with desirable separation properties alongside manufacturing tools for processing these materials into membranes at scale. The first set of innovations in the 1960s combined the high-salt–rejecting properties of cellulose acetate with the nonsolvent-induced phase separation manufacturing process (*1*). Twenty years later, the discovery of aromatic polyamide materials manufactured through interfacial polymerization led to the thin-film composite (TFC) membrane that delivered 10-fold improvements in both water productivity

[1]University of Connecticut, Chemical and Biomolecular Engineering, Storrs, CT, USA. [2]Stanford University and National Alliance for Water Innovation at Lawrence Berkeley National Laboratory, CA, USA. Email: mauter@stanford.edu; jeffrey.mccutcheon@uconn.edu

sponges. This further challenges the historically accepted notion that there has been a simple, stepwise increase in nervous system complexity through the course of animal evolution.

Burkhardt *et al.* provide critical understanding about the structure of the ctenophore nerve net that goes right to the heart of these questions. The authors report that the ctenophore nerve net is unlike the nervous systems of other animals. The difference is particularly relevant to debates at the dawn of neurobiology. In the late 19th century, Golgi proposed that the nervous system is a syncytial continuum, with the neurons directly connected with shared cell membranes and cytoplasm (*10*). This is known as the reticulate theory of nervous system structure. Ramón y Cajal proposed instead that neurons are distinct cells (*11*). This is known as the neuronal doctrine, and the discovery of synapses seemed to settle the debate in favor of this view.

Burkhardt *et al.* used serial block face scanning electron microscopy to make three-dimensional ultrastructural reconstructions of a ctenophore subepithelial nerve net. They observed that this nerve net is not formed by neurons connecting to each other with synapses. Instead, the processes of the neurons are directly fused to each other, forming a syncytial continuum. There are synapses elsewhere, including where the nerve net connects to effector cells, but the subepithelial nerve net itself is not formed with synaptic connections.

The findings of Burkhardt *et al.* suggest that Ramón y Cajal's neuronal doctrine and Golgi's reticulate theory are not universally exclusive hypotheses. Most animals with nervous systems (cnidarians and bilaterians) conform to the neuronal doctrine of separate cells that communicate through synapses. The subepithelial nerve net of this ctenophore species consists of fused neurons, as in Golgi's reticulate theory. It was never a question, then, of whether all animal nervous systems conform to the neuronal doctrine or the reticulate theory but rather of describing which animals conform to which theory.

There is much that remains unknown about the anatomy, physiology, genome biology, and natural history of nonbilaterian animals. This creates the dual illusions that nonbilaterians are simpler than they are (because studies tend to focus on bilaterian traits they lack rather than the many distinctive traits that they have) and that these animals are more similar to each other than they actually are (because superficial similarities are often prioritized over clear differences) (*12*). Studies such as that of Burkhardt *et al.* are important for dispelling such illusions. For example, ctenophores and cnidarians are both transparent and squishy (they are "jellyfish" in the broad sense), and both have nerve nets. Historically, they have therefore been placed together in the animal tree of life as Coelenterata. This grouping has been taken as evidence that Ctenophora cannot be a sister group to all other animals. But many animals that live suspended in the open ocean (as many ctenophores and cnidarians do), including some annelids and molluscs, have independently converged in soft transparent bodies, which indicates that so-called jellyfish traits are not good support for Coelenterata (*12*). This leaves the nerve net as one of the only traits uniting Ctenophora and Cnidaria into the single group Coelenterata. Burkhardt *et al.* show that resemblances of nerve nets between Ctenophora and Cnidaria are also superficial, and they remove some of the last remaining evidence for Coelenterata.

The findings of Burkhardt *et al.* help drive home the point that the gain of the nervous system should not be marked as a singular event in the history of animal evolution. Instead, the evolution of many constituent traits that can together make up a nervous system should be considered, including morphology, molecular inventory, and physiology (*13*). Some animals may have lost nervous system components, as may be the case in sponges. Sponges are filter feeders, which tend to have reduced nervous systems even within Bilateria. Other animals may have convergently evolved superficially similar nervous system features, such as the nerve nets of ctenophores and cnidarians, when faced with the same functional challenges.

It is exciting that such fundamental observations about animal anatomy, like those described by Burkhardt *et al.*, can still have such big implications for the study of animal evolution. The work of Burkhardt *et al.* shows how much potential lies at the growing intersection of comparative morphology, phylogenetics, physiology, and genomics. Answering the most important open questions about early animal evolution will require the integration of all these diverse approaches and perspectives. ■

REFERENCES AND NOTES

1. P. Burkhardt *et al.*, *Science* **380**, 293 (2023).
2. C. Nielsen, *R. Soc. Open Sci.* **6**, 190638 (2019).
3. M. Srivastava *et al.*, *Nature* **454**, 955 (2008).
4. M. Srivastava *et al.*, *Nature* **466**, 720 (2010).
5. C. L. Smith *et al.*, *Curr. Biol.* **24**, 1565 (2014).
6. J. M. Musser *et al.*, *Science* **374**, 717 (2021).
7. R. Feuda *et al.*, *Curr. Biol.* **27**, 3864 (2017).
8. C. W. Dunn *et al.*, *Nature* **452**, 745 (2008).
9. Y. Li *et al.*, *Mol. Biol. Evol.* **38**, 4322 (2021).
10. C. Golgi, *Sulla fina anatomia degli organi centrali del sistema nervosa* (Reggio Emilia Tipi di Stefano Calderini e Compagno, 1885).
11. S. Ramón y Cajal, *Rev. Trim. Histol. Norm. Patol* **1**, 1 (1888).
12. C. W. Dunn, S. P. Leys, S. H. D. Haddock, *Trends Ecol. Evol.* **30**, 282 (2015).
13. B. J. Liebeskind *et al.*, *Trends Ecol. Evol.* **31**, 127 (2016).

10.1126/science.adh0542

MATERIALS SCIENCE

Fixing the desalination membrane pipeline

Materials discovery alone has not translated into lower-cost water treatment

By **Jeffrey R. McCutcheon**[1] *and* **Meagan S. Mauter**[2]

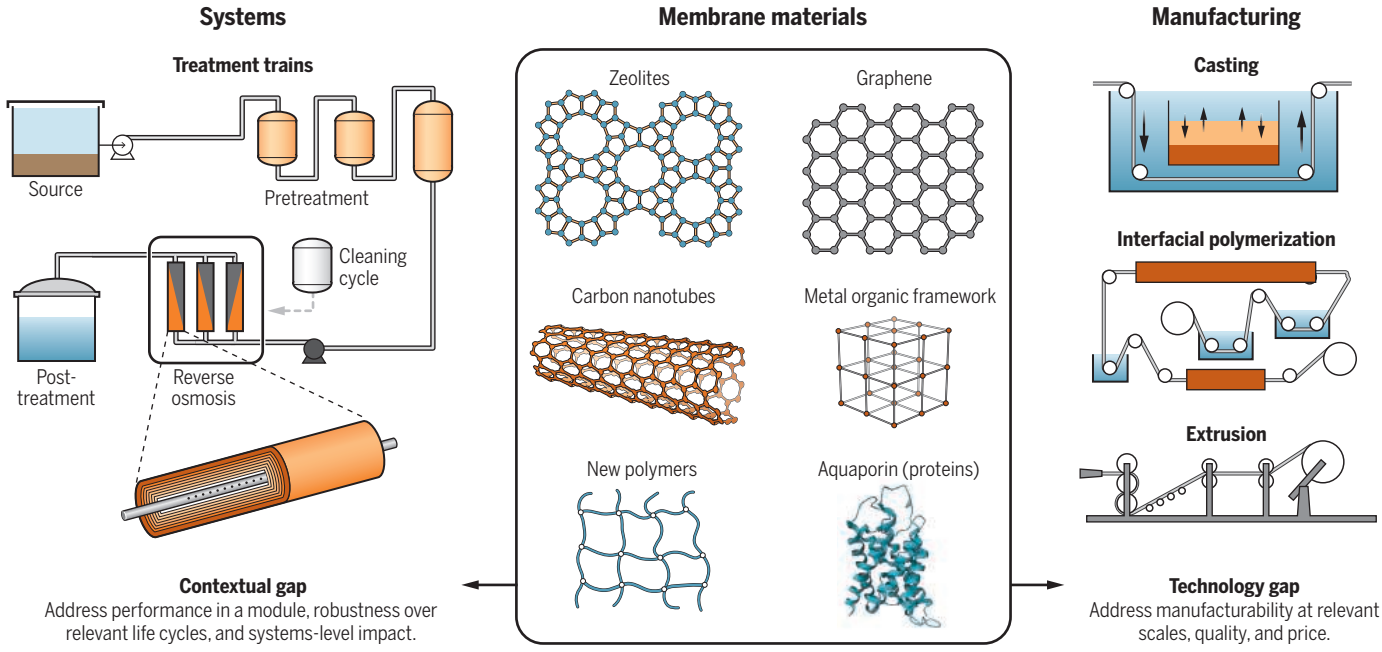
Global water scarcity is motivating the expanded treatment of seawater, brackish water, and wastewater. Robust treatment trains typically include semipermeable reverse osmosis (RO) membrane barriers that allow the passage of clean water while retaining the majority (>99%) of salts, dissolved organics, and pathogens. Despite considerable research effort to optimize membrane chemistry, morphology, and module designs for diverse source-water and end-use applications, most treatment trains deploy RO membrane modules that closely resemble those developed for seawater desalination over 50 years ago. The enduring dominance of these traditional RO membranes reveals a broader need within the water treatment community to reassess the innovation pipeline for membranes for desalination and water treatment.

Past breakthroughs in membrane-based processes for desalination and water treatment were enabled by the joint discovery of new materials with desirable separation properties alongside manufacturing tools for processing these materials into membranes at scale. The first set of innovations in the 1960s combined the high-salt–rejecting properties of cellulose acetate with the nonsolvent-induced phase separation manufacturing process (*1*). Twenty years later, the discovery of aromatic polyamide materials manufactured through interfacial polymerization led to the thin-film composite (TFC) membrane that delivered 10-fold improvements in both water productivity

[1]University of Connecticut, Chemical and Biomolecular Engineering, Storrs, CT, USA. [2]Stanford University and National Alliance for Water Innovation at Lawrence Berkeley National Laboratory, CA, USA. Email: mauter@stanford.edu; jeffrey.mccutcheon@uconn.edu

Materials innovation in a systems and manufacturing context

Research on emergent membrane materials for reverse osmosis in water treatment trains has often failed to consider the systems-level context or potential for manufacturing scalability. Fixing the desalination membrane technology pipeline requires a more holistic approach to materials discovery that addresses contextual and technology gaps.



and salt rejection (*2*). Forty years later, the polyamide TFC membrane remains the gold standard RO membrane and is deployed in nearly every RO membrane treatment train, from the largest seawater desalination plant to the point-of-use RO system under the kitchen sink.

Although TFC membranes represented a step change for seawater desalination membrane performance, they are far from a perfect solution for the diverse range of nontraditional source waters that are now treated. The charge density that makes polyamide so effective for rejecting dissolved ions does not translate to efficient rejection of small neutral molecules (e.g., *N*-nitrosodimethylamine, urea, or boron) that is critical for water reuse (*3*). Polyamide membranes are also susceptible to fouling and scaling, degrade in the presence of oxidative cleaning products, and only withstand moderate pressures (~65 bar) before performance degradation or failure.

Today, desalination system designers have engineered process solutions to membrane fouling, poor rejection, oxidative degradation, and low burst pressures. Extensive pretreatment systems reduce fouling, dechlorination prior to RO reduces oxidative damage, advanced oxidation processes destroy poorly rejected organic molecules in the permeate, and thermal solutions are deployed for brine concentration. These interventions increase the specific energy consumption and the overall levelized cost (the combined operational and amortized capital costs) of desalination and water treatment (*4*).

Simultaneously, the membrane research community has proposed several new classes of materials that address these problems. Reports of membranes with increased water permeability, superb selectivity, fouling resistance, and chemical tolerance are pervasive. One popular approach enhances the properties of the polymeric matrix by embedding a second, nonpolymeric material with distinct transport, charge, or interfacial properties to form a "mixed matrix" membrane. Proposed fillers include carbon nanotubes, graphene, graphene oxide, zeolites, aquaporin, and metal organic frameworks. A second approach considers materials as alternatives to polyamide membranes. These alternative materials are intended for use as standalone membranes (not mixed matrix) and can be categorized into organic polymers and inorganic materials. Examples of polymer alternatives include other condensation polymers such as polyester (*5*) and epoxide-based membranes (*6*). Inorganic materials include ceramic, carbon molecular sieves, graphene, or graphene oxide. Unfortunately, none of these materials have yet displaced polyamide TFC membranes in large-scale applications (*7*). These materials are currently unable to be manufactured at scale; cast into defect-free, ultrathin membranes; or rolled into modules that are readily integrated into treatment skids.

These failures reveal a critical gap in the desalination membrane innovation pipeline. Membrane materials discovery, design, and synthesis have become divorced from membrane manufacturing, membrane component design, and process systems engineering. Fixing the membrane innovation pipeline will require three key interventions: a strong system-level contextualization during the materials discovery process, rigorous methods for testing new materials under relevant conditions and life cycles, and early-stage screening for manufacturability at relevant scales (see the figure).

To contextualize new membrane materials discovery, the value proposition of a material in a desalination membrane should be quantified in the context of both the unit process and the entire water treatment train (*8*). For example, aligned carbon nanotube membranes with near-frictionless pores promised to enhance water permeability (*9*), but membrane productivity is mostly limited by poor mass transfer of concentrated salt away from the membrane surface. In short, the wrong problem was being solved. Optimized and efficiently manufactured membrane spacer designs would have been a more productive research avenue for addressing low water productivity. Process systems optimization models coupled with technoeconomic and lifecycle assessment are foundational tools for identifying impactful innovation targets at the unit process and treatment train levels. These tools

GRAPHIC: K. HOLOSKI/*SCIENCE*

prioritize innovation needs by quantifying the system-level benefits of innovations such as membranes with improved organics rejection, better chlorine resistance, or higher burst pressures. They also allow scientists to define upper limits on the value of membrane performance improvements and explore trade-offs in capital and operating costs (*10*).

Once a system-level need is identified, potential material solutions must be tested under relevant conditions. Seawater desalination facilities expect to replace membranes on a 5- to 7-year cycle, but few researchers assess materials robustness over extended durations or repeated cleaning cycles. For example, although membrane surfaces modified with hydrophilic polymers or nanoscale textures might inhibit fouling in a 3-hour laboratory experiment on synthetic feedwater, testing in real waters over long periods of time can yield very different results. Organics, extracellular polymeric substances, or precipitated inorganic materials quickly cover sophisticated chemistries and textures, rendering them ineffective or even detrimental to performance if they inhibit routine cleaning. Successful fouling management approaches acknowledge that the membrane surface changes the moment the membrane is used and instead focus on new methods to sense and manage fouling in real time.

Manufacturability at scale is another essential criterion for high-impact desalination membrane development. Very few recently proposed desalination membrane materials have scalable or economically viable manufacturing pathways for serving the municipal drinking water market. For example, despite many years of effort, manufacturers have never fabricated inorganic materials at a quality, scale, and price point that are attractive for desalination applications. Furthermore, the impossibility of defect-free fabrication techniques over square kilometer areas renders single-layer inorganic materials a poor fit for deployment in desalination applications. Mixed matrix membranes offer some degree of processability and scalability of a polymer, but conventional manufacturing processes, such as casting or extrusion, make directional or location control of the filler material challenging. Likewise, emergent polymeric materials, such as self-assembled polymers, intrinsically porous polymers, and polyelectrolyte complexes, are exciting classes of potentially easy-to-manufacture materials, but the continued reliance on decades-old manufacturing processes for polymers may prevent these materials from exhibiting their desired properties in a membrane. New or modified manufacturing processes must be developed to address these weaknesses while also enabling production rates on the order of meters per minute at widths of up to 40 inches, the industry standard for current desalination membranes.

There are opportunities to fix these weaknesses in the membrane innovation pipeline for water treatment. Access to design tools that help researchers identify high-impact materials needs and R&D priorities for water treatment membranes would be particularly valuable (*10*). Collective national efforts, such as the National Alliance for Water Innovation (NAWI) in the United States, have provided baselines for component and unit process costs of diverse treatment trains, and developing an open-source water treatment costing platform promises to be one such resource (*11*). Additional progress would be facilitated by full-cost transparency from the water sectors' vertically fragmented industrial partners. Accurate estimates of component, unit process, and facility costs and life-spans are needed to effectively guide academic and government agencies toward R&D priorities.

"The status quo of overselling a membrane made at millimeter scale and promoting an unrealistic narrative...must end."

Access to manufacturing science infrastructure is also needed to address these weaknesses. Most membrane research happens in laboratories that have limited, if any, access to membrane manufacturing tools. To explore the effectiveness of new materials in membranes, an understanding of both conventional and emergent manufacturing techniques (*12*, *13*) applied to these materials is needed. To this end, scale-up and prototyping facilities for bulk material production, membrane manufacturing, and element assembly are required to support membrane researchers' attempts to address manufacturability throughout the research, development, demonstration, and deployment (RDD&D) pipeline. The US Department of Energy is investing in manufacturing demonstration infrastructure through its national laboratory facilities that will support this pipeline and the membrane manufacturing ecosystem. Roll-to-roll manufacturing tools are currently available at the National Renewable Energy Laboratory, Oak Ridge National Laboratory (Manufacturing Demonstration Facility), and Argonne National Labs (Materials Engineering Research Facility). Similar facilities outside of the United States are supported in Singapore at the Separation Technologies Applied Research and Translation Centre and the Singapore Membrane Technology Center at Nanyang Technological University. Such capabilities also exist in Australia at the Commonwealth Scientific and Industrial Research Organisation (CSIRO). In parallel, historic heavyweights in the membrane manufacturing space could more openly articulate manufacturing requirements, and experienced scientists from the roll-to-roll, extrusion, printing, polymer processing, and nanomaterials application areas could lend their expertise to water treatment membrane manufacturing. Such facilities and partnerships come with increased cost of research. This requires funding agencies to adequately support membrane manufacturing innovation and incorporate manufacturing research infrastructure and other systems-level experts when possible.

The status quo of overselling a membrane made at millimeter scale and promoting an unrealistic narrative about how it will change water treatment must end. Such efforts lead to incremental improvements in membrane performance and distract researchers from exploring manufacturing science and process systems-level questions that would yield more impactful results. Repairing the desalination membrane innovation pipeline starts by grounding materials discovery research in systems-level performance improvements, manufacturability, and life-cycle performance. Doing so would substantially improve the membrane science community's focus on applied research. ■

REFERENCES AND NOTES

1. S. Loeb, *ACS Symp. Ser.* **153**, 1 (1981).
2. J. E. Cadotte, R. J. Petersen, R. E. Larson, E. E. Erickson, *Desalination* **32**, 25 (1980).
3. T. Fujioka *et al.*, *Separ. Purif. Tech.* **116**, 426 (2013).
4. A. Sim, M. S. Mauter, *Environ. Sci. Water Res. Technol.* **7**, 748 (2021).
5. Y. Yao *et al.*, *Nat. Sustain.* **4**, 138 (2021).
6. R. Verbeke *et al.*, *Environ. Sci. Technol. Lett.* **8**, 818 (2021).
7. J. R. McCutcheon, *Joule* **3**, 1168 (2019).
8. S. K. Patel *et al.*, *Energy Environ. Sci.* **13**, 1694 (2020).
9. B. J. Hinds *et al.*, *Science* **303**, 62 (2004).
10. A. V. Dudchenko, T. V. Bartholomew, M. S. Mauter, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2022196118 (2021).
11. M. S. Mauter, D. D. Dionysiou, J.-H. Kim, *ACS ES T Eng.* **2**, 271 (2022).
12. M. R. Chowdhury, J. Steffes, B. D. Huey, J. R. McCutcheon, *Science* **361**, 682 (2018).
13. X. Qian *et al.*, *J. Membr. Sci.* **645**, 120041 (2022).

ACKNOWLEDGMENTS

The authors thank A. Dudchenko for contributions to this manuscript. Support is provided by the National Science Foundation (CMMI 2001544) and the National Alliance for Water Innovation (NAWI) funded by the US Department of Energy, Office of Energy Efficiency and Renewable Energy (EERE), Advanced Manufacturing Office, under Funding Opportunity Announcement Number DE-FOA-0001905. M.S.M.'s contribution to this publication was as a paid consultant to NAWI and was not part of her Stanford University duties or responsibilities.

10.1126/science.ade5313

RETROSPECTIVE

William Steffen (1947–2023)

Earth system scientist with passion for people and planet

By **Johan Rockström**[1] *and* **Katherine Richardson**[2]

William ("Will") Lee Steffen, influential Earth system scientist who quantified human impacts on the global environment and shaped public understanding of the planetary crisis, died on 29 January. He was 75. Will's critical insights contributed to our current understanding of Earth as a complex adaptive system characterized by interactions among a range of biogeochemical and physical components, and he warned that modern society risks exceeding the limits of planetary stability. At numerous organizations focusing on global change research, he spearheaded interdisciplinary worldwide sustainability initiatives, established bridges between climate science and policy, and communicated results with the public.

Born on 25 June 1947 in the tiny town of Clearwater, Nebraska, and raised in Spencer, Iowa, Will completed his bachelor's in chemical engineering at the University of Missouri in 1970 (supported by an oil company scholarship, a fact he later found ironic given his focus on anthropogenic climate change). He then moved to the University of Florida, where he received a master's degree in education in 1972; met and married his wife, Carrie; and—after a US Peace Corps tour to the Fiji Islands—earned his PhD in chemistry in 1975. Will did a postdoc at Cornell University and then accepted a research fellowship in chemistry in 1977 at the Australian National University (ANU) in Canberra.

In 1980, Will joined the Commonwealth Scientific and Industrial Research Organisation in Canberra. His work measuring land-atmosphere CO_2 flux laid the foundations for his systems approach to global environmental change. In 1990, he moved to the International Geosphere-Biosphere Programme (IGBP). From 1998 to 2004, based in Stockholm, he served as IGBP's executive director. Returning to Canberra in 2004, Will served as a science adviser to the government until 2011. He became director of ANU's Fenner School of Environment and Society in 2007 and served as the inaugural director of ANU's Climate Change Institute from 2008 to 2012. Will's vast experience provided him with an extensive network across scientific disciplines, and he propelled Earth system science forward by bringing together various lines of research, all focusing on different aspects of global environmental change.

In 2011, Will was appointed to Australia's Climate Commission, established by the government to communicate climate science to the public. When an administration skeptical of climate change dissolved the commission in 2013, Will—who was not afraid of confrontation where warranted—and his fellow commissioners launched the independent organization Climate Council, formed and supported through a crowdfunding campaign.

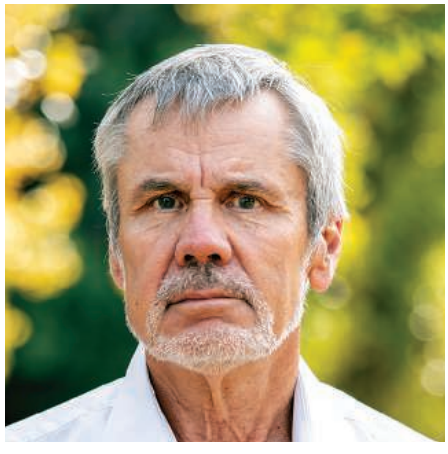


PHOTO: THE AUSTRALIAN NATIONAL UNIVERSITY

Humans' potential to change the environment at the planetary level fascinated Will. At a 2000 IGBP meeting, when chemistry Nobel laureate Paul Crutzen exclaimed in frustration that we are no longer living in the Holocene but in the "Anthropocene," Will encouraged him to publicly suggest the name change. Eager to understand when this change in geological epoch might have occurred, Will led a project that compiled the first-ever time series of human pressures on the planet. The data—recorded in *Global Change and the Earth System: A Planet Under Pressure*, the now-classic 2005 volume in an IGBP series—revealed the exponential rise in global environmental change, establishing the "great acceleration" of the effects of humans on the environment beginning in the 1950s. Will was also a member of the Anthropocene Working Group of the Subcommission on Quaternary Stratigraphy that voted in 2019 to introduce the Anthropocene as a new geological epoch.

Although based in Australia, Will maintained strong links to Europe, including affiliations at the Stockholm Resilience Centre and the Potsdam Institute for Climate Impact Research, allowing us to work together closely. His personality was an asset when building bridges between scientific disciplines and negotiating complex issues. In 2009, Will, the two of us, and colleagues developed the Planetary Boundaries framework, which defines a quantitative safe operating space for humanity on Earth by integrating the science on tipping points with our understanding of Earth system stability. Will later led the 2015 scientific update of the framework, which showed threats to climate and biosphere boundaries (including biodiversity, land use change, and nutrient loading).

Concerned that anthropogenic impacts could cause a loss of Earth's resilience, in 2018 Will invited a diverse group of scholars to address the question of how the Earth system might respond if human activities resulted in a 2°C increase in global average temperature. The group warned that anthropogenic warming could trigger a warming cascade. This "Hothouse Earth" hypothesis has generated substantial scientific debate.

Recognizing that neither Earth's resources nor its limits are distributed equitably, Will was committed to Earth system justice. He called for legal frameworks acknowledging that all people are connected and interdependent and that collective actions are therefore required to safeguard resources such as ice sheets, forests, the ocean, and biodiversity. The Planetary Boundaries also show the need for a fair distribution of remaining ecological space and the moral imperative of securing a livable Earth for future generations. These dimensions of Earth system justice prompted a proposal to establish a United Nations charter recognizing a stable planet as a "common heritage for humankind," an effort in which Will was heavily engaged.

Will's immersion in and devotion to his adopted country were complete. He became a citizen and voraciously studied the country's natural history and the history of its Indigenous peoples. An avid hiker, he reveled in climbing Australia's hills.

We will remember Will as an extremely likable person and cherished friend who combined scientific excellence with humility and sincerity. He was a strong supporter of a "whole Earth system" approach in science and policy, integrating people and planet. Through his work, he not only provided novel Earth system insights but also strove to identify potential pathways for humanity's future on Earth. ■

[1]Potsdam Institute for Climate Impact Research, 14473 Potsdam, Germany. [2]Globe Institute, University of Copenhagen, 2100 Copenhagen, Denmark.
Email: johan.rockstrom@pik-potsdam.de; kari@sund.ku.dk

10.1126/science.adh9882

POLICY FORUM

RESEARCH SECURITY

Managing United States–China university relations and risks

US universities must proactively address potential concerns

By **Richard Lester[1], Lily Tsai[2], Suzanne Berger[2], Peter Fisher[3], M. Taylor Fravel[2], David Goldston[4], Yasheng Huang[5], Daniela Rus[6]**

The intensifying geopolitical rivalry between the United States and China is clouding the outlook for cross-border academic exchange and collaboration in science and technology. Technological competition is a principal focus of this rivalry, and pressures are building in both countries to erect higher barriers to academic research collaborations and to restrict the flow of students and scholars between the two countries. A major challenge for US universities is how to manage these pressures while preserving open scientific research, open intellectual exchange, and the free flow of ideas and people. New federal regulations designed to strengthen research security on US university campuses are now being introduced. Yet federal policies, no matter how well crafted, cannot be a substitute for actions by universities themselves. We share an approach developed at the Massachusetts Institute of Technology (MIT) to make clear the lines that should not be crossed and the principles that should govern academic relations with China.

US research universities have long benefited from their ability to attract some of the world's most talented students, scholars, and innovators, many of them from China. Now, as government officials confront the immediate challenges posed by the Chinese leadership's coercive actions at home and around the world, US research universities must prepare for an extended period of adversarial relations and potential conflict between the United States and China. Can the values that underpin the excellence of US research universities survive the struggle with China? What are the nation's goals for these universities? And what role should the universities themselves play in shaping the course of academic relations with China?

Similar questions confront governments and universities elsewhere. Amid concerns about the risks to national security and academic freedom, UK universities have been warned about their "strategic dependency" on Chinese partnerships (*1*). The European Commission has recently published a "toolkit" to help universities mitigate foreign influence in research and innovation (*2*). In Japan, the government has introduced new rules that require security reviews before universities can accept foreign students and researchers (*3*). The G7 governments recently declared their intention to work together to enhance research security without undermining academic freedom and open science (*4*).

Yet governments should not lose sight of longer-term domestic and global interests in research and innovation. For example, senior Biden administration officials have recently underscored the importance of attracting talented Chinese students to the United States (*5*). In 2019, 41% of all science, technology, engineering, and mathematics (STEM) PhDs graduating from US universities were temporary visa holders, with China accounting for more of these graduates than the next nine foreign countries combined (*6*). Most Chinese PhD graduates stay in the United States, helping to advance US research and innovation (*7*). But rising United States–China tensions as well as US border control enforcement and well-publicized investigations into alleged campus intellectual property theft may be affecting Chinese graduate student applications and enrollment in STEM programs at US schools. Although the overall impact of these factors is unclear, some United States–based faculty report that top-rated STEM students at Chinese universities, who in previous years would have applied to graduate departments at leading American universities, are instead choosing to stay in Ch

United States–China tensions are also affecting cross-national collaborations between US and Chinese researchers—in recent years by far the most important axis of international collaboration for US researchers when measured by jointly authored publications (*8*). Going forward, US researchers will presumably have even more to gain from such collaborations as China's investment in science and technology continues to grow rapidly (*9*). But some US faculty who have previously collaborated with colleagues in China report that they are now holding back from joint research.

Faculty and student concerns over rising bilateral tensions have been aggravated by a series of arrests and failed prosecutions of Chinese-origin university researchers accused of enabling scientific espionage. These actions have prompted accusations that the US government is criminalizing normal scientific and academic exchange and engaging in racial stereotyping (*10*). Perceptions of bias and discrimination have helped convince many outstanding young Chinese scientists at US universities to pursue their careers in other countries (*11*).

The strained relations between the US government and the US academic community add to the importance of self-initiated efforts by universities, reflective of their institutional values. Some China-related challenges are in any case better addressed by the academic community itself, and university actions—such as upgrading campus research security, identifying clearly the kinds of interactions with China that should be out of bounds, and establishing processes for deciding on difficult cases—can also build confidence among policy-makers and may help to avoid federal government overreach in policies and regulations.

To be clear, there is an urgent need for an integrated government policy framework for academic relations with China that addresses immigration, research security, and research collaboration. But universities should develop their own policies, processes, and risk management frameworks, informed by their deeper knowledge of educational and research practices and institutional values and shaped by their role as guarantors of the intellectual autonomy of their faculty.

Even as the economic and military rivalry between the United States and China continues to build, US research universities and the nation more broadly can benefit from academic exchange with China. Ending academic relations would weaken the foundations of US science, technology, and innovation and would harm US economic development and national security.

[1]Department of Nuclear Science and Engineering, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. [2]Department of Political Science, MIT, Cambridge, MA, USA. [3]Department of Physics, MIT, Cambridge, MA, USA. [4]MIT Washington Office, Washington, DC, USA. [5]Sloan School of Management, MIT, Cambridge, MA, USA. [6]Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, MA, USA. Email: rklester@mit.edu

LINES THAT WILL NOT BE CROSSED

In 2021, the authors of this article were asked by MIT President L. Rafael Reif to chart a path for MIT's future relations with China. The resulting approach (*12*), now being implemented, is designed to help MIT advance knowledge and the needs of the United States and the world—without damaging US interests in national security or the economy, without endangering human rights, and in ways consistent with the core values of our institution.

In developing this approach, our group consulted extensively with experts in academia and government both in the United States and internationally, as well as with many members of the MIT community. We assumed that the complex and challenging international security environment will persist and that relations between the United States and China may deteriorate further. We recognized that, although continued academic relations with China will bring benefits, engagement brings its own risks, and that new approaches to managing these risks are needed.

We take seriously the concern that the Chinese government—and some other foreign governments—are targeting US university research and technology to gain advantage, mostly through legal means but sometimes illegally or improperly. We also take seriously the obstacles to academic collaborations presented by Chinese government policies that restrict academic autonomy on Chinese university campuses, increase the risk of seizure of intellectual property deemed to be in China's national interest, and attempt to exert influence over Chinese students and scholars in the West. We recognize too that when researchers at US universities collaborate with individuals or institutions in countries with authoritarian or autocratic governments, the good intentions of their direct collaborators are not enough to assure good outcomes.

Our new strategy is the latest stage in the ongoing development of policies for China at MIT that began several years ago. An important milestone occurred in 2019, when a process was introduced for case-by-case reviews of all proposed China-related research, educational, and other formal engagements, as well as engagements with certain other countries, from the perspective of risks to national security, civil and human rights, and economic competitiveness (*13*). In some cases, government regulations prescribe how these risks should be managed. For example, federal regulations require principal investigators (PIs) supported by US government funding agencies to demonstrate that their work for the government is adequately protected against theft and to disclose international collaborations that are related to that work. But regulatory compliance is often not enough to determine whether the proposed activities should be undertaken at all. MIT's elevated risk review process provides guidance on proposed activities that would not violate federal rules but nevertheless require careful assessment of risks and benefits to determine whether they should proceed. An important aspect of the process is to consider the risks of not undertaking proposed engagements as well as the risks of doing so.

Lines that will not be crossed

The Institute and its faculty are called:

- Not to host researchers and visiting students who are known to be employed by Chinese military and security institutions or who are graduates of China's civilian national defense universities
- Not to enter into research collaborations with China's civilian national defense universities, military research institutes, or national defense key laboratories at civilian universities
- Not to enter into relationships with Chinese corporate or other entities that are known to provide systems, products, or services with military applications to the Chinese armed forces, or for which there is credible evidence that their activities are contributing to the suppression of human rights in China
- Not to participate in Chinese talent recruitment programs that are designed to transfer US technology to China

The reviews involve faculty and administrative committees, and the process is coordinated by the Office of the Provost (the Associate Provost for International Activities). The toughest cases are referred to a small group of senior administrators, and decisions are necessarily based on judgment rather than precedent or standard rules. For research collaborations, decisions are made with the active participation of the PIs, who typically have the best understanding of the benefits of the proposed collaboration for research as well as the technical capabilities of their partners. A key role of the PIs is to describe the benefits that all the participants in the proposed collaboration might expect to realize, as well as potential benefits to the nation and the world.

But PIs are generally not as well informed about national security and human rights risks. Information is also needed to understand the context in which the potential Chinese collaborators are operating, including the ways in which organizations and individuals in China are connected to the Chinese government or the Chinese Communist Party and the obligations they have to them. Again, PIs usually do not have this information and do not know where to get it. The process relies on inputs from country and regional experts at MIT and elsewhere; from MIT's Washington, DC, office; and occasionally from ad hoc faculty committees that may be convened for advice on difficult cases. As a result of this process, some proposed engagements have been rejected, many have been approved, and for others, specific conditions have been applied or modifications required—in some cases, to ensure reciprocity.

The new strategy goes beyond these ex ante risk assessments and covers all aspects of MIT's interactions with China, including research security on campus, informal collaborations, the appointment of postdoctorates and visiting scientists, and executive and professional education. It provides practical guidance to the MIT community on these and other issues within a framework defined by the core mission, goals, and values of MIT. The principal goals include ensuring that all members of the MIT community, including those of Chinese origin, can thrive and do their best work without fear of external influence, bias, or discrimination; enabling our faculty, staff, and students to work with leading Chinese researchers and institutions on problems that are important to both countries and to the world; and educating our students about Chinese science, technology, innovation, business, history, culture, politics, and economics—knowledge whose benefits to our students, and more broadly to the United States, will only grow.

The new strategy also describes lines that should not be crossed in MIT's engagements with China (see the box). Other guidance for the MIT community covers technology licensing, data protection, and travel to China. Regarding upgrades to campus research security, although PIs are generally responsible for ensuring that all members of their research groups understand the norms and expectations concerning the sharing of information outside the group, the university should provide training and other guidance to help PIs with these tasks.

MAPPING SPACES FOR COLLABORATION

Many academic leaders in the United States and Europe have expressed interest in our report, and some are developing and implementing approaches of their own. A challenge for all of them, even those with extensive in-house specialist knowledge, is how to acquire the information needed for risk assessments, especially concerning China's policies, regulations, and practices in research, education, and innovation. There are opportunities for universities to work together to develop shared information resources. They may also soon be able to consult a new unclassified information service under development by the National Science Foundation that covers foreign research partners and projects that could pose security risks (*14*).

By helping to map out spaces for productive exchange and collaboration, these measures by universities will assist their own faculty, who otherwise will be less and less inclined to pursue connections with China. Each institution needs to develop an approach adapted to its own culture and internal processes. But there is value in cooperation, and MIT is working closely with university associations such as the Association of American Universities and the Association of Public and Land-grant Universities on this issue.

Although motivated by the particular context of China, many of these issues and approaches are also important for systematically managing collaborations that involve a broader range of countries in what is an increasingly complex and dynamic international environment. Again, the likely benefits must be clearly identified and the risks managed effectively, and for most universities, this will entail the development of new risk management capabilities.

Alongside this risk management agenda, US research universities must also advocate effectively to preserve two key factors that have contributed to their global leadership but are now themselves at risk as United States–China tensions mount. First is the ability to continue admitting the best and most promising students in all disciplines from around the world. When these people come to the United States to study and work, it strengthens US research and education immeasurably; investing in these students builds trust, friendships, and confidence in US research and innovation that can last a lifetime. Ensuring the greatest possible access for individuals of great ability, regardless of nationality, is the strategy that will allow US universities to produce the highest benefits for the United States and for people everywhere.

Second, the system of open scientific research that is the foundation of knowledge, education, and innovation in US research universities must be sustained and strengthened. Restricting scientific discourse stymies scientific progress by preventing researchers from building on and challenging each others' work. Erecting barriers around specific areas of academic research will deny the United States, as well as others, the benefits that result from scientific progress. Although special precautions may be needed in some new areas of research because of national security concerns, the United States has more to lose than to gain if sweeping restrictions on the conduct and publication of academic research are implemented.

To that end, we urge federal agencies to avoid blurring the distinction between open and unrestricted research. For example, they should exercise caution in expanding the reach of the Controlled Unclassified Information program (*15*), which establishes an intermediate category of government-owned information that is unclassified but subject to extra safeguards. Federal agencies should also give up the growing practice of requiring universities to apply nationality or national origin criteria to determine who should be permitted to work on their research projects. The government can and should vet which individuals are admitted to the United States, but once admitted, they should be able to participate in any unclassified research project, except if participation would violate export controls (which restrict the transfer of certain sensitive technological information to foreign persons in the United States). Preventing certain members of university communities from working in specific research fields or from studying particular academic subjects because of their nationality is deeply problematic and corrosive.

"In the current geopolitical environment, there is a risk of self-inflicted damage..."

In the current geopolitical environment, there is a risk of self-inflicted damage to the principles of openness, tolerance, and nondiscrimination that differentiate the United States from its rival. In this difficult environment, US universities have an important role in articulating and defending the values that have enabled them to flourish, in welcoming excellent Chinese students and scholars to their campuses, and in enabling their faculty and students to work safely with Chinese peers on shared intellectual challenges. Even if the overall trend in relations between the two countries is toward less rather than more engagement, there are important areas of research and education in which the academic community, the nation, and the world would be better off with more rather than less United States–China scientific collaboration.

US research universities should now work to establish a comprehensive, ongoing dialog with the federal government on the China issue. Universities should be proactive in addressing new problems as they arise, while vigorously advocating for themselves as institutions with values to uphold and with value to provide to the nation and the world. ■

REFERENCES AND NOTES

1. News Centre, "China's influence on UK research has grown ten-fold in past 20 years," King's College London, 9 March 2021; https://www.kcl.ac.uk/news/chinas-influence-on-uk-research-has-grown-ten-fold-in-past-20-years.
2. Directorate-General for Research and Innovation, "Commission publishes a toolkit to help mitigate foreign interference in research and innovation," European Commission, 18 January 2022; https://research-and-innovation.ec.europa.eu/news/all-research-and-innovation-news/commission-publishes-toolkit-help-mitigate-foreign-interference-research-and-innovation-2022-01-18_en.
3. Nikkei staff writers, "Japan to limit foreign students' access to security-linked tech," Nikkei Asia, 20 October 2021; https://asia.nikkei.com/Politics/Japan-to-limit-foreign-students-access-to-security-linked-tech.
4. D. Matthews, R. L. Hudson, "G7 science ministers urge democracies to unite research efforts," *Science Business*, 14 June 2022; https://sciencebusiness.net/news/g7-science-ministers-urge-democracies-unite-research-efforts.
5. G. Raimondo, "Remarks by U.S. Secretary of Commerce Gina Raimondo on the U.S. Competitiveness and the China Challenge," US Department of Commerce, 30 November 2022; https://www.commerce.gov/news/speeches/2022/11/remarks-us-secretary-commerce-gina-raimondo-us-competitiveness-and-china.
6. National Science Foundation, National Center for Science and Engineering Statistics, Data Tables, Tables 18 and 26, (2019); https://ncses.nsf.gov/pubs/nsf21308/data-tables.
7. National Science Board, Science and Engineering Indicators, Table 3-22 (2017); https://ncses.nsf.gov/pubs/nsb20198/immigration-and-the-s-e-workforce.
8. Nature Index, 2022 Annual Tables; https://www.nature.com/nature-index/annual-tables/2022.
9. J. Brainard, D. Normile, *Science* 10.1126/science.ade4585 (2022).
10. E. Guo, J. Aloe, K. Hao, *MIT Technology Review* 2 December 2021; https://www.technologyreview.com/2021/12/02/1040656/china-initative-us-justice-department.
11. Y. Xie *et al.*, arXiv:2209.10642 [physics.soc-ph] (2022).
12. MIT China Strategy Group, "University engagement with China: an MIT approach" (MIT, 2022); https://global.mit.edu/about/report-by-the-mit-china-strategy-group.
13. MIT Global Support Resources, "Elevated-risk project review process" (MIT, 2019); https://globalsupport.mit.edu/planning-agreements/elevated-risk-project-review-process.
14. "US science agency plans new centre for research security," *Science Business* 17 November 2022; https://sciencebusiness.net/news/us-science-agency-plans-new-centre-research-security.
15. US General Services Administration, Controlled unclassified information; https://www.gsa.gov/reference/controlled-unclassified-information-cui.

10.1126/science.adg5619

Potential buyers examine guns for sale at The Nation's Gun Show in Chantilly, Virginia.

PHOTO: JABIN BOTSFORD/THE WASHINGTON POST/GETTY IMAGES

BOOKS *et al.*

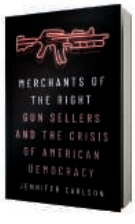
SCIENCE AND SOCIETY

Gun gatekeepers and their philosophies

Individualism, conspiracism, and partisanship fuel American gun culture, argues a sociologist

By **Matthew Lacombe**

In *Merchants of the Right*, Jennifer Carlson takes on a topic of crucial importance: the relationship between conservative gun culture and the core commitments of American democracy. Along the way, she sheds fascinating new light on the factors that galvanized the largest gun-buying spree in the country's history in 2020 and shaped how many Americans responded to the tumult of that year. Balancing engaging prose with a sober tone, Carlson shows how American gun culture has developed over the past 50 years, articulates the "civic toolkit" it provides its adherents, and vividly describes how that toolkit was applied during a year marked by pandemic, protests, and contentious politics.



Merchants of the Right: Gun Sellers and the Crisis of American Democracy
Jennifer Carlson
Princeton University Press, 2023. 288 pp.

Carlson relies on interviews with 50 gun sellers located in diverse areas of the country. Conducted between April and August 2020, these interviews provide insights into how gun sellers made sense of the year. Carlson's use of gun sellers as a window into gun culture during the pandemic is astute and clever; these individuals are not only enmeshed in gun culture, they in many ways act as its gatekeepers.

Carlson's interviews focus on how adherents of the "dominant strains of gun rights" (a phrase used to clarify that not all gun owners hold this outlook) conceive of American civic life and democracy—an approach she argues has come to be reflected in conservative politics writ large. The book contends that this culture provides individuals with a civic toolkit consisting of three interrelated tenets, which inform their political outlooks generally and their reactions to the pandemic specifically.

The first is armed individualism. Conceiving of guns as sources of personal protection and political empowerment, this tenet simultaneously views guns as helpful in the face of government underreach and as security against government overreach. Central democratic questions about collective responsibilities are boiled down to a belief that everyone is ultimately responsible for protecting themselves. Applied to the events of 2020, gun buying was a way to overcome the government's failure to provide physical safety and economic stability—to deal with the perceived chaos of the year—as well as protection against attempts by the government to use the crisis to seize control.

The second part of Carlson's framework is conspiracism. Animated by a populistic skepticism of experts, this tenet emphasizes that individuals need to come to their own conclusions in light of the misleading information those in power are said to provide on behalf of their own agendas. Conspiracy beliefs provide a way to assert control by explaining otherwise intractable conditions—an assertion reinforced by gun ownership. We are all familiar with how this line of thinking manifested in 2020: Conspiracy theories about the pandemic, racial justice protests, and the election proliferated, as did gun buying—all in response to a combination of threats that left many Americans feeling anxious.

The final tenet is an intense form of partisanship that encourages the denigration of opponents and even questions their democratic worthiness. Gun ownership, as a political identity, is attached to a wide range of conservative beliefs and (to an even greater extent) animosity toward liberals, who not only threaten gun rights but also are seen as irrational, irresponsible, and insufficiently independent. Liberal-coded measures—like mask wearing—should be rejected, faith that Democrats play fair in elections should be low, and, in the most extreme case, insurrection may be acceptable.

On the topic of partisanship, Carlson also explores how gun sellers' civic toolkit shaped their reactions to 2020's first-time gun buyers, who disproportionately came from sociopolitical demographics outside of the conservative, white, male center of gun culture. Were gun sellers—the culture's gatekeepers—accepting of these new, often liberal gun owners? The answer, by and large, is no; gun sellers viewed many liberal gun owners as "unfit for Second Amendment rights," hypocritical, and in need of political reeducation.

"Gun sellers enlisted partisanship…to parse out who could claim the identity of gun owner and all that it implied politically, socially, even morally," writes Carlson. New gun owners would only be accepted into the community if they adhered to all three of its core tenets, thereby putting sharp boundaries around membership in it—boundaries that are paradoxically harmful to gun sellers materially and, at least in some ways, harmful to the gun rights cause politically.

Through an examination of conservative gun culture, Carlson helps explain the rise, spread, and consequences of armed right-wing populism and the anti-system conspiracy beliefs and extreme partisanship associated with it. She shows not how gun culture has been mobilized on behalf of gun rights, but instead how its adherents view and practice democracy—and how those views and practices have influenced conservative politics more broadly. ■

The reviewer is at the Department of Political Science, Case Western University, Cleveland, OH 44106, USA.
Email: matthew.lacombe@case.edu

10.1126/science.adh2084

HISTORY OF SCIENCE

The era of the intoxicated experimentalist

A historian revisits psychoactive self-experimentation's 19th-century heyday

By **Lucas Richert**

In 1912, the amateur scientist and medical historian Victor Robinson analyzed the effects of hashish on his friends, colleagues, and himself and concluded that cannabis held value as a medical tool and was generally safe. But he also believed that developing rock-solid evidence around potential drugs and medicines was a fraught business. "Old drugs, like old folks," he proclaimed in "An Essay on Hasheesh," published in *Medical Review of Reviews*, "must give way to the new, and even the therapeutic master-builders must beware when the young generation of healing-agents knocks on the door of health."

With *Psychonauts*, author Mike Jay, who is among our finest big-picture analysts and popular historians of global intoxicants, colorfully contextualizes how individuals like Robinson self-experimented with potential "healing agents." Over 300-plus pages, Jay offers the reader an occasionally uncomfortable bird's-eye view of what it took to self-assess the risks, efficacy, and mind-expanding properties of all manner of substances—and how "psychonauts" like Robinson navigated the roiling seas of pharmacological uncertainty, commercial imperatives, and the pitfalls of addiction.

The word "psychonaut" harkens to the 1950s and refers to explorations of the mind and altered states of being, deriving from the roots of "psyche" (spirit or mind) and "nautes" (sailor); yet Jay accurately points to the much longer tradition of self-experimentation. His analysis, set largely in the 19th century in Euro-America, appropriately carries forward into the 20th, terminating in the mid-1970s. In this way, Jay connects Alexander "Sasha" Shulgin with Sigmund Freud, Humphry Davy, Thomas De Quincey, William James, and other renowned psychonauts, providing a throughline from the past to the present.

Jay is on familiar and comfortable territory with *Psychonauts*. In earlier books, *Emperors of Dreams* and *High Society*, Jay exposed readers to the entanglements of drugs, science, and culture during the 19th century. More recently, *Mescaline: A Global History of the First Psychedelic* charted peyote's journey across the planet as an object of ritual, religion, and biomedicine. Jay's *Psychonauts* similarly blends social and biomedical history; the result is a wide-ranging and lavishly illustrated account of multiple drugs and self-experimentation that lacks the paint-by-numbers trappings of many academic monographs.

***Ether Dreams*, an 1868 illustration in a book by Louis Figuier, depicts a vivid ether-induced hallucination.**

The scope of scientific practice is a major feature of *Psychonauts*. How is the science of drugs conducted? Who is doing it? What counts as expertise? To address these questions, over the course of eight chapters, Jay identifies how individuals from all walks of life tested the effects of various mind-altering agents on themselves, and he explores why the methodology of self-experimentation (turning the researcher into both observer and subject) ultimately matters in the history of science.

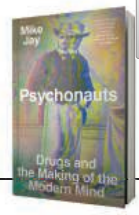
Objective science, we learn, was a product of professionalization and nascent techniques in the mid-1800s, and this so-called progress recalibrated expectations about evidence around intoxicants and medications. As Jay recounts the variety of individuals carrying out drug testing, he transitions expertly from major figures to "their less famous contemporaries," meaning his protagonists come from disparate realms of basic science, medicine, pharmacy, in addition to literature, history, philosophy. It also means that the reader is exposed to previously underexplored historical actors beyond Freud or James. Here, one of the book's few missteps rests in Jay's early vow to share how women and persons of color were present in this history. This promise—save for some scarce instances—goes mostly unfulfilled.

Psychonauts, more generally, arrives at an auspicious moment in the history of intoxicants and drugs, a time when far-reaching regulatory changes are rattling the foundations of global drug policy. With the psychedelic class of drugs, for example, recent years have witnessed an uptick in venture capital investment, increased patent activity and interest from the US Food and Drug Administration and the National Institutes of Health, as well as sustained research activities in the biomedical establishment. While gesturing toward this renaissance, Jay's primary interest is not to advance specific recommendations about psychedelic science in the present; rather, he sketches historical parallels around scientific experimentation with altered consciousness that enhance modern discussions.

Beyond its fascination with the psychedelic community and those steeped in drug and pharmaceutical history, *Psychonauts* has an even wider applicability. It furnishes insights about replication debates in scientific circles, particularly the contemporary reproducibility crisis. It also casts a bright light on the contentious evolution of our mental health crisis and the enduring struggle to develop efficacious psychopharmacological interventions. Lastly, *Psychonauts* lends critical perspectives into how citizen scientists and activists operating on the margins of society or outside the mainstream biomedical establishment may produce beneficial knowledge. To be sure, while *Psychonauts* centers on "old drugs," "the therapeutic master-builders" of the 19th and 20th centuries, and a "young generation of healing-agents," one of its biggest strengths is that it expands our minds in other critical ways. ■

10.1126/science.adh3338

Psychonauts: Drugs and the Making of the Modern Mind
Mike Jay
Yale University Press, 2023.
376 pp.



The reviewer is at the School of Pharmacy, University of Wisconsin, Madison, WI 53705, USA.
Email: lucas.richert@wisc.edu

LETTERS

The invasive fall armyworm (*Spodoptera frugiperda*) has damaged corn crops in the Global South.

PHOTO: BAOQIAN LYU

Biological pest control protects pollinators

Since 2016, the fall armyworm (*Spodoptera frugiperda*) has invaded extensive areas in the Global South (*1*). Native to the Neotropics, this migratory noctuid moth annually colonizes corn fields, where its larvae voraciously feed upon vegetative and reproductive tissues. In Africa alone, *S. frugiperda* inflicts US$9.4 billion of damage per year (*2*). As a result, many small farms have increased their use of insecticides to mitigate pest-induced losses (*1*). Although these chemicals protect corn crops in the short term, they also put pollinators at risk. Biological pest management would protect pollinators and, in turn, long-term food security.

For wild and domesticated bee pollinators, corn pollen and liquids excreted from leaves constitute a substantial, although temporally restricted, resource even in diverse farming landscapes (*3*). Bee colonies rely extensively on these dietary resources during periods of pollen deficiencies in simplified agroecosystems or under extended drought, when floral nectar is scarce (*4*, *5*). In tropical settings with continuous corn cropping, corn pollen features prominently among honeybee foraging resources year-round (*6*).

Insecticide use has been shown to compromise bee-mediated pollination in temperate regions (*7*). Given pollinators' extensive exposure to corn pollen in the tropics and subtropics, insecticide use on corn crops will likely have the same effect. Bee colony health is particularly vulnerable in settings with insufficient sources of pollen or nectar, which can be exacerbated by climate change (*8*). Considering that bees sustain human food security and nutrition (*9*), the cascading One Health impacts of insecticide-centered *S. frugiperda* control cannot be overlooked.

Supplanting insecticides with agroecological and biodiversity-based alternatives can either prevent or control *S. frugiperda* outbreaks (*10*, *11*). Increased usage of manure or compost, cropland diversification, conservation of natural habitat patches, and resistant maize cultivars would counteract pest buildup. Meanwhile, microbial biopesticides, botanicals, and scheduled releases of micro-wasps or insect-killing nematodes can provide curative control (*1*).

Biodiversity-based strategies also strengthen natural pest regulation. Given its protein, carbohydrate, and micronutrient content, corn pollen is a coveted food resource for myriad beneficial arthropods (*12*). In the case of short-lived, annual crops, such as maize, pollen feeding allows carnivorous ladybirds, ants, parasitic wasps, and ground-foraging sheet-web spiders to sustain their populations and provide (cost-free) biological control services. Many of these organisms provide effective *S. frugiperda* biological control in its native habitats in the Americas and are slowly but steadily adapting to this pest in its invasive range (*1*). Replacing chemical insecticides with biodiversity-driven alternatives would uphold societal wellbeing in some of the world's most food-insecure regions.

Baoqian Lyu[1]*, Shuchang Wang[1], Kris A. G. Wyckhuys[2,3,4], Zhuo Liu[1]

[1]Environment and Plant Protection Institute, Chinese Academy of Tropical Agricultural Sciences, Key Laboratory of Integrated Pest Management of Tropical Crops, Ministry of Agriculture and Rural Affairs, Haikou, Hainan 571101, China. [2]Chrysalis Consulting, Danang, Vietnam. [3]Institute for Plant Protection, China Academy of Agricultural Sciences, Beijing, China. [4]School of Biological Sciences, University of Queensland, Saint Lucia, QLD, Australia.
*Corresponding author.
Email: lvbaoqian@hotmail.com

REFERENCES AND NOTES

1. M. Kenis *et al.*, *Entomol. Gen.* 10.1127/entomologia/2022/1659 (2022).
2. R. Eschen *et al.*, *CABI Agricult. Biosci.* **2**, 30 (2021).
3. J. Zawislak, G. Lorenz, J. Adamczyk, R. Wiedenmann, N. K. Joshi, *Environ. Adv.* **5**, 100116 (2021).
4. N. Höcherl, R. Siede, I. Illies, H. Gätschenberger, J. Tautz, *J. Insect. Physiol.* **58**, 278 (2012).
5. G. Di Pasquale *et al.*, *PLOS One* **11**, e0162818 (2016).
6. V. Chaimanee, P. Chantawannakul, K. Khongphinitbunjong, T. Kamyo, J. S. Pettis, *J. Apicult. Res.* **58**, 720 (2019).
7. D. Goulson, E. Nicholls, C. Botías. E. L. Rotheray, *Science* **347**, 1255957 (2015).
8. C. L. Mogren, R. G. Danka, K. B. Healy, *Insects* **10**, 21 (2019).
9. M. R. Smith *et al.*, *Environ. Health. Perspect.* **130**, 127003 (2022).
10. K. Wyckhuys *et al.*, *Curr. Opin. Env. Sust.* **58**, 101208 (2022).
11. J. P. Deguine *et al.*, in *Advances in Agronomy, Volume 178*, D. L. Sparks, Ed. (Academic Press, 2023), pp. 1–59.
12. J. G. Lundgren, *Relationships of Natural Enemies and Non-Prey Foods* (Springer, 2009).

COMPETING INTERESTS

K.A.G.W. is the chief executive officer of Chrysalis Consulting, a firm that provides tailored support to biological control and biodiversity-friendly agriculture.

10.1126/science.adh3467

Effectively implementing biosecurity policies

In her Science Insider story "U.S. scientists brace for greater scrutiny of risky research" (3 February, p. 422), J. Kaiser explains that recently issued recommendations (*1*) from the National Science Advisory Board for Biosecurity (NSABB) would require government oversight for all pathogen research. Although we strongly support the NSABB recommendations, we recognize that the disjointed federal oversight of biosafety and insufficient support could stymie their implementation. To operationalize effective biosecurity oversight systems, the US government must provide regulators and local institutional staff with the appropriate tools and resources.

The existing biosecurity framework includes multiple agencies, including the Department of Health and Human Services (DHHS), that have unique authorities and responsibilities. For the NSABB recommendations to work, the DHHS will need to unify oversight under a single agency, establish an intra-agency biosecurity governance framework, or consider an alternative

biosecurity oversight mechanism. Any new US government biosecurity policies must also account for differences in dual-use procedures within various settings, such as universities, pharmaceutical companies, public health laboratories, and other organizations that use biotechnology. The policy must address the impact on workload and training for staff including biosafety professionals, compliance personnel, and researchers. Because NSABB is recommending expanded oversight for a broader range of experiments, biosecurity reviews will need to be integrated into life sciences research, including privately funded research, on an ongoing basis.

Such changes will require federal and institutional investment. The Federal Select Agent Program (*2*), dual-use (*3*), and pandemic pathogen (*4*) policies are unfunded mandates. To fully assess the fiscal impact of these recommendations, a baseline and periodic financial evaluation needs to be conducted before policies change. Supporting biosafety and biosecurity efforts could include an increased cost ceiling on federal and private grants and dedicated funding at institutions. Any future policy changes should allocate funding for personnel, infrastructure, and compliance operations. In addition, the agencies implementing these recommendations must be funded to cover the costs of training, educational resources, communication, audits, and enforcement.

Moving forward, the NSABB should include more members that have practical experience running institutional biosecurity oversight programs. Members should represent a variety of sectors to ensure a broader perspective in how these policies affect local research enterprises. Before implementation, there should be a thoughtful and thorough review of their impact on a range of stakeholders. As with the development of these NSABB recommendations, every step of the decision-making process should be transparent to the public and concerned citizens should have a voice in the outcome.

David R. Gillum[1]*, Rebecca L. Moritz[2], Antony Schwartz[3]
[1]Department of Environmental Health and Safety, Arizona State University, Tempe, AZ, USA. [2]Office of Research Collaboration and Compliance, Colorado State University, Ft. Collins, CO, USA. [3]Occupational and Environmental Safety Office, Duke University, Durham, NC, USA. *Corresponding author. Email: David.Gillum@asu.edu

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the views or positions of the organizations they represent.

REFERENCES AND NOTES

1. National Science Advisory Board for Biosecurity, "Proposed biosecurity oversight framework for the future of science" (2023); https://osp.od.nih.gov/wp-content/uploads/2023/03/NSABB-Final-Report-Proposed-Biosecurity-Oversight-Framework-for-the-Future-of-Science.pdf.
2. Federal Select Agent Program, "2021 annual report of the Federal Select Agent Program" (2022); https://www.selectagents.gov/.
3. US Government, "United States Government policy for institutional oversight of life sciences DURC" (2014); https://www.phe.gov/s3/dualuse/Pages/InstitutionalOversight.aspx.
4. US Department of Health and Human Services, "Framework for guiding funding decisions about proposed research involving enhanced potential pandemic pathogens" (2017); https://www.phe.gov/s3/dualuse/Documents/P3CO.pdf.

COMPETING INTERESTS

D.R.G. is a past president of American Biological Safety Association (ABSA) International. R.L.M. is current president of ABSA International. A.S. is on the ABSA International Council.

10.1126/science.adh5519

China bans electric capture of earthworms

The earthworm is widely distributed, with 4000 species in the world and more than 300 species in China (*1*, *2*), where they are in demand as fishing bait, livestock feed, and components of traditional medicine (*3*). In recent years, poachers in China have started using soil electrocution to capture earthworms, putting earthworm populations and ecosystems at risk. In February, China prohibited the practice. The decision is a necessary first step to protect agriculture, but to ensure that electric capture ceases, China must follow up with legislation to support the implementation of the ban.

Known as "ecosystem engineers," earthworms play a vital role in the biogeochemical cycle (*4*). They can affect the biological, chemical, and physical processes of the ecosystem through feeding, digestion, excretion, and burrowing. The earthworm population, which is closely related to soil microorganisms, soil animals, and soil enzyme activities, can increase soil organic matter, improve soil structure, and increase soil fertility (*5*).

In the past, earthworms in China were captured manually, which limited the number that could be caught in a short amount of time and did not threaten the earthworm population or the soil ecosystem. In contrast, applying electric current to the soil can catch about 150 kg of earthworms in a day (*6*, *7*). Removing earthworms at this scale threatens the species with local extinction and robs the targeted ecosystem of the benefits they provide (*8*).

This capture method threatens other species as well. The electric current may kill or harm soil organisms beyond earthworms. The changes could disrupt the food cycle for birds, arthropods, and mammals that depend on earthworms as a food source (*9*, *10*). The sharp decrease of earthworms in the soil may also affect soil fertility and productivity, reducing crop yields (*11*).

Electric earthworm capture has quickly become widespread. Because earthworms are not included on the list of protected wild animals in China, there have been few legal ramifications for using the approach, despite the damage it causes. In February, the Chinese government took a step toward better agricultural stewardship when it unveiled its agriculture plan (*12*), which explicitly prohibits electric capture of earthworms, poaching of phaeozem, and other activities that damage the soil ecosystem. However, the financial benefits of large-scale earthworm poaching may provide poachers with an incentive to continue using electric capture despite the ban. To implement the ban effectively, China should pass legislation that specifies how poachers will be penalized if they continue to use soil electrocution. In addition, law enforcement agencies should be authorized to take action if they find poachers disregarding the policy.

Jiangwei Zhu[1]*, Li Fu[2], Meixia Guo[3], Min Wang[4], Yan Zhao[5]
[1]Co-Innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing 210037, China. [2]College of Materials and Environmental Engineering, Hangzhou Dianzi University, Hangzhou 310018, China. [3]Henan Joint International Research Laboratory of Environmental Pollution ControlMaterials, College of Chemistry and Chemical Engineering, Henan University, Kaifeng 475004, China. [4]Institute of Process Equipment and Environmental Engineering, Northeastern University, Shenyang 110189, China. [5]Shanghai Institute of Quality Inspection and Technical Research, Shanghai 200233, China.
*Corresponding author.
Email: 2686355689@qq.com

REFERENCES AND NOTES

1. C. A. Edwards, *Earthworm Ecology* (CRC Press, ed. 2, 2004).
2. J. Huang *et al*., *J. China Agric. Univ*. **11**, 9 (2006).
3. P. Marco *et al*., *Sci. Tot. Environ.* **734**, 139460 (2020).
4. R. Xiao *et al*., *Environ. Int*. **158**, 106924 (2022).
5. A. M. Salwan, K. A. A. T. Duraid, Z. A. M. Huda, *Acta Ecol. Sin.* **41**, 512 (2021).
6. Q. Li, "The profit chain of wild earthworms," *Xinhuanet* (2022); http://www.news.cn/2022-07/20/c_1128846404.htm [in Chinese].
7. Y. Guan, "Catching earthworms by electricity is unbearable to the land," *Beijing Justice Network Media* (2022); https://baijiahao.baidu.com/s?id=1735662464697469185&wfr=spider&for=pc [in Chinese].
8. Q. Wu, "Use electricity to catch earthworms," *Kepu Times* (2022); https://baijiahao.baidu.com/s?id=1760714762233736441&wfr=spider&for=pc [in Chinese].
9. G. G. Brown, *Plant Soil* **170**, 209 (1995).
10. S. Datta *et al*., *Environ. Sci. Pollut. Res*. **23**, 8227 (2016).
11. C. A. Edwards *et al*., *Biology and Ecology of Earthworms* (Springer, 2022).
12. Y. Lu, "No. 1 central document explicitly prohibits electric capture of earthworms," *Red Star News* (2023); https://www.thepaper.cn/newsDetail_forward_21945235 [in Chinese].

10.1126/science.adh3383

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

PHOTO: RACHEL HOLSER, NMFS 23188

Although northern elephant seals do sleep on land, like the one pictured here in California, they can also sleep while diving to 300 meters underwater.

SLEEP

Sleeping deep

Sleep is essential, but not all mammals live in environments where long periods of time asleep are possible. Marine mammals encounter especially challenging conditions for sleep when they are at sea. Using advanced remote monitoring techniques, Kendall-Bar *et al*. found that wild northern elephant seals can sleep for less than 2 hours per day at sea and do so while diving to depths of around 300 meters. Unlike other marine mammals, they enter full REM sleep, with accompanying paralysis, but they do so at depths below those occupied by their predators. —SNV *Science*, adf0566, this issue p. 260

STRUCTURAL BIOLOGY

A retrotransposon zooms in on its target

Much of the human genome consists of repetitive sequences called transposons that can copy and paste themselves into new DNA locations. The most common, and only demonstrably active, type of transposon in humans is the non–long terminal repeat retrotransposon, which uses an RNA intermediate that is converted into DNA directly at the site of a new insertion. Wilkinson *et al*. used cryo–electron microscopy to show what determines the sequence that is inserted and where it is inserted. They also show that the specificity can be altered, suggesting that these retrotransposons could be engineered as gene insertion tools. —DJ

Science, adg7883, this issue p. 301

CANCER EVOLUTION

A tale of two cancers

Already threatened by habitat loss and past persecution, Tasmanian devil populations have been decimated over the past few decades by a fatal transmissible cancer referred to as DFT1. Some recent research has suggested that the growth and spread of this cancer has slowed, and it may have become endemic, leading to lower mortality. However, Stammnitz *et al*. describe the evolution of a second transmissible facial cancer in Tasmanian devils, DFT2. This lineage evolved more recently, mutates more rapidly, and is a high-level threat to this species. —SNV

Science, abq6453, this issue p. 283

PROTEIN DESIGN

Complex architectures by top-down design

Designing artificial protein complexes based on components with established structural properties can impose limits on the properties of the final complex. Lutz *et al*. developed methods for generating complex protein architectures that adhere to preordained parameters using reinforcement learning. They demonstrated this approach by generating designs that fill arbitrary volumes, including a symmetrical connector between previously designed protein rings. A small protein designed to assemble into 60-subunit icosahedra may be useful for presenting antigens in vaccines or signaling molecules in multivalent agonist complexes, as the authors demonstrate in preliminary biological experiments. —MAF

Science, adf6591, this issue p. 266

SEXUAL SELECTION

A passing advantage

Female choice plays a large role in shaping populations. Across many species, females have been shown to prefer males with traits that are rare or uncommon. How this preference is maintained over time has remained an open question.

Potter *et al*. looked across generations in Trinidadian guppies and found that females do have a clear preference for rare males, and that they acquire a further fitness benefit through sons that also have these rare traits. Once rare traits become more common, however, this fitness benefit dissipates such that rare traits in the father eventually become common, leaving the grandsons to be less preferred. —SNV

Science, ade5671, this issue p. 309

OPTIC FIBERS

Twist and confinement

Optical fibers form the backbone of the modern information age, with information encoded in various properties of light: wavelength, polarization, intensity, and phase. These modes rely on total internal reflection within the fiber. To further increase capacity, propagating light in different spatial modes within the fiber is possible, but this is complicated by mode mixing, in which the information is corrupted. Using light with orbital angular momentum, Ma *et al*. demonstrate an optical fiber designed with a topological feature that confines the light separate from that of total internal reflection, thus avoiding the detrimental issues related to mode mixing. These results suggest a route for developing higher-capacity optical networks. —ISO

Science, add1874, this issue p. 278

Representative modes in an optical fiber illustrated as spiral patterns

MATERIALS SCIENCE

The secret of ancient Maya masonry

The addition of organic materials in the making of lime plasters is key to achieving the remarkable stability of pre-Columbian mortars and stuccoes despite exposure to a hot and humid tropical environment. Studying Maya plasters from Copán, Honduras, Rodriguez-Navarro *et al.* found organic materials and nanostructures of calcite biominerals analogous to those found in seashells. By testing lime plaster replicas in which bark extracts from local trees in Copan were added, the authors showed that the presence of these organic additives enhanced the toughness and weather resistance of the plasters. Optimized lime-based plasters may be useful for heritage conservation and modern sustainable construction. —JKD

Sci. Adv. (2023) 10.1126/sciadv.adf6138

INFLUENZA

Fighting flu

A major goal for influenza vaccines is to elicit broadly reactive immune responses that can protect against many strains of the virus. Two studies report results of a phase 1 clinical trial testing a vaccine that may get closer to this goal. Widge *et al*. demonstrated that immunization with one or two doses of an H1 hemagglutinin–stabilized stem nanoparticle (H1ssF) vaccine was safe in recipients and elicited durable neutralizing antibody responses. Andrews *et al*. found that memory B cell responses elicited by H1ssF vaccination were broadly cross-reactive and targeted two conserved epitopes on the H1 stem. These studies highlight the potential of H1ssF and similar stem-only immunogens as influenza vaccines. —CSM

Sci. Transl. Med. (2023) 10.1126/scitranslmed.ade4790, 10.1126/scitranslmed.ade4976

FRAMEWORK MATERIALS

Stabilizing COFs with NO

The assembly of covalent organic frameworks (COFs) relies on reversible covalent bond formation to heal any structural defects, but such bonds could also lead to low stability in sorption applications, especially if reactive gases such as nitric oxide (NO) are being adsorbed. Emmerling *et al*. found that NO exposure actually passivated unreacted terminal amines in COFs through denitrogenation. Whereas imide and thiazole linkages were largely unreactive, NO caused conversion of imine and amine linkages that changed the COF surface area and crystallinity. The amine linkages could be converted to *N*-diazeniumdiolates ("NONOate") linkages at room temperature. These modified COFs released NO under physiological solution conditions and offer a platform for the delivery of NO in vivo as a signaling molecule. —PDS

J. Am. Chem. Soc. (2023) 10.1021/jacs.2c11967

NEURODEGENERATION

The hormonal X factor

Postmenopausal women comprise the majority of Alzheimer's disease patients (70%). This population has been shown to have increased toxic neuronal tau protein deposition compared with age-matched men. However, how sex, age, and hormone replacement therapy (HRT) affect tau deposition in cognitively unimpaired individuals remains unclear. Using tau imaging in cognitively normal postmenopausal women and age-matched men, Coughlan *et al.* showed that women had higher tau deposition in several brain areas. Early

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

MATERIALS SCIENCE

Improving water treatment membranes

Converting brackish or waste water to clean drinking water requires semipermeable reverse osmosis membranes that permit water flow while retaining contaminants. Despite research that claims to improve on existing membrane materials, the widely used gold standard reverse osmosis membrane has not changed in decades. In a Perspective, McCutcheon and Mauter discuss the problems with translating the development of new materials to commercial application. In particular, they highlight the need to contextualize materials development within the entire water treatment process, test new materials under relevant conditions, and assess manufacturability. Changing evaluations in this way could improve the application of materials science to the treatment of contaminated water. —GKA

Science, ade5313, this issue p. 242

SUSTAINABILITY

Facing coupled environmental crises

Humanity is facing major social and ecological impacts from climate change and biodiversity loss. These two crises are intertwined, with common causes and effects on one another. Pörtner *et al*. review the results of a joint meeting of members of the International Panels on Climate Change and Biodiversity and Ecosystem Services. They discuss the connections between biodiversity loss and climate change and propose potential solutions for addressing them as interconnected problems. Drastic reductions in greenhouse gas emissions, protection of multiuse landscapes and seascapes, and policies for providing equitable access to natural resources can help to ensure future ecological function and human well-being. —BEL

Science, abl4881, this issue p. 256

GENOME EDITING

Base editing in a single treatment

Spinal muscular atrophy is the leading genetic cause of infant death. It arises from the lack of a protein called survival motor neuron (SMN). Drugs that increase SMN are effective but require repeated dosing or may fade over time. Arbab *et al*. identified genome-editing strategies that permanently correct SMN protein levels to normal levels by converting a partially active gene encoding SMN into a fully active form. In a mouse model, treatment with base editors that efficiently and precisely make this change increased life span and rescued motor function. A one-time combination treatment of a base editor and a current spinal muscular atrophy drug further improved outcomes in mice. —DJ

Science, adg6518, this issue p. 257

CANCER

RNA surveillance turns oncogenic

Precise RNA regulation is a quality control mechanism that is important for normal development and to prevent disease. Aberrant RNAs require identification and destruction to avoid translation of defective proteins. Insco *et al*. report that cyclin-dependent kinase 13 (CDK13), which activates an RNA surveillance mechanism to degrade abnormal RNAs, also has a tumor-suppressor function (see the Perspective by Fisher). When CDK13 was mutated in an animal melanoma model, accumulation and translation of aberrant RNAs resulted in more aggressive malignancy. Analysis of other cancer types revealed similar CDK13 mutations and showed that additional RNA surveillance genes were recurrently mutated in human tumors. These findings suggest that RNA surveillance may have a previously unrecognized tumor-suppressive role. —PNK and YN

Science, abn7625, this issue p. 258; see also adh4051, p. 240

METABOLISM

Linking AMPK to organelle biogenesis

The kinase AMPK is a key sensor that helps to control energy homeostasis. Malik *et al*. reveal the mechanism by which AMPK controls the transcription factor TFEB to increase gene transcription and to support mitochondrial and lysosomal biogenesis. AMPK appears to act by direct phosphorylation of folliculin-interacting protein 1 (FNIP1). FNIP is part of a complex that acts as a GTP-activating protein for the GTPases RagC and RagD, which regulate the mechanistic target of rapamycin complex 1 protein kinase signaling complex on the lysosomal surface. This results in release of TFEB from the lysosome, allowing it to act at the nucleus. —LBR

Science, abj5559, this issue p. 259

QUANTUM MECHANICS

Schrödinger's cats, kittens, and lions

The idea of Schrödinger's cat being both alive and dead at the same time—its fate revealed only upon inspection—came from a thought experiment that pointed out an absurdity in the interpretation of quantum mechanics at the time. However, because such superposition states have now been prepared in many different quantum systems, the question is where do the classical and quantum worlds part company? Bild *et al*. prepared, observed, and controlled cat states of a 16-microgram mechanical resonator. Being able to control the size of the superposition states, they effectively created a menagerie of quantum states, thus providing a platform to explore the boundary between the quantum and classical behavior. —ISO

Science, adf7553, this issue p. 274

NEUROEVOLUTION

View of an ancient brain

The evolutionary origin of nervous systems remains a fundamental question in biology. A hallmark of nervous systems is that they are composed of discrete cells (neurons) that communicate through synapses. Ctenophores, a sister group to all animals with nervous systems, play a key role in comparative studies into the evolutionary origin(s) of neurons and their connections. To establish neuronal circuits that facilitate ctenophore behavior, Burkhardt *et al.* used high-resolution three-dimensional electron microscopy, revealing that nerve-net neurons are not separate entities, but rather are interconnected through continuous neurite plasma membranes without evidence of synapses (see the Perspective by Dunn). The findings offer a new perspective on the evolution of neuronal networks and neurotransmission. —MMa

Science, ade5645, this issue p. 293; see also adh0542, p. 241

CONSERVATION

Private land protection

To protect high rates of biodiversity, a large amount of global land must be under some sort of protection. In regions where public land is not prioritized nor widely distributed, it is possible that protection of private land could contribute to species conservation. In Brazil, a native vegetation law instituted decades ago has provided an opportunity for evaluation of the

role of private land in conserving species. De Marco *et al*. looked at mammal species protected by these private set-asides in the Brazilian Cerrado and found that they covered up to 25% of species ranges (see the Perspective by Machado and Aguiar). Such areas play an even more important role when ecologically intact or restored. —SNV

Science, abq7768, this issue p. 298
see also adh1840, p. 238

IMMUNOLOGY

The master of MAIT cell metabolism

Mucosal-associated invariant T (MAIT) cells are a type of innate-like T cells that recognize bacterial metabolites. Upon activation, they proliferate and produce cytokines to promote host defense. Kedia-Mehta *et al*. show that MAIT cell proliferation depends on MYC-associated pathways involving amino acid transport and glycolysis. Furthermore, MAIT cells from patients with obesity show disrupted function and engagement of these pathways. These metabolic pathways may be relevant to the development of MAIT cell–based therapies. —AEB

Sci. Signal. (2023)
10.1126/scisignal.abo2709

NEUROIMMUNOLOGY

Dispensable dural lymphatics

Lymphatic vessels in the dura, the outermost layer of the meninges, provide a vascular path for immune cells connecting the meninges with the systemic circulation. Dural lymphatics have been proposed as a gateway that T cells targeting CNS autoantigens use to access the brain and spinal cord. The formation and maintenance of dural lymphatics can be abrogated by genetic or pharmacologic interference with vascular endothelial growth factor C (VEGF-C) or its receptor, VEGFR3. Li *et al*. found that atrophy of dural lymphatics by VEGFR3 blockade in mice was insufficient to block autoimmune neuroinflammation initiated by immunization with myelin autoantigens or transfer of encephalitogenic T cells. These findings suggest that therapies aimed at disrupting dural lymphatics are unlikely to attenuate human autoimmune neuroinflammatory diseases such as multiple sclerosis. —IRW

Sci. Immunol. (2023)
10.1126/sciimmunol.abq0375

Potter *et al*. looked across generations in Trinidadian guppies and found that females do have a clear preference for rare males, and that they acquire a further fitness benefit through sons that also have these rare traits. Once rare traits become more common, however, this fitness benefit dissipates such that rare traits in the father eventually become common, leaving the grandsons to be less preferred. —SNV

Science, ade5671, this issue p. 309

OPTIC FIBERS

Twist and confinement

Optical fibers form the backbone of the modern information age, with information encoded in various properties of light: wavelength, polarization, intensity, and phase. These modes rely on total internal reflection within the fiber. To further increase capacity, propagating light in different spatial modes within the fiber is possible, but this is complicated by mode mixing, in which the information is corrupted. Using light with orbital angular momentum, Ma *et al*. demonstrate an optical fiber designed with a topological feature that confines the light separate from that of total internal reflection, thus avoiding the detrimental issues related to mode mixing. These results suggest a route for developing higher-capacity optical networks. —ISO

Science, add1874, this issue p. 278

Representative modes in an optical fiber illustrated as spiral patterns

MATERIALS SCIENCE

The secret of ancient Maya masonry

The addition of organic materials in the making of lime plasters is key to achieving the remarkable stability of pre-Columbian mortars and stuccoes despite exposure to a hot and humid tropical environment. Studying Maya plasters from Copán, Honduras, Rodriguez-Navarro *et al.* found organic materials and nanostructures of calcite biominerals analogous to those found in seashells. By testing lime plaster replicas in which bark extracts from local trees in Copan were added, the authors showed that the presence of these organic additives enhanced the toughness and weather resistance of the plasters. Optimized lime-based plasters may be useful for heritage conservation and modern sustainable construction. —JKD

Sci. Adv. (2023) 10.1126/sciadv.adf6138

INFLUENZA

Fighting flu

A major goal for influenza vaccines is to elicit broadly reactive immune responses that can protect against many strains of the virus. Two studies report results of a phase 1 clinical trial testing a vaccine that may get closer to this goal. Widge *et al*. demonstrated that immunization with one or two doses of an H1 hemagglutinin–stabilized stem nanoparticle (H1ssF) vaccine was safe in recipients and elicited durable neutralizing antibody responses. Andrews *et al*. found that memory B cell responses elicited by H1ssF vaccination were broadly cross-reactive and targeted two conserved epitopes on the H1 stem. These studies highlight the potential of H1ssF and similar stem-only immunogens as influenza vaccines. —CSM

Sci. Transl. Med. (2023) 10.1126/scitranslmed.ade4790, 10.1126/scitranslmed.ade4976

FRAMEWORK MATERIALS

Stabilizing COFs with NO

The assembly of covalent organic frameworks (COFs) relies on reversible covalent bond formation to heal any structural defects, but such bonds could also lead to low stability in sorption applications, especially if reactive gases such as nitric oxide (NO) are being adsorbed. Emmerling *et al*. found that NO exposure actually passivated unreacted terminal amines in COFs through denitrogenation. Whereas imide and thiazole linkages were largely unreactive, NO caused conversion of imine and amine linkages that changed the COF surface area and crystallinity. The amine linkages could be converted to *N*-diazeniumdiolates ("NONOate") linkages at room temperature. These modified COFs released NO under physiological solution conditions and offer a platform for the delivery of NO in vivo as a signaling molecule. —PDS

J. Am. Chem. Soc. (2023) 10.1021/jacs.2c11967

NEURODEGENERATION

The hormonal X factor

Postmenopausal women comprise the majority of Alzheimer's disease patients (70%). This population has been shown to have increased toxic neuronal tau protein deposition compared with age-matched men. However, how sex, age, and hormone replacement therapy (HRT) affect tau deposition in cognitively unimpaired individuals remains unclear. Using tau imaging in cognitively normal postmenopausal women and age-matched men, Coughlan *et al.* showed that women had higher tau deposition in several brain areas. Early

PLANT BREEDING

Ethiopian farmers inform wheat breeding

Local farming practices are best understood by the farmers themselves. Gesesse *et al.* combined Ethiopian farmers' knowledge with genetic data on durum wheat to assist breeding. Smallholder farmers from three sites were skilled at assessing and differentiating among durum wheat varieties. Farmer preferences varied in relation to yield, morphology, and phenology depending on locality. Preference scores were used in genomic selection and association analyses to identify alleles underlying performance. Some loci colocated with yield, but others were independent, indicating that they may relate to alternative traits such as those relating to potential for local adaptation. —MRS

Proc. Natl. Acad. Sci. U. S. A. (2023)
10.1073/pnas.2205774119

Detail of a wheat field in the highlands near Lalibela, Amhara Region, Ethiopia

menopause and late initiation of HRT were associated with increased neuronal tau compared with late menopause and early HRT. —MMa

JAMA Neurol. (2023)
10.1001/jamaneurol.2023.0455

IMMUNE ADAPTATION

Old selection affects modern health

The immune system is a constant and repeated target of natural selection. Kerner *et al.* used a wide array of ancient and modern European DNA to assess selection on immune-related alleles in humans. The authors modeled allele frequency trajectories for over 1 million variants, finding 89 independent loci predicted to be under positive selection since the Neolithic period. Many of these variants were associated with hematopoietic traits such as platelet and reticulocyte counts, autoimmune disease, and infection. This finding indicates that these traits have been directly affected by selection in Europeans and supports the idea that increased protection from infectious disease may concomitantly increase the risk for autoimmunity. —CNS

Cell Genom. (2023)
10.1016/j.xgen.2022.100248

MICROBIOTA

Not all fiber is the same

A healthy diet, we are told, demands a high fiber intake to sustain a healthy microbiota. However, fiber is a heterogenous substance, and its composition is likely to influence the range of microorganisms that ferment it. Using fecal batch cultures, Solvang *et al.* tested microbial growth on a variety of edible plant extracts. At 72 hours, fermentation products, bacterial abundance, and community composition were measured. More complex substrates supported a more variable community, which also varied by fiber source. For example, high arabinan from beetroot and high galactan from carrots predicted specific microbial enrichment. Thus, knowledge of dietary fiber composition will be valuable in designing diets that optimize desirable composition and function in a patient's microbiota. —CA

Environ. Microbiol. (2023)
10.1111/1462-2920.16368

MATERIALS SCIENCE

Recovering metals from lithium-ion batteries

With the growth in electric vehicles and grid storage, there is a need for methods to efficiently recover valuable metals from lithium-ion batteries once they reach their end of useful life. Quintana *et al.* developed a liquid extraction process that uses rotating reactors to stir, separate, and capture metal ions. Using either a vertical or horizontal configuration, the reactors contain a higher pH aqueous feed and a lower pH aqueous acceptor phase that are kept separate by an inorganic solvent. The organic layer contains low concentrations of extractants that shuttle manganese and cobalt from the feed, and because they will cycle multiple times, much less of these elements is needed compared with conventional extraction techniques. —MSL

Adv. Mater. (2023)
10.1002/adma.202211946

SCIENCE EDUCATION

An arbitrary gatekeeper

Another blow is delivered against the predictive power of the Graduate Record Examination (GRE). Feldon *et al.* performed a meta-analysis to assess the utility of the GRE as a predictor across outcome variables for graduate students in the United States. The team also assessed changes in observed effects over time as related to increasing diversity in the graduate student population. Their results showed that 61.6% of the reported effects showed no predictive value of GRE scores on student outcomes. Sample composition effects by race or ethnicity were notable but nonsignificant, with increasing proportions of people of color within a study sample associated with poorer predictive validity, a result that should be of great interest to institutions working to increase diversity. This study provides additional evidence that removing GRE requirements for graduate applications eliminates a source of unnecessary inequity. —MMc

J. High. Educ. (2023)
10.1080/00221546.2023.2187177

REVIEW SUMMARY

SUSTAINABILITY

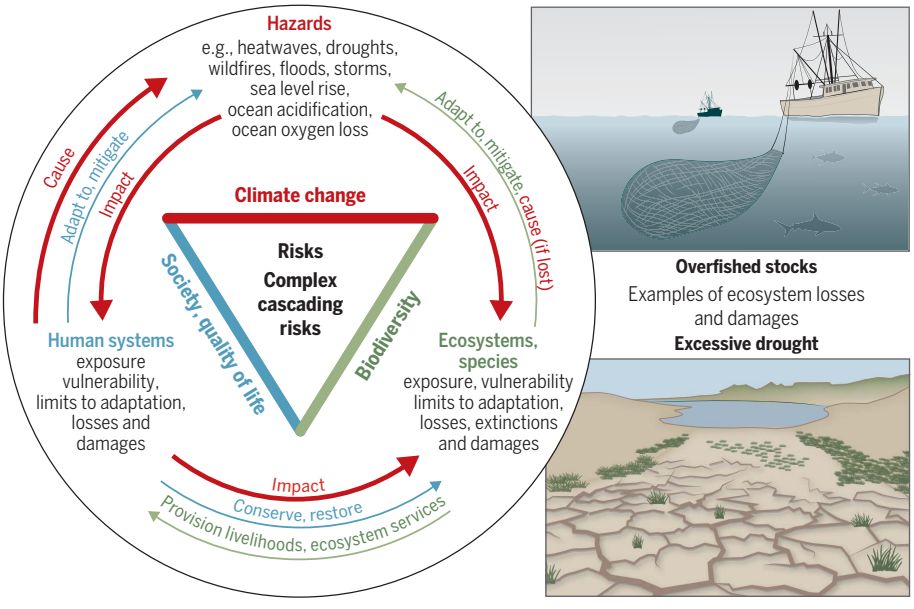
Overcoming the coupled climate and biodiversity crises and their societal impacts

H.-O. Pörtner*, R. J. Scholes†, A. Arneth, D. K. A. Barnes, M. T. Burrows, S. E. Diamond, C. M. Duarte, W. Kiessling, P. Leadley, S. Managi, P. McElwee, G. Midgley, H. T. Ngo, D. Obura, U. Pascual, M. Sankaran, Y. J. Shin, A. L. Val

BACKGROUND: Two intertwined crises threaten human well-being: Climate change—arising from human-induced greenhouse gas emissions, including those from the loss of biomass and biodiversity—is raising temperatures beyond those of the Holocene, when human civilization evolved and expanded globally. Mean warming and the increasing frequency and severity of extreme events in turn disturb ecosystem functioning, cause habitat loss to humans and biodiversity, and exacerbate the unprecedented loss of biodiversity already caused by human-induced habitat degradation, overexploitation of natural resources, and pollution. Both crises reduce nature's contributions to people, which sustain well-being, livelihoods, economies, and development prospects and also support climate change adaptation and mitigation. Failing to act will increase human vulnerability, including poverty, food insecurity, involuntary displacement, and political instability and conflict. The coupled global climate and biodiversity crises and their societal impacts concern land, freshwater, and ocean ecosystems alike but are insufficiently tackled by current actions, as identified by assessments of both the IPCC (Intergovernmental Panel on Climate Change) and IPBES (Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services). None of the 20 2011–2020 Aichi biodiversity targets and none of the mileposts on climate trajectories intended to limit warming to 1.5°C have been met.

ADVANCES: Climate, biodiversity, and societal challenges are intertwined but are often treated as singular problems. Solutions exist with cobenefits across sectors. In intact and functional ecosystems, nature is efficient at carbon capture (by photosynthesis) and sequestration (long-term removal from the carbon cycle), provided that warming is limited to 1.5°C or below through ambitious emissions reductions. Strengthening the biosphere on land, in fr… water, and in the ocean will support cli… change mitigation, adaptation, biodiversity, human well-being, and livelihoods. Mounting scientific evidence points to the need to prioritize protection of remaining undamaged carbon- and species-rich environments and to implement targeted restoration projects, with more attention to effectively sustaining biodiversity and fairly distributed societal cobenefits. Three critical objectives for future spatial planning include a habitable climate, self-sustaining biodiversity, and sustained provisioning of nature's contributions to people to support development and a good quality of life for all. Coordinated efforts among science and policy can identify and help navigate development pathways toward climate resilience for both human society and biodiversity.

OUTLOOK: New global biodiversity, climate, and sustainability targets envisioned for 2030 and 2050 will likely fail if drivers behind climate change and biodiversity loss remain insufficiently addressed or concrete actions to meet current political agreements and goals do not increase in pace and scale. The following actions are urgently needed. (i) Ambitious emissions reduction, combined with suitable adaptation measures. (ii) Effective protection of an average of 30 to 50% of surface areas across a mosaic of interspersed and interconnected land, ocean, and freshwater "scapes." These cover a gradient, from pristine ecosystems; to spaces shared by humans and wild species and sustainably used; to spaces under intensive uses such as cities, which nonetheless can harbor substantial biodiversity in terrestrial and aquatic spaces. Efforts need to consider the specific spatial demands for healthy ecosystems. (iii) Building of development pathways and the underpinning of political, economic, and social institutions (including norms and rules) on visions such as collective responsibility, sustainable and circular uses of natural resources, avoidance of overconsumption and waste, and more equitable and participatory development regionally and globally. (iv) The enabling of just and equitable access to and benefits from natural assets across societies, groups, and individuals, securing good quality of life. Transformative action can overcome siloed approaches through institutional and individual change, achieving sustainability for nature and people, as well as human, ecosystem, and planetary health. ■



Climate, biodiversity, and human society are coupled through dynamic interactions across scales. Human-caused exploitation and climate change are increasingly threatening biodiversity and nature's contributions to people, causing losses and damages exemplified through overfished stocks and excessive drought that harm productive habitats.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: hans.poertner@awi.de
†Deceased.
Cite this article as H.-O. Pörtner *et al.*, *Science* **380**, eabl4881 (2023). DOI: 10.1126/science.abl4881

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abl4881

REVIEW

SUSTAINABILITY

Overcoming the coupled climate and biodiversity crises and their societal impacts

H.-O. Pörtner[1,2]*, R. J. Scholes[3]†, A. Arneth[4], D. K. A. Barnes[5], M. T. Burrows[6], S. E. Diamond[7], C. M. Duarte[8,9], W. Kiessling[10], P. Leadley[11], S. Managi[12], P. McElwee[13], G. Midgley[14], H. T. Ngo[15,16], D. Obura[17,18], U. Pascual[19,20,21], M. Sankaran[22], Y. J. Shin[23], A. L. Val[24]

Earth's biodiversity and human societies face pollution, overconsumption of natural resources, urbanization, demographic shifts, social and economic inequalities, and habitat loss, many of which are exacerbated by climate change. Here, we review links among climate, biodiversity, and society and develop a roadmap toward sustainability. These include limiting warming to 1.5°C and effectively conserving and restoring functional ecosystems on 30 to 50% of land, freshwater, and ocean "scapes." We envision a mosaic of interconnected protected and shared spaces, including intensively used spaces, to strengthen self-sustaining biodiversity, the capacity of people and nature to adapt to and mitigate climate change, and nature's contributions to people. Fostering interlinked human, ecosystem, and planetary health for a livable future urgently requires bold implementation of transformative policy interventions through interconnected institutions, governance, and social systems from local to global levels.

Climate change and biodiversity loss arguably represent the most substantial challenges to ecosystem health and supply of nature's contributions to people (NCP) (*1*). They also directly and indirectly affect human health and well-being and the functioning and sustainability of human societies at the global scale (*2*, *3*). Increases in greenhouse gas emissions now exceed 55 billion tonnes of carbon dioxide equivalents (Gt CO_2e) year^{-1} and have resulted in warming by more than 1.1°C above preindustrial times, altered precipitation regimes, sea level rise, more extreme weather events, and oxygen depletion and acidification of aquatic environments (*4*). Climate change and other anthropogenic drivers of biodiversity loss—such as agricultural expansion and intensification, pollution, overfishing, and invasive alien species introductions—independently and together stress and disrupt ecological communities (*2*). Human activities have modified nearly 75% of Earth's land surface and 66% of its ocean area (*3*, *5*), leading to the loss of more than 80 and 50% of wild mammal and plant biomass, respectively (*6*), and threatening more species with extinction than ever before in human history (*7*). Moreover, climate change and biodiversity loss are inextricably connected, with each stressor contributing to and exacerbating the effects of the other.

Climate change is causing species shifts and biodiversity loss globally because species are confined to limited thermal performance ranges and therefore geographical ranges. Climate warming increasingly causes species to move (Fig. 1, A and C) to higher latitudes (poleward), higher altitudes (on land), or deeper water. Range shifts lead to changes in species interactions, with cascading effects on species abundances, species composition, and ecological functions. When species cannot track the changing climate or when suitable habitat shrinks, warming leads to mortality and biodiversity loss, as has been observed in the tropics (Fig. 1B). Some species in coral reefs, savannahs, rain forests, high-latitude and -altitude ecosystems, and Mediterranean systems show evidence of exceedance of tolerance and adaptation limits (*2*). Mountains and continental boundaries may become climate dead ends because climate belts shift off the top of mountains, at the ends of a continent, or on islands. Extinction risks are high for species with narrow geographic ranges or specific habitat requirements, and those in the tropics and in polar regions (*2*). Climate change surpassing species' adaptation limits constrains options to protect those species from extinction (*2*). Overall, climate change and sea level rise are projected to exacerbate the direct impacts of human activities, causing further losses in biomass, habitats, and species (*2*, *8*–*10*).

Conversely, biodiversity loss contributes to climate change through loss of wild species and biomass. This reduces carbon stocks and sink capacity in natural and managed ecosystems, increasing emissions (*2*, *4*). Habitat loss and disturbance by human activities (such as through deforestation or expansion of livestock production or pollution) are the primary cause of biodiversity loss, increasingly exacerbated by climate change. Feedbacks of biodiversity loss on climate and of climate change on biodiversity include temperature-induced changes in photosynthetic capacity and carbon storage, modified reflectivity of the land surface, altered formation of clouds and atmospheric dust, and shifted biogeochemical cycling of nutrients and carbon, which in turn influence the concentration of greenhouse gases in the atmosphere (*11*–*16*).

Biodiversity loss and climate change are thus both drivers and consequences of one another, tightly linked to human activities or demographic change and resulting in negative impacts on NCP, human health and well-being, as well as societal functioning [for example, (*17*, *18*)]. Further, the human species is also losing habitat, through sea level rise and, more widely, through prolonged exposures to lethal temperature-humidity combinations outside humans' thermal niche, especially at but not limited to the lower latitudes (*2*). Yet these interlinkages are not sufficiently recognized in policies and multilateral agreements to limit climate change and halt biodiversity loss and societal consequences, particularly as the United Nations Framework Convention on Climate Change (UNFCCC) and the Convention on Biological Diversity (CBD) have largely addressed these issues independently. Policies and actions that simultaneously address climate, biodiversity, and society (the "nexus") can help to avoid dangerous trade-offs and maximize cobenefits for humankind (*18*).

Accordingly, various aspects of the "solution space" are developed in the following sections. The discussion includes (i) strategies to

[1]Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany. [2]Department of Biology and Chemistry, University of Bremen, Bremen, Germany. [3]Global Change Institute, University of the Witwatersrand, Johannesburg, South Africa. [4]Atmospheric Environmental Research, Karlsruhe Institute of Technology (KIT), Garmisch-Partenkirchen, Germany. [5]British Antarctic Survey, Natural Environment Research Council, Cambridge, UK. [6]Scottish Association for Marine Science, Oban, Argyll, UK. [7]Department of Biology, Case Western Reserve University, Cleveland, OH, USA. [8]Red Sea Research Centre (RSRC), King Abdullah University of Science and Technology, Thuwal, Saudi Arabia. [9]Computational Bioscience Research Centre (CBRC), King Abdullah University of Science and Technology, Thuwal, Saudi Arabia. [10]Geozentrum Nordbayern, Friedrich-Alexander-Universität, Erlangen, Germany. [11]Laboratoire d'Ecologie Systématique Evolution, Université Paris-Saclay, CNRS, AgroParisTech, 91400 Orsay, France. [12]Urban Institute, Kyushu University, Fukuoka, Japan. [13]Department of Human Ecology, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA. [14]Global Change Biology Group, Botany and Zoology Department, University of Stellenbosch, 7600 Stellenbosch, South Africa. [15]Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES), Bonn, Germany. [16]Food and Agriculture Organization of the United Nations, Viale delle Terme di Caracalla, Rome, Italy. [17]Coastal Oceans Research and Development–Indian Ocean (CORDIO) East Africa, Mombasa, Kenya. [18]Global Climate Institute, University of Queensland, Brisbane, QLD 4072, Australia. [19]Basque Centre for Climate Change (BC3), Leioa, Spain. [20]Basque Foundation for Science (Ikerbasque), Bilbao, Spain. [21]Centre for Development and Environment, University of Bern, Bern, Switzerland. [22]National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bellary Road, Bangalore, Karnataka, India. [23]Marine Biodiversity, Exploitation and Conservation (MARBEC), Institut de Recherche pour le Développement (IRD), Université Montpellier, Insititut Français de Recherche pour l'Exploitation de la Mer (IFREMER), CNRS, 34000 Montpellier, France. [24]Brazilian National Institute for Research of the Amazon, 69080-971 Manaus, Brazil.

*Corresponding author. Email: hans.poertner@awi.de

†Deceased.

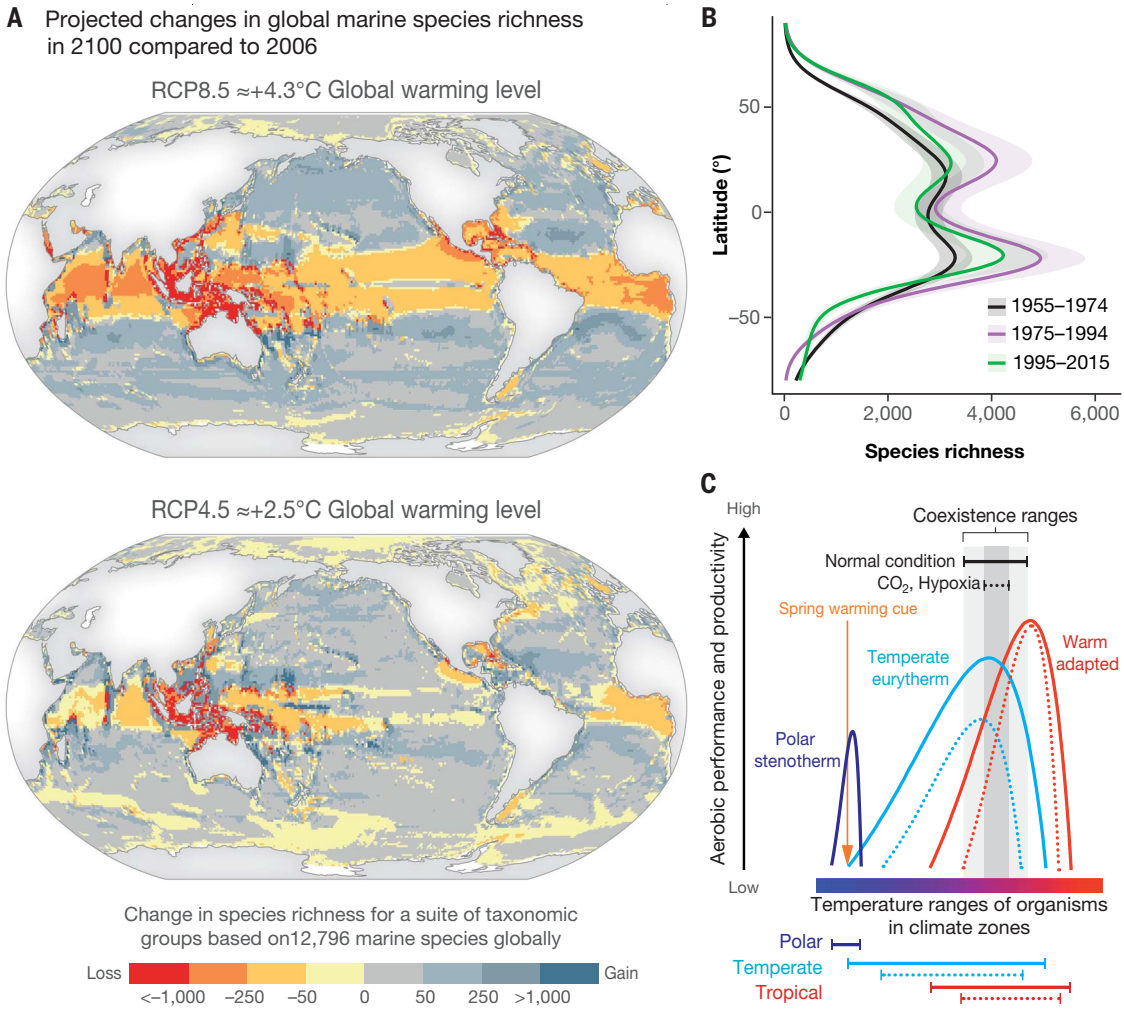


Fig. 1. Latitudinal shifts in marine species richness (12,796 species from 23 phyla) illustrate the responses of species to global warming. (**A**) Projected shifts reflect a scenario-dependent impoverishment of species richness at low latitudes. [Data are from (*10*), reproduced from (*2*).] (**B**) The projection is supported by present-day observations of reduced biodiversity at the equator compared with slightly higher Northern and Southern latitudes (48,661 marine animal species included across latitudes) [data are from (*149*)], reflecting that heat limits are already surpassed for Metazoa in warm ocean waters (*150*). Parallel phenomena are projected to occur on land (*2*). (**C**) The conceptual graph illustrates how species' thermal performance curves result from adaptation to the temperature regime of their biogeographical area. Overlapping curves define the temperature range and limits of species' coexistence and interactions. Shifts in coexistence and ecosystem dynamics would result from differential latitudinal shifts and from species-specific constraints on thermal ranges elicited by hypoxia and/or CO_2 enrichment (ocean acidification) (*8*, *18*, *151*). Surpassing such limits may trigger fatal tipping points in ecosystems and dependent societies.

address biodiversity loss and climate change, and their impacts on each other; (ii) proposed actions to address both crises together with consideration of societal needs; underpinned by (iii), the nexus between climate, biodiversity, and society, as well as the enablers and urgency of actions leading to transformative change.

Protecting biodiversity in a changing climate by multifunctional "scapes"

Conservation of biodiversity through spatial planning, conservation, and restoration is increasingly seen as an existential requirement for humankind (as addressed by CBD 2022). It is a flagship target of many conservation and restoration programs, including the Kunming-Montreal Global Biodiversity Framework of the CBD, the UN Decade on Ecosystem Restoration (2021–2030), and the Declaration on Forests and Land Use at the UNFCCC Conference of the Parties (COP) 26 (*19*). Global estimates of the average surface area required to be effectively protected for self-sustaining biodiversity and long-term ecosystem health range from 30 to 50% for land and ocean (*20*–*22*), but to integrate this target with multiple social objectives of human well-being, including equity and food security, is challenging (*23*–*26*). Further research is needed to derive biome-specific guidance and targets (*27*, *28*). For example, much higher protection, even 80%, might be needed to secure the Amazon's capacity to maintain its own regional climate regime and thereby biodiversity and global carbon sequestration capacity (*29*–*32*). For maintaining self-sustaining biodiversity in a habitable climate, the conservation of biodiversity- and carbon-rich natural ecosystems across latitudes—such as diverse forests, peatlands, and seagrass meadows—have highest priority. The conservation of spatially intact ecosystems can reduce risks to health, such as the risk of a pandemic, and implementation of some managed ecosystem types reduces climate risks to food security, such as through agroforestry (*24*, *33*–*36*). Conservation of, for example, coral reefs secures multifunctional benefits that support the UN Sustainable Development Goals (SDGs) (*37*). More encompassing conservation actions, extending beyond the historical focus on protected areas and species-based approaches, are thus required to safeguard biodiversity, support climate change mitigation and adaptation (*28*, *38*, *39*), and sustain NCP (*1*, *28*, *40*), including secure livelihoods for Indigenous Peoples and local communities (*41*). These ecosystem benefits include enhancing water and air quality; reducing soil erosion; ensuring pollination of crops; supporting food security and livelihoods; and regulating floods, regional climates, and microclimates. Climate change is already affecting access to or availability of many of these benefits (*40*). It is likely that with successful conservation and restoration efforts, human physical and mental health will also benefit (*2*).

A substantial conservation gap exists toward agreed (for example, Aichi biodiversity targets) and even more so, envisioned goals. As of 2020,

only 15% of land was declared as protected, and only a fraction of that area was effectively protected, given threats of encroachment, loss of legal status, and shifts of biodiversity into surrounding areas that are more intensively used (*42*, *43*). Protection rates are still below 10% in the ocean, again with a smaller fraction being highly protected (*23*). Moreover, the dichotomy between protected and unprotected zones has resulted in failures to address connectivity, spatial mosaic dynamics, and human needs. This calls for complementing conservation with extensive restoration efforts.

Restoration intervenes when high degrees of ecosystem degradation prevent spontaneous recovery. Restoration of ecosystems with native species enhances climate resilience of biodiversity but, in light of shifted biogeographies under climate change, may include new species assemblages to match future climatic conditions. Restoration can involve biodiversity offsetting, which aims to mitigate the negative impacts of infrastructure development, agricultural expansion, or mining by restoring biodiversity elsewhere. In light of shifting climate zones, restoration may need to do the same under climate change. However, if biodiversity-offsetting programs limit local people's access or cause loss of services on which their livelihoods depend, this can create negative impacts on adaptation (*44*). Social-ecological trade-offs and disconnects between the loss of local and "off-stage" (distant, diffuse, and delayed) biodiversity benefits can be minimized if the scale, type, and distribution of NCP are considered in restoration and offsetting initiatives (*45*, *46*).

A new generation of conservation and restoration actions that focus on multifunctional connected "scapes" (Fig. 2) is necessary to address biodiversity and climate interactions and their nexus with human development, health, and well-being (*47*). Increasingly, biodiversity management envisages a continuum of protection requirements from highly protected areas, through shared scapes, to highly human-dominated scapes (*48*). This integrative approach aims at the maintenance of self-sustaining biodiversity that is functionally intact and resilient with respect to species composition and in providing stable ecosystem services and that also meets human socioeconomic and cultural needs, including material provisioning (such as agriculture, forestry, and fisheries), nonmaterial contributions (such as cultural identity, spirituality and inspiration, and recreation), and regulatory contributions (such as pollination, water and soil protection, and microclimate). Such multifunctional scapes would include mosaics of protected areas, climate refugia and migration corridors, spaces shared by biodiversity and humans and modified by and for human use (including lands managed by Indigenous peoples and local communities), and profoundly transformed ecosystems such as urban and intensively farmed areas (Fig. 2) (*47*, *48*). Implementing mixed-use land- and seascapes could enhance biodiversity and climate cobenefits and also generate diverse, inclusive, and equitable benefits to communities and people. In managed ecosystems, increasing crop diversity within fields (for example, by using varietal mixtures, intercropping, or agroforestry), adding temporal diversity through extended crop rotations, and diversifying agricultural practices at landscape levels dilutes risk and increases resilience to climate change (*49*, *50*). In forests, management practices that promote biodiversity such as natural regeneration, mixed species stands, and selective harvests can also increase resilience and secure livelihoods (*51*, *52*).

The proportion and area of the different types of scapes would vary locally to meet different

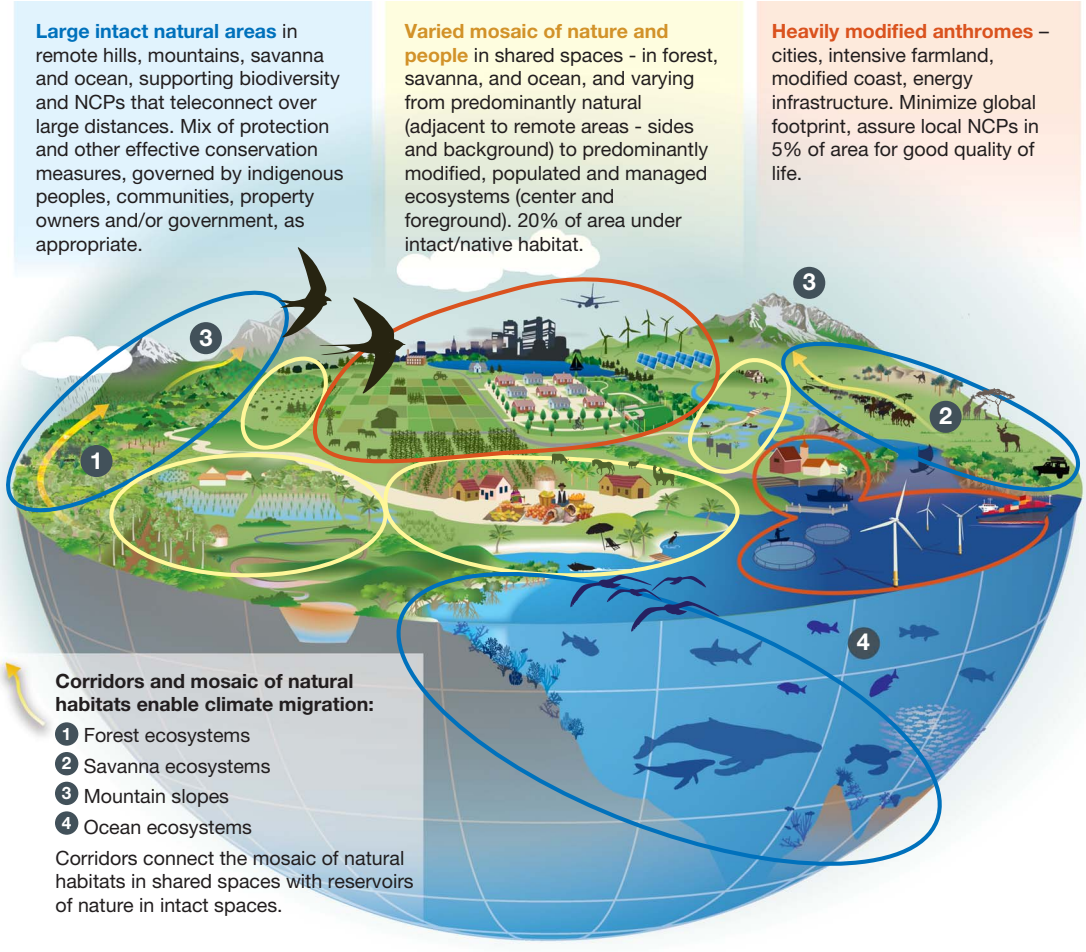


Fig. 2. Multifunctional connected and structurally and visibly distinct scapes across land, freshwater, and marine biomes. Scapes include large, intact wilderness spaces (blue circles), shared spaces (yellow circles), and anthromes (red circles). In spaces "shared" by people and biodiversity, the mosaic of intact natural habitat provides critical nature's contributions to people (NCP). Corridors of natural habitat (yellow arrows) facilitate species migration up elevational gradients. This multifunctional concept can assist integrating global and large-scale targets within local geographies (*18*).

local-to-global objectives for biodiversity, livelihoods, and well-being, embedded in polycentric governance models that have emerged to manage cross-sectoral complexity (*48*, *53*). Critically, planning must be closely coordinated across local jurisdictions and at larger scales to jointly meet global goals and objectives (*48*). Equitable participation and leadership by Indigenous peoples and local communities is essential to these efforts, given their important and long-term role in managing biodiverse scapes and in successful conservation initiatives (*54*, *55*). Extending these conservation measures requires substantial scaling up of financial investment (*56*), connecting efforts across scapes (*57*), and seeking positive feedbacks between biodiversity conservation and social benefits and burdens (*58*).

Key ecological and social safety measures require consideration when conceiving and implementing protection and restoration programs. The mosaic scapes approach (Fig. 2) would combine protection and sustainable uses, considering human and ecosystem needs per region and ecosystem. Protection and conservation initiatives work well when they involve stakeholder engagement and support, especially at the local scale, and where recognition of values and rights and the distribution of cobenefits (and potential trade-offs and costs) are perceived as just (*2*, *26*, *39*). Involvement in and access to restored areas is critical for local communities with ecosystem-based livelihoods and can support conservation success.

Conservation and restoration activities need to be paralleled by reducing excess human consumption habits and pollution. For example, measures across sectors that reduce the demand for energy through efficiency policies can reduce current emissions by more than 50% (*59*). Sustained reductions and eventual elimination of waste and pollution can be enabled through the progressive introduction of a circular economy as part of the envisioned holistic transformation that builds on system transitions toward sustainability in land, ocean, coastal, and freshwater ecosystems; urban and rural infrastructure; energy; industry; and society (*2*, *60*). Conservation and restoration successes depend on limiting the degree of global warming, ideally to below 1.5°C (*2*).

Nature-based solutions to mitigate climate change

The protection, sustainable management, and restoration of ecosystems not only benefit biodiversity but can also aid climate change mitigation and other societal challenges (*17*, *61*). Maximizing biodiversity protection and conservation as well as contributing to climate change mitigation through nature-based solutions (with large adaptation cobenefits) entails urgently avoiding and reversing losses and degradation of ecosystems (*2*, *62*, *63*). Protection and restoration of well-connected areas are not only key strategies for halting and reversing biodiversity loss (Fig. 2) (*64*, *65*), they are also among the most rapidly implemented measures that can secure and rebuild carbon-rich ecosystems. Restoration also has long-term cobenefits—for example, water quality and poverty reduction—especially when spatial planning is optimized for such criteria (*66*). Many locations best for protecting biodiversity and nature's contributions to people are coincident with currently high carbon storage and high capacity for ongoing sequestration (*18*, *41*, *49*, *67*). Terrestrial examples include intact tropical rainforest, wetlands, peatlands, grasslands, and savannahs (*49*). In the ocean, healthy mangrove forests, salt marshes, kelp forests, and seagrass meadows are important (*41*, *67*), but so are undisturbed sediments, as well as deep water and newly colonized polar blue carbon habitats emerging from ice melt (*68*). The role, scale, and efficacy of carbon storage through nature-based solutions varies considerably with regional and local contexts (Table 1) (*23*, *69*).

Although ecosystem restoration is often more cost effective than technological mitigation measures, nature-based solutions should not be mistaken as a substitute for substantial emissions reductions from energy, transportation, agriculture, infrastructure, building, and industrial sectors (*70*). The rate at which nature-based solutions can contribute to mitigation over time is too modest compared with current emission rates (*61*, *71*). Nonetheless, because of substantial carbon storage over time (Table 1), they have an important role in long-term climate stabilization (*71*, *72*).

However, as a source of concern, the rate and effectiveness of nature-based solutions is constrained by the degree of warming. Rising temperature progressively exceeds performance optima and adaptation limits of photosynthesis and its carbon assimilation in several ecosystems (following the principles outlined in Fig. 1). Warming thus constrains the potential of nature-based solutions in terrestrial systems, exacerbated by drought (*69*, *73*–*75*), and in some ocean systems (*71*, *76*, *77*). Therefore, climate change may even turn some ecosystems from carbon sinks to carbon sources. Maintaining a high capacity of carbon storage thus depends on limiting global warming to below 1.5°C, noting that the capacity of carbon binding and storage will decline and risks to carbon stocks increase with every 0.1°C on top of current warming (*2*, *67*, *76*). Warming by 1.5°C will already weaken nature-based solutions and represents a threshold in the transition from "safe" to "dangerous" climate change (*2*, *78*). Restricting warming to 1.5°C through ambitious emissions reductions, together with reducing other and interacting pressures that drive biodiversity decline, is thus a prerequisite for the success of conservation and restoration efforts involving natural and managed ecosystems (*3*, *23*, *66*, *71*, *78*–*81*) and a guardrail for avoiding catastrophic change (*82*).

Provided emissions are successfully reduced, the effectiveness of nature-based solutions can be optimized when planned for longevity, when social equity is adequately addressed, and when solutions are not narrowly focused on rapid carbon sequestration (*83*) but also on cobenefits beyond mitigation. With careful planning and implementation, including spatial aspects and social safeguards, nature-based solutions can support multiple UN SDGs also in cities (*84*, *85*). By enhancing ecosystem adaptive capacity and resilience, they can also reduce ecosystem degradation and the associated exacerbation of climate change. Higher genetic, species, and ecosystem diversity reduces both social and ecological risks and keeps adaptation options open (*86*).

Synergies and trade-offs in climate change mitigation, adaptation, and biodiversity conservation

Biodiversity conservation and climate change mitigation actions that are well-managed and considered together tend to display more synergies than trade-offs (Fig. 3). For example, both bioenergy and reforestation projects can be implemented carefully to contribute to human development needs, alongside pronounced and rapid reductions in fossil fuel emissions (*87*, *88*). The most robust path to limiting climate change while safeguarding biodiversity depends on identifying the strongest win-win solutions by region and avoiding those with negative interactions (*89*). Management of natural or seminatural ecosystems through reintroduction of keystone species or altered wildfire frequency can strengthen biodiversity while enhancing climate mitigation and adaptation. However, although most actions to conserve biodiversity are positive or neutral for climate, some potential climate mitigation and adaptation actions will have negative effects on biodiversity unless managed well (Fig. 3). Both ecosystem protection and restoration may also interfere with food production unless carefully realized (*24*, *25*). For wetland conservation, the quantification of carbon storage and greenhouse gas flux rates is uncertain with respect to the balance between methane (CH_4) sources and CO_2 sinks (*81*, *90*), but options are available to reduce their large CH_4 emissions, such as by restoring tidal flows (*91*).

Devoting vast land areas globally to the production of biomass for bioenergy is integral to many mitigation scenarios (*49*, *78*). Yet unintended negative side effects arise from their area requirements and competition for space, especially in terrestrial ecosystems, or from afforestation of natural grasslands (*87*, *92*, *93*).

Table 1. Examples for climate change mitigation aided through carbon uptake and storage with biodiversity cobenefits.

System and type*	Option	Mitigation potential	Biodiversity cobenefits	Data source
Forest protection	Reducing deforestation and forest degradation	0.4 to 5.8 Gt CO_2e year^{-1}	Soil and forest vegetation preserved	(*18*, *49*, *81*)
Forest restoration and reforestation	Forest expansion (for example, on less than half of degraded tropical forest area (3.69 million km^2)	1.5 to 10.1 (5.5 for tropical example by 2030) Gt CO_2e year^{-1}	Conservation of hundreds of vertebrate species	(*63*, *153*)
Land restoration	Restoring 15% of degraded land	up to 300 Gt CO_2 long term	Avoiding 60% of extinctions	(*66*)
Forest management	For example, selective logging	0.4 to 2.1 Gt CO_2e year^{-1} globally	Improved regeneration, protected forest soils	(*49*)
Agroforestry	Strengthening resilience, reducing soil tillage, and maintaining plant cover	0.1 to 5.7 Gt CO_2e year^{-1}	Resilience building, healthy soils	(*154*, *155*)
Peatland restoration	Protecting and growing carbon stores	0.15 to 0.81 Gt CO_2e year^{-1} up to 2050	Distinct biodiversity regained	(*81*, *156*, *157*)
Grassland and savannah restoration and management	Restoration of degraded grassland, rotational and light grazing	Reducing wildfires, strengthens soil carbon, 2.3 to 8 Gt CO_2e year^{-1}	Protection of megafauna, increase of grass-species and soil biodiversity	(*16*, *158*, *159*)
Total land management	Sustainably managed lands†	≈5 Gt CO_2e year^{-1} or higher by 2030	Global-scale synergies	(*18*)
Urban management	Green and blue spaces	Small contribution	Biodiversity refugia in cities	(*48*, *160*)
Marine protection	Blue carbon habitats (coastal marine)§	0.14 to 0.46 Gt CO_2e year^{-1}	Avoidance of coastal habitat loss	(*41*)
Marine restoration	Blue carbon habitats (coastal marine)§	0.62 to 1.06 Gt CO_2e year^{-1}	Expansion of coastal habitat	(*41*)
Protection of sediments and rebuilding marine species stocks	Protecting large sediment carbon stores, carbon export by abundant species stocks plankton to whales	Strengthening natural pathways of carbon export, large long-term potential	Restoring natural species stocks, enhanced nutrient cycling	(*23*, *161*, *162*)
Expanding ecosystems	High latitude retreat of ice opens new habitat	New biomass through colonization	Expansion of natural coastal habitat	(*71*, *163*)
Sustainably managed cultures	Seaweed use and farming	≈1500 tons CO_2e km^{-2} year^{-1}, up to 0.24 Gt CO_2e year^{-1} by 2050	Creating protected habitat	(*164*, *165*)

*Four types of nature-based solutions: (i) protection, (ii) restoration, (iii) management, or (iv) expansion of ecosystems in and around cities or across the wider landscape. †Release of methane can offset carbon sequestration. ‡Emissions reductions supported by plant-based human diets and reduced methane emissions from livestock, such as by adding farmed red seaweed to ruminant diets, and from rice paddies and reduced nitrous oxide (N_2O) from fertilizer application (*49*, *81*, *164*, *166*). §Blue carbon encompasses all biologically driven carbon fluxes and storage in marine systems that are amenable to management (*167*).

Projected CO_2 uptake rates through bioenergy or monoculture forest regrowth by 2050, which are similar in magnitude to double today's existing terrestrial carbon sink (*49*, *94*), are unrealistic. Further, relying on tree biomass for long-term carbon sequestration is risky, particularly in monocultures with high vulnerability to heat, drought, storms, fire, or pest outbreak (*95*). Trade-offs of bioenergy crops or afforestation at large scales by using monoculture or nonindigenous species (*96*) cause displacement of other land uses locally or elsewhere. Associated risks include carbon and biodiversity losses, negative impacts on food security, as well as dispossession of local access to land and to a wide range of NCP (*97*–*101*).

Deployment of renewable energy infrastructure can substantially contribute to mitigation, but hydropower, solar and wind energy, and storage for these intermittent sources of energy can also have negative impacts on biodiversity if their scale and design are not carefully implemented (*102*). For example, nearly 4000 major dam projects are currently underway or planned for energy production, irrigation, or flood control (*103*), but large dams generally have detrimental impacts on freshwater biodiversity (*104*) unless reduced through sustainable management (*105*). Likewise, the negative impacts of wind turbines on bats and birds, and of solar energy on the terrestrial or aquatic areas they cover, can be reduced through careful placement and operation (*106*, *107*). Technological mitigation measures can also exert harm to the environment and to biodiversity through the vast amount of materials required, such as metals or toxic waste products (*108*), which argues for strong environmental and social sustainability criteria. As one safe alternative, placing solar panels on half of the world's roofs could meet humankind's entire (as of 2018) electricity demand (*109*). Expanding circular economies will reduce the net use of materials, enable reuse, and avoid waste (*110*). Enhanced energy savings and efficiency also reduce the extent to which infrastructure needs to be deployed.

Climate adaptation policies can also incur large biodiversity impacts. For example, building sea walls to limit impacts of sea level rise on coastal infrastructure, adding irrigation capacity to reduce climate change impacts on agriculture, or introducing exotic tree species in anticipation of increasing climatic stress on forests impose substantial risks of large aggregate losses of biodiversity from these actions (*86*, *111*). As such, there is an urgent need

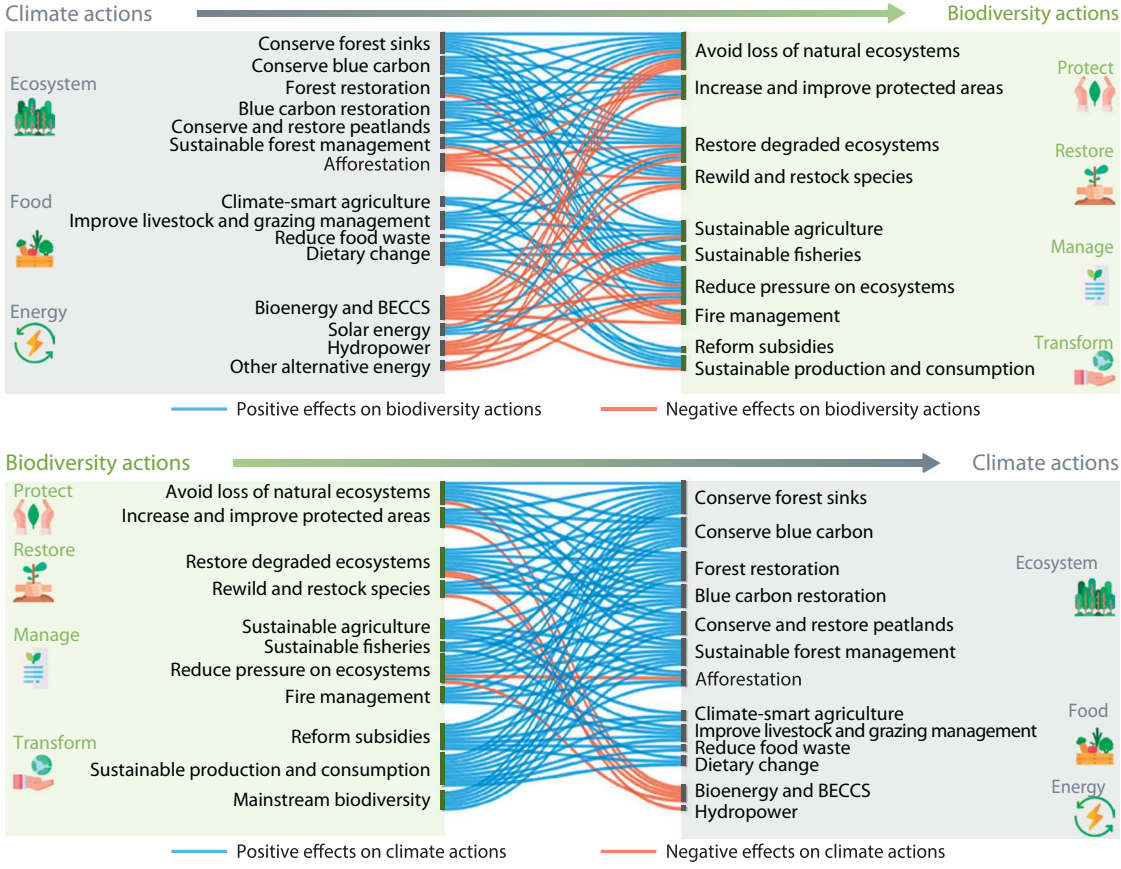


Fig. 3. Positive and negative effects (Top) of actions to mitigate climate change on actions to reduce biodiversity loss and (Bottom) of actions reducing biodiversity loss on those mitigating climate change. Orange lines indicate potential negative effects, and blue lines indicate potential positive effects. Interactions may shift over time and may surpass thresholds that trigger unforeseen consequences, positive or negative [after (*18*)].

to assess the cumulative impacts of these adaptation (and maladaptation) measures on biodiversity, evaluate more positive alternatives (for example, those that rely on nature-based solutions), and plan deployment with biodiversity impacts explicitly accounted for (*2*, *86*).

In both land and ocean systems, options exist to manage trade-offs by combining nature-based solutions and technology-based measures for climate change mitigation and adaptation, while sustaining biodiversity (*112*). For example, solar panel fields integrated with cropping or grazing systems can create multiple benefits: Grazing underneath solar panels can enhance soil carbon, and positive spillover effects into neighboring fields have been observed when habitat for pollinators is created underneath solar panels (*108*). Solar power generation on land is more efficient on an area basis than production of bioenergy crops and could thus contribute to reducing competition for land (*113*). Photovoltaic cells on the surface of water bodies can reduce evaporation, which could be beneficial to hydroelectric reservoirs in arid regions, considering that floating photovoltaics may also affect the water body's physical, chemical, and biological properties, particularly air exchange and thus oxygen supply (*114*). Offshore wind in combination with hydrogen generation can be powerful for both mitigation and biodiversity conservation if negative impacts on mobile species (such as birds) can be minimized (*115*); offshore turbines can create artificial substrate, with beneficial effects on marine biodiversity (*106*).

Implementing a nexus approach

Lack of appropriate governance mechanisms to anticipate and address the nexus between climate, biodiversity, and society can lead to the surpassing of biodiversity and climate tipping points with dire consequences for nature and society (*2*, *71*, *78*, *81*). Cascading effects, such as from melting of the polar ice sheets or forest dieback, can lead to breaching of socially acceptable limits and thresholds, especially regarding already vulnerable and marginalized societies (*116*). For example, reduced stability of crop yields may trigger food crises and subsequently lead to political instability, mass migration, and other crises (*2*, *81*). The uncertainty in the projections of tipping point timing and locations suggests that the best approach to avoid them is to stay far away from even the lowest thresholds that would trigger irreversible change (*117*) and implement quantified safety margins toward "safe and just" futures (*118*). These considerations emphasize the need for rapid progress in emissions reductions and biodiversity protection (*2*), paralleled by rapid transformative action in all systems (*2*).

Given the complexity of the nexus, there is thus a need for effective governance approaches that are appropriate for implementing and meeting protection and restoration goals, including multifunctional land, freshwater, and ocean scapes and providing equitable benefits that avoid trade-offs (*18*, *119*). This requires a "nexus approach" to governance, which at a minimum would involve mainstreaming of biodiversity into climate policy and vice versa, and of both into initiatives to advance human development and good quality of life. Such an approach would also benefit from socioecological systems perspectives, which identify interacting subsystems involving resources, users, and governance at multiple scales. Three key insights should inform related decisions and their implementation: first, all climate-biodiversity interactions have social considerations, with implications for both intra- and intergenerational equity and justice (*120*); second, telecoupling challenges abound, in which

off-stage effects also manifest far from the location of the intervention on different time scales (*46*, *121*); and third, trade-offs and feedbacks among system components, as well as threshold effects and nonlinear outcomes for biodiversity, climate, and societal relationships are the norm rather than the exception (*119*).

Taking a nexus approach can help guide policy interventions that have the potential to generate double or triple wins and avoid unintended consequences while facilitating more deep and rapid transformative change (*119*, *122*). Assessing the range of viable solutions requires recognition of differences in both environmental conditions and social and cultural situations and their vulnerabilities. For example, when defining suitable conditions for human society, ecosystems, their biodiversity, and the planet, the interrelationships between systems and between human, ecosystem, and planetary health need to be considered (*2*). Shifts in planetary functioning due to climate change (for example, responses of ocean currents and atmospheric streams) influence regional and local climate regimes, which in turn set environmental conditions for human well-being as well as ecosystem functioning, on which human health and well-being depends (*118*). Cascading unfavorable outcomes (such as habitat loss) can affect all levels and systems (species, ecosystems, and humans), requiring in turn multiscalar action (*2*).

Just as environmental characteristics differ from place to place and requirements for biodiversity differ across biomes, motivations, interests, preferences, and values embedded in institutions, societies, and cultures are also highly diverse (*58*, *123*). Biodiversity conservation interventions that take into account this diversity, such as those that deliberately include Indigenous peoples and local communities, have better outcomes on both human well-being and biological indicators (*54*). This observation calls for robust and transparent governance mechanisms aimed at inclusion that are able to overcome unequal power relations among stakeholders. A shift toward coordinated bottom-up pathways may enable both the local granularity and the global coherence needed to address nexus challenges across scales (*37*, *48*, *124*). Thus, it is crucial to identify interventions that are universal in terms of intent but sufficiently flexible and adaptive to different socioecological contexts, including historically and culturally embedded institutional and political governance structures (*125*, *126*). Some promising governance initiatives are already emerging, including jurisdictional approaches, experimental policy mixes, and rights-based approaches (*127*, *128*).

These approaches often stress the need for integrated and multifunctional interventions from toolboxes of adaptive solutions instead of single silver bullets. More innovative and flexible governance approaches can also give room to local autonomy to choose between alternative pathways to transformative change (*129*). Yet there is currently insufficient cross-sectoral policy coherence and integration, as seen in separate UN conventions such as UNFCCC and CBD, which often debate the interests of national governments who in turn have their own siloed approaches. Overall, although integrated solutions for the nexus exist that have cobenefits in terms of sustainable development and meeting basic needs of the poor and vulnerable (*130*), designing and financing these approaches faces numerous barriers because they require the cooperation of multiple actors across scales (*119*, *129*, *131*).

Enabling urgent action for transformative change

Limiting warming to 1.5°C or even to below 2°C and halting biodiversity loss require rapid action, which entails transformative change through transformative governance. This includes drawing on collaborative solutions across integrated systems, involves a broad range of actors and diversity of values about nature, engages different knowledge systems through more equitable approaches, and adaptively manages complex interactions (*119*, *132*, *133*). Transformative change implies deep shifts away from current ways of governing and decision-making that are often made on the basis of single-issue, short-term priorities. For example, current pledges of net zero emissions by companies and countries run the risk of exacerbating inequitable outcomes if they push nature-based solutions that are inappropriate (such as tree planting in natural grasslands), have few social benefits or increase vulnerability, or fail to be accompanied by simultaneous fossil fuel emissions reductions (*61*). Previous experiences from REDD+ programs (set up by UNFCCC to initiate "reduced emissions from deforestation and degradation" of forests) show that interventions need to be codesigned across stakeholders to create fair compensation mechanisms for equitable distribution of benefits and costs and to avoid risks to the most vulnerable, otherwise the effectiveness can be compromised (*21*, *49*).

Part of the move to transformative governance entails better harnessing of positive social tipping points that, when reached and exceeded, can accelerate transformative change toward desirable biodiversity-climate interactions and help navigate through strong and often unavoidable trade-offs within the nexus (Fig. 4) (*134*). Social tipping dynamics require synergizing technological, political, and behavioral processes toward structural reorganization, leading to targeted and "contagious" interventions with positive impacts on climate, biodiversity and human well-being (*116*, *135*). Such interventions depend on social contexts, cultural or behavioral specificities, and political as well as institutional systems (such as norms and regulations) (*119*). Ideally, goals for social tipping points should focus on structural reorganization through targeted interventions across institutions, technology, and behavior with positive impacts on climate, biodiversity, and human well-being (*116*, *135*). Some examples include the spread of carbon-neutrality pledges across businesses and cities, rapid adoption of new technologies such as electric vehicles, or the strengthening of climate and biodiversity education and civil society engagement (*119*, *136*–*138*). Because the locations of tipping points are moving targets, owing partly to the interconnectedness of the climate-biodiversity-social system, social tipping interventions need to be flexible and adaptive (*119*, *134*). Activating such social tipping interventions necessarily involves a shift of individual and collectively shared social values away from individualism and materialism to principles such as responsibility, stewardship, and justice (*59*, *123*). Such shifts in values will require structural actions as well, such as eliminating subsidies that presently support fossil fuel uses (*139*) or activities that harm biodiversity such as deforestation (*140*–*142*) or overfishing (*143*); such subsidies currently prevent the take-off of alternative approaches and technologies that would immediately aid climate change mitigation by avoiding further emissions. Transformative governance would also entail creating coalitions of support to confront recalcitrant and powerful interest groups (such as economic actors benefiting from fossil fuel subsidies) whose goal is keeping economic privileges and the political status quo and who lobby against transformative agendas (*144*).

Transformative change requires the use of levers and decision points that have the potential to alter future social-ecological trajectories (Fig. 4) (*59*). "Deep" leverage points (decision points that are harder to act on but have stronger potential effects) include moving away from valuing nature in narrow (market-based instrumental) ways and exploring alternative collective visions of development and good quality of life, while empowering individuals to enhance a sense of collective responsibility, especially toward more vulnerable natural systems and people (*123*). Interventions that activate deep leverage points could include (i) reconceptualizing economic development away from aggregate value of market flows (such as gross domestic product) to others that recognize a more inclusive understanding of changes in wealth (*145*) as well as the multiple values of nature (*123*) for a good quality of life within climate, biodiversity, and social limits; (ii) establishing deliberative governance instruments to empower civil society to take decisions by emphasizing their role as citizens as opposed to consumers (*123*, *146*); and (iii) recognition of Indigenous peoples' and community conserved

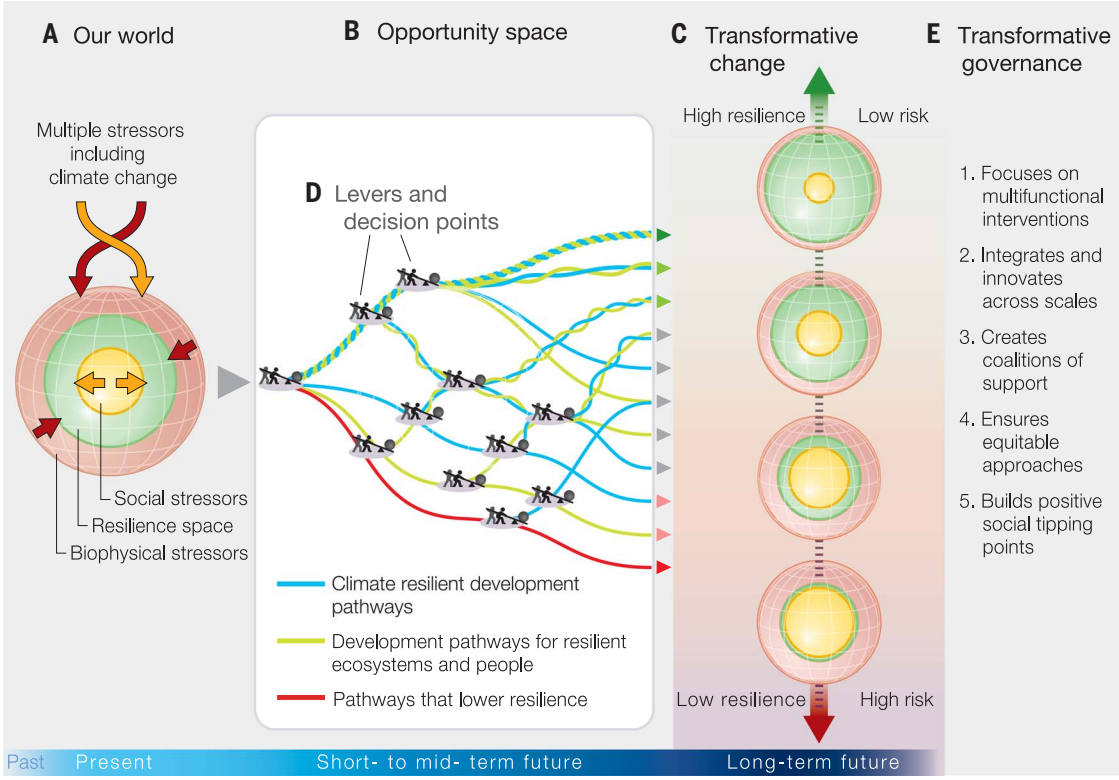


Fig. 4. Biodiversity-climate interactions in social-ecological systems with an explicit depiction of resilience outcomes for alternative pathways. (**A**) Alternative pathways start from the present world and cross the opportunity space (**B**), reaching different levels of resilience depending on the success of transformative change (**C**). It is assumed that high resilience for biodiversity and ecosystems and human society is paralleled by low risk as a condition for social stability and human well-being. (**D**) Decision points reached over time and associated levers relating to governance and policy interventions have the potential to keep the system on desired trajectories toward cobenefits and high resilience. (**E**) These transitions are enabled by transformative governance that triggers adaptive and inclusive decision-making at relevant time points, including positive social tipping interventions that can help shift toward positive biodiversity-climate interactions and associated social outcomes (*2*). [Figure adapted from (*152*).]

territories and areas, initiated, designed, and governed by local communities, along with full development of "other effective area-based conservation measures" (*147*) that can lead to the conservation of natural and modified ecosystems and their biodiversity and associated benefits, including climate benefits (*69*).

Conclusions

Achieving the scale and scope of transformative change needed to concurrently meet the goals of the UNFCCC, the CBD, and the UN Agenda 2030 and its SDGs will rely on far-reaching mobilization and transformation actions of a type never before attempted, within a rapidly closing time window as required to keep global warming below 1.5°C and to secure a livable future (*2*). Healthy ecosystems and a healthy planet are preconditions for flourishing life on Earth. Strengthening biodiversity in all systems will also support long-term climate stabilization. Securing human, ecosystem, and planetary health and societal well-being is facilitated through implementation of multifunctional connected scapes and other integrated actions across the climate-biodiversity nexus, building on an analysis of critical assets (*148*) and their protection and expansion through restoration. Securing a livable future will require rapid action and commitment not only from countries through actions in their national territories but also from emergent coalitions and governance models at all levels. Addressing conventions in concert with one another, such as the Paris Agreement of the UNFCCC (alongside SDG 13) and the Kunming-Montreal Global Biodiversity Framework (together with SDGs 14 and 15), will put society on the pathway to a positive vision of good quality of life in harmony with nature (*37*). Engaging in deeper transformative change will require effective incentives and social tipping points, capacity building with broad education and outreach initiatives, institutional change and improved cooperation across sectors and jurisdictions, and shifts in values to support intergenerational justice as well as equity more broadly. This vision of inclusive, integrative, and adaptive decision-making can overcome societal and political inertia (*2*, *3*) and thereby help society to avoid crossing biophysical tipping points and their worst impacts, as well as help build a more just and sustainable world (*2*).

REFERENCES AND NOTES

1. S. Díaz *et al.*, Assessing nature's contributions to people. *Science* **359**, 270–272 (2018). doi: 10.1126/science.aap8826; pmid: 29348221
2. Intergovernmental Panel on Climate Change, "Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change" in, H.-O. Pörtner *et al.*, Eds. (Cambridge Univ. Press., 2022), p. 3056.
3. Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES), "Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services," E. S. Brondizio, J. Settele, S. Díaz, H. T. Ngo, Eds. (IPBES, 2019); https://doi.org/10.5281/zenodo.3831674.
4. Intergovernmental Panel on Climate Change, "Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change," V. Masson-Delmotte *et al.*, Eds. (Cambridge Univ. Press, 2021).
5. J. E. M. Watson *et al.*, Protect the last of the wild. *Nature* **563**, 27–30 (2018). doi: 10.1038/d41586-018-07183-6; pmid: 30382225
6. Y. M. Bar-On, R. Phillips, R. Milo, The biomass distribution on Earth. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 6506–6511 (2018). doi: 10.1073/pnas.1711842115; pmid: 29784790
7. S. Díaz *et al.*, Pervasive human-driven decline of life on Earth points to the need for transformative change. *Science* **366**, eaax3100 (2019). doi: 10.1126/science.aax3100; pmid: 31831642
8. H. O. Pörtner *et al.*, Ocean systems in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth*

Assessment Report of the Intergovernmental Panel on Climate Change, C. B. Field *et al*., Eds. (Cambridge Univ. Press, 2014), pp. 411–484; https://www.ipcc.ch/site/assets/uploads/2018/02/WGIIAR5-Chap6_FINAL.pdf.
9. R. Warren, J. Price, E. Graham, N. Forstenhaeusler, J. VanDerWal, The projected effect on insects, vertebrates, and plants of limiting global warming to 1.5°C rather than 2°C. *Science* **360**, 791–795 (2018). doi: 10.1126/science.aar3646; pmid: 29773751
10. J. García Molinos *et al*., Climate velocity and the future global redistribution of marine biodiversity. *Nat. Clim. Chang.* **6**, 83–88 (2016). doi: 10.1038/nclimate2769
11. D. F. Zhao *et al*., Environmental conditions regulate the impact of plants on cloud formation. *Nat. Commun.* **8**, 14067 (2017). doi: 10.1038/ncomms14067; pmid: 28218253
12. S. Chen *et al*., Plant diversity enhances productivity and soil carbon storage. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4027–4032 (2018). doi: 10.1073/pnas.1700298114; pmid: 29666315
13. S. Wang, M. Maltrud, S. Elliott, P. Cameron-Smith, A. Jonko, Influence of dimethyl sulfide on the carbon cycle and biological production. *Biogeochemistry* **138**, 49–68 (2018). doi: 10.1007/s10533-018-0430-5
14. T. W. Crowther *et al*., The global soil community and its influence on biogeochemistry. *Science* **365**, eaav0550 (2019). doi: 10.1126/science.aav0550; pmid: 31439761
15. S. Xu *et al*., Species richness promotes ecosystem carbon storage: Evidence from biodiversity-ecosystem functioning experiments. *Proc. Biol. Sci.* **287**, 20202063 (2020). doi: 10.1098/rspb.2020.2063; pmid: 33234078
16. O. J. Schmitz *et al*., Animals and the zoogeochemistry of the carbon cycle. *Science* **362**, eaar3213 (2018). doi: 10.1126/science.aar3213; pmid: 30523083
17. E. Cohen-Shacham, C. Janzen, S. Maginnis, G. Walters, *Nature-Based Solutions to Address Global Societal Challenges* (IUCN, 2016); https://portals.iucn.org/library/node/46191.
18. H.-O. Pörtner *et al*., "Scientific outcome of the IPBES-IPCC co-sponsored workshop on biodiversity and climate change" (IPBES Secretariat, 2021); doi: 10.5281/zenodo.5031995.
19. UK Government, "Policy paper – World leaders summit on 'Action on forests and land use'" (2021); https://www.gov.uk/government/publications/cop26-world-leaders-summit-on-action-on-forests-and-land-use-2-november-2021/world-leaders-summit-on-action-on-forests-and-land-use.
20. E. Wilson, *Half-Earth: Our Planet's Fight for Life* (Liveright, 2016).
21. E. Dinerstein *et al*., An ecoregion-based approach to protecting half the terrestrial realm. *Bioscience* **67**, 534–545 (2017). doi: 10.1093/biosci/bix014; pmid: 28608869
22. J. E. M. Watson *et al*., Set a global target for ecosystems. *Nature* **578**, 360–362 (2020). doi: 10.1038/d41586-020-00446-1; pmid: 32071453
23. E. Sala *et al*., Protecting the global ocean for biodiversity, food and climate. *Nature* **592**, 397–402 (2021). doi: 10.1038/s41586-021-03371-z; pmid: 33731930
24. Z. Mehrabi, E. C. Ellis, N. Ramankutty, The challenge of feeding the world while conserving half the planet. *Nat. Sustain.* **1**, 409–412 (2018). doi: 10.1038/s41893-018-0119-8
25. J. Schleicher *et al*., Protecting half of the planet could directly affect over one billion people. *Nat. Sustain.* **2**, 1094–1096 (2019). doi: 10.1038/s41893-019-0423-y
26. N. Zafra-Calvo *et al*., Progress toward equitably managed protected areas in Aichi Target 11: A global survey. *Bioscience* **69**, 191–197 (2019). doi: 10.1093/biosci/biy143; pmid: 30914829
27. C. A. Soto-Navarro *et al*., Towards a multidimensional biodiversity index for national application. *Nat. Sustain.* **4**, 933–942 (2021). doi: 10.1038/s41893-021-00753-z
28. S. Díaz *et al*., Set ambitious goals for biodiversity and sustainability. *Science* **370**, 411–413 (2020). doi: 10.1126/science.abe1530; pmid: 33093100
29. D. M. Lapola *et al*., Pervasive transition of the Brazilian land-use system. *Nat. Clim. Chang.* **4**, 27–35 (2014). doi: 10.1038/nclimate2056
30. T. E. Lovejoy, C. Nobre, Amazon tipping point. *Sci. Adv.* **4**, eaat2340 (2018). doi: 10.1126/sciadv.aat2340; pmid: 29492460
31. O. L. Phillips, R. J. W. Brienen; RAINFOR collaboration, Carbon uptake by mature Amazon forests has mitigated Amazon nations' carbon emissions. *Carbon Balance Manag.* **12**, 1 (2017). doi: 10.1186/s13021-016-0069-2; pmid: 28413845
32. A. Antonelli *et al*., Amazonia is the primary source of Neotropical biodiversity. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 6034–6039 (2018). doi: 10.1073/pnas.1713819115; pmid: 29760058
33. IPBES, "Workshop Report on Biodiversity and Pandemics of the Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES)," in P. Daszak *et al*., Eds. (IPBES Secretariat, 2020; doi: 10.5281/zenodo.4147317.
34. A. L. Val, Biodiversity—The hidden risks. *An. Acad. Bras. Cienc.* **92**, e20200699 (2020). doi: 10.1590/0001-3765202020200699; pmid: 32428091
35. R. K. F. Unsworth, L. M. Nordlund, L. C. Cullen-Unsworth, Seagrass meadows support global fisheries production. *Conserv. Lett.* **12**, e12566 (2019). doi: 10.1111/conl.12566
36. F. A. Ascioti, M. C. Mangano, C. Marcianò, G. Sarà, The sanitation service of seagrasses – Dependencies and implications for the estimation of avoided costs. *Ecosyst. Serv.* **54**, 101418 (2022). doi: 10.1016/j.ecoser.2022.101418
37. D. O. Obura, Getting to 2030—Scaling effort to ambition through a narrative model of the SDGs. *Mar. Policy* **117**, 103973 (2020). doi: 10.1016/j.marpol.2020.103973
38. G. G. Gurney, S. Mangubhai, M. Fox, M. Kiatkoski Kim, A. Agrawal, Equity in environmental governance: Perceived fairness of distributional justice principles in marine co-management. *Environ. Sci. Policy* **124**, 23–32 (2021). doi: 10.1016/j.envsci.2021.05.022
39. H. D. Jonas *et al*., Equitable and effective area-based conservation: Towards the conserved areas paradigm. *Parks* **27**, 71–84 (2021). doi: 10.2305/IUCN.CH.2021.PARKS-27-1HJ.en
40. R. Chaplin-Kramer *et al*., Global modeling of nature's contributions to people. *Science* **366**, 255–258 (2019). doi: 10.1126/science.aaw3372; pmid: 31601772
41. P. I. Macreadie *et al*., Blue carbon as a natural climate solution. *Nat. Rev. Earth Environ.* **2**, 826–839 (2021). doi: 10.1038/s43017-021-00224-1
42. Secretariat of the Convention on Biological Diversity, "Global Biodiversity Outlook 5" (SCBD, Montreal, Canada, 2020).
43. R. E. Golden Kroner *et al*., The uncertain future of protected lands and waters. *Science* **364**, 881–886 (2019). doi: 10.1126/science.aau5525; pmid: 31147519
44. J. P. G. Jones *et al*., Net gain: Seeking better outcomes for local people when mitigating biodiversity loss from development. *One Earth* **1**, 195–201 (2019). doi: 10.1016/j.oneear.2019.09.007
45. J. T. Erbaugh, J. A. Oldekop, Forest landscape restoration for livelihoods and well-being. *Curr. Opin. Environ. Sustain.* **32**, 76–83 (2018). doi: 10.1016/j.cosust.2018.05.007
46. U. Pascual *et al*., Off-stage ecosystem service burdens: A blind spot for global sustainability. *Environ. Res. Lett.* **12**, 075001 (2017). doi: 10.1088/1748-9326/aa7392
47. H. Locke *et al*., Three global conditions for biodiversity conservation and sustainable use: An implementation framework. *Natl. Sci. Rev.* **6**, 1080–1082 (2019). doi: 10.1093/nsr/nwz136; pmid: 34691979
48. D. O. Obura *et al*., Integrate biodiversity targets from local to global levels. *Science* **373**, 746–748 (2021). doi: 10.1126/science.abh2234; pmid: 34385386
49. P. Smith *et al*., Which practices co-deliver food security, climate change mitigation and adaptation, and combat land degradation and desertification? *Glob. Chang. Biol.* **26**, 1532–1575 (2020). doi: 10.1111/gcb.14878; pmid: 31637793
50. FAO, *The State of the World's Biodiversity for Food and Agriculture* (FAO Commission on Genetic Resources for Food and Agriculture Assessments, 2019); http://www.fao.org/3/CA3129EN/CA3129EN.pdf.
51. W. R. L. Anderegg *et al*., Hydraulic diversity of forests regulates ecosystem resilience during drought. *Nature* **561**, 538–541 (2018). doi: 10.1038/s41586-018-0539-7; pmid: 30232452
52. F. Bussotti, M. Pollastrini, V. Holland, W. Brüggemann, Functional traits and adaptive capacity of European forests to climate change. *Environ. Exp. Bot.* **111**, 91–113 (2015). doi: 10.1016/j.envexpbot.2014.11.006
53. K. Carlisle, R. L. Gruby, Polycentric systems of governance: A theoretical model for the commons. *Policy Stud. J.* **47**, 927–952 (2019). doi: 10.1111/psj.12212
54. N. M. Dawson *et al*., The role of Indigenous peoples and local communities in effective and equitable conservation. *Ecol. Soc.* **26**, art19 (2021). doi: 10.5751/ES-12625-260319
55. S. T. Garnett *et al*., A spatial overview of the global importance of Indigenous lands for conservation. *Nat. Sustain.* **1**, 369–374 (2018). doi: 10.1038/s41893-018-0100-6
56. A. Seidl, K. Mulungu, M. Arlaud, O. van den Heuvel, M. Riva, The effectiveness of national biodiversity investments to protect the wealth of nature. *Nat. Ecol. Evol.* **5**, 530–539 (2021). doi: 10.1038/s41559-020-01372-1; pmid: 33462490
57. M. Ward *et al*., Just ten percent of the global terrestrial protected area network is structurally connected via intact land. *Nat. Commun.* **11**, 4563 (2020). doi: 10.1038/s41467-020-18457-x; pmid: 32917882
58. U. Pascual *et al*., Biodiversity and the challenge of pluralism. *Nat. Sustain.* **4**, 567–572 (2021). doi: 10.1038/s41893-021-00694-7
59. K. M. A. Chan *et al*., Levers and leverage points for pathways to sustainability. *People Nat.* **2**, 693–717 (2020). doi: 10.1002/pan3.10124
60. E. H. Arruda, R. A. P. B. Melatto, W. Levy, D. de Melo Conti, Circular economy: A brief literature review (2015–2020). *Sustain. Oper. Comput.* **2**, 79–86 (2021). doi: 10.1016/j.susoc.2021.05.001
61. N. Seddon, Harnessing the potential of nature-based solutions for mitigating and adapting to climate change. *Science* **376**, 1410–1416 (2022). doi: 10.1126/science.abn9668; pmid: 35737796
62. P. Smith *et al*., How do we best synergize climate mitigation actions to co-benefit biodiversity? *Glob. Chang. Biol.* **28**, 2555–2577 (2022). doi: 10.1111/gcb.16056; pmid: 34951743
63. Y.-J. Shin *et al*., Actions to halt biodiversity loss generally benefit the climate. *Glob. Chang. Biol.* **28**, 2846–2874 (2022). doi: 10.1111/gcb.16109; pmid: 35098619
64. R. A. Magris *et al*., Biologically representative and well-connected marine reserves enhance biodiversity persistence in conservation planning. *Conserv. Lett.* **11**, e12439 (2018). doi: 10.1111/conl.12439
65. S. L. Maxwell *et al*., Area-based conservation in the twenty-first century. *Nature* **586**, 217–227 (2020). doi: 10.1038/s41586-020-2773-z; pmid: 33028996
66. B. B. N. Strassburg *et al*., Global priority areas for ecosystem restoration. *Nature* **586**, 724–729 (2020). doi: 10.1038/s41586-020-2784-9; pmid: 33057198
67. O. Hoegh-Guldberg *et al*., The human imperative of stabilizing global climate change at 1.5°C. *Science* **365**, eaaw6974 (2019). doi: 10.1126/science.aaw6974; pmid: 31604209
68. N. Bax *et al*., Perspective: Increasing blue carbon around Antarctica is an ecosystem service of considerable societal and economic value worth protecting. *Glob. Chang. Biol.* **27**, 5–12 (2021). doi: 10.1111/gcb.15392; pmid: 33064891
69. M. L. Noon *et al*., Mapping the irrecoverable carbon in Earth's ecosystems. *Nat. Sustain.* **5**, 37–46 (2021). doi: 10.1038/s41893-021-00803-6
70. N. Seddon *et al*., Getting the message right on nature-based solutions to climate change. *Glob. Chang. Biol.* **27**, 1518–1546 (2021). doi: 10.1111/gcb.15513; pmid: 33522071
71. Intergovernmental Panel on Climate Change (IPCC), "IPCC Special Report on the Ocean and Cryosphere in a Changing Climate," H. O. Pörtner *et al*., Eds. (Cambridge Univ. Press, 2019). doi: 10.1017/9781009157964
72. German Advisory Council on Global Change (WBGU), "Beyond Climate Neutrality. Policy Paper 12," policy paper (WBGU, 2021); https://www.wbgu.de/fileadmin/user_upload/wbgu/publikationen/politikpapiere/pp12_2021/pdf_und_ebook/wbgu_pp12_navi_en.pdf.
73. K. A. Duffy *et al*., How close are we to the temperature tipping point of the terrestrial biosphere? *Sci. Adv.* **7**, eaay1052 (2021). doi: 10.1126/sciadv.aay1052; pmid: 33523891
74. D. Gampe *et al*., Increasing impact of warm droughts on northern ecosystem productivity over recent decades. *Nat. Clim. Chang.* **11**, 772–779 (2021). doi: 10.1038/s41558-021-01112-8
75. L. V. Gatti *et al*., Amazonia as a carbon source linked to deforestation and climate change. *Nature* **595**, 388–393 (2021). doi: 10.1038/s41586-021-03629-6; pmid: 34262208
76. H.-G. Hoppe, K. Gocke, R. Koppe, C. Begler, Bacterial growth and primary production along a north-south transect of the Atlantic Ocean. *Nature* **416**, 168–171 (2002). doi: 10.1038/416168a; pmid: 11894092
77. L. Kwiatkowski *et al*., Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections. *Biogeosciences* **17**, 3439–3470 (2020). doi: 10.5194/bg-17-3439-2020
78. Intergovernmental Panel on Climate Change, "Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty", V. Masson-Delmotte *et al*., Eds. (Cambridge Univ. Press, 2018). doi: 10.1017/9781009157940

79. M. Jung *et al*., Areas of global importance for conserving terrestrial biodiversity, carbon and water. *Nat. Ecol. Evol.* **5**, 1499–1509 (2021). doi: 10.1038/s41559-021-01528-7; pmid: 34429536
80. D. Leclère *et al*., Bending the curve of terrestrial biodiversity needs an integrated strategy. *Nature* **585**, 551–556 (2020). doi: 10.1038/s41586-020-2705-y; pmid: 32908312
81. Intergovernmental Panel on Climate Change, "Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems", P. R. Shukla *et al*., Eds. (Cambridge Univ. Press, 2019). doi: 10.1017/9781009157988
82. L. Kemp *et al*., Climate Endgame: Exploring catastrophic climate change scenarios. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2108146119 (2022). doi: 10.1073/pnas.2108146119; pmid: 35914185
83. C. A. J. Girardin *et al*., Nature-based solutions can help cool the planet - if we act now. *Nature* **593**, 191–194 (2021). doi: 10.1038/d41586-021-01241-2; pmid: 33981055
84. I. Palomo *et al*., Assessing nature-based solutions for transformative change. *One Earth* **4**, 730–741 (2021). doi: 10.1016/j.oneear.2021.04.013
85. S. Goodwin, M. Olazabal, A. J. Castro, U. Pascual, Global mapping of urban nature-based solutions for climate change adaptation. *Nat. Sustain.* (2023). doi: 10.1038/s41893-022-01036-x
86. M. D. Morecroft *et al*., Measuring the success of climate change adaptation and mitigation in terrestrial ecosystems. *Science* **366**, eaaw9256 (2019). doi: 10.1126/science.aaw9256; pmid: 31831643
87. S. Fuss *et al*., Negative emissions-Part 2: Costs, potentials and side effects. *Environ. Res. Lett.* **13**, 063002 (2018). doi: 10.1088/1748-9326/aabf9f
88. S. Roe *et al*., Contribution of the land sector to a 1.5 °C world. *Nat. Clim. Chang.* **9**, 817–828 (2019). doi: 10.1038/s41558-019-0591-9
89. P. McElwee *et al*., The impact of interventions in the global land and agri-food sectors on Nature's Contributions to People and the UN Sustainable Development Goals. *Glob. Chang. Biol.* **26**, 4691–4721 (2020). doi: 10.1111/gcb.15219; pmid: 32531815
90. T. Spencer *et al*., Global coastal wetland change under sea-level rise and related stresses: The DIVA Wetland Change Model. *Global Planet. Change* **139**, 15–30 (2016). doi: 10.1016/j.gloplacha.2015.12.018
91. J. A. Rosentreter *et al*., Half of global methane emissions come from highly variable aquatic ecosystem sources. *Nat. Geosci.* **14**, 225–230 (2021). doi: 10.1038/s41561-021-00715-2
92. R. C. R. Abreu *et al*., The biodiversity cost of carbon sequestration in tropical savanna. *Sci. Adv.* **3**, e1701284 (2017). doi: 10.1126/sciadv.1701284; pmid: 28875172
93. K. D. Holl, P. H. S. Brancalion, Tree planting is not a simple solution. *Science* **368**, 580–581 (2020). doi: 10.1126/science.aba8232; pmid: 32381704
94. J. Rogelj *et al*., Scenarios towards limiting global mean temperature increase below 1.5°C. *Nat. Clim. Chang.* **8**, 325–332 (2018). doi: 10.1038/s41558-018-0091-3
95. W. R. L. Anderegg *et al*., Climate-driven risks to the climate mitigation potential of forests. *Science* **368**, eaaz7005 (2020). doi: 10.1126/science.aaz7005; pmid: 32554569
96. S. Lewis, C. Wheeler, E. Mitchard, A. Koch, Restoring natural forests is the best way to remove atmospheric carbon. *Nature* **568**, 25–28 (2019). doi: 10.1038/d41586-019-01026-8; pmid: 30940972
97. K. Dooley, S. Kartha, Land-based negative emissions: Risks for climate mitigation and impacts on sustainable development. *Int. Environ. Agreement Polit. Law Econ.* **18**, 79–98 (2018). doi: 10.1007/s10784-017-9382-9
98. F. Humpenöder *et al*., Large-scale bioenergy production: How to resolve sustainability trade-offs? *Environ. Res. Lett.* **13**, 024011 (2018). doi: 10.1088/1748-9326/aa9e3b
99. A. Arneth *et al*., "Chapter 1: Framing and Context" in *Climate Change and Land: An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems*, P. R. Shukla *et al*., Eds. (IPCC, 2019); https://www.ipcc.ch/site/assets/uploads/sites/4/2019/12/04_Chapter-1.pdf.
100. O. Reisman-Berman, T. Keasar, N. Tel-Zur, Native and non-native species for dryland afforestation: Bridging ecosystem integrity and livelihood support. *Ann. For. Sci.* **76**, 114 (2019). doi: 10.1007/s13595-019-0903-2
101. B. A. Polidoro *et al*., The loss of species: Mangrove extinction risk and geographic areas of global concern. *PLOS ONE* **5**, e10095 (2010). doi: 10.1371/journal.pone.0010095; pmid: 20386710
102. J. M. Pérez-García *et al*., Priority areas for conservation alone are not a good proxy for predicting the impact of renewable energy expansion. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2204505119 (2022). doi: 10.1073/pnas.2204505119; pmid: 35878057
103. J. Boulange, N. Hanasaki, D. Yamazaki, Y. Pokhrel, Role of dams in reducing global flood exposure under climate change. *Nat. Commun.* **12**, 417 (2021). doi: 10.1038/s41467-020-20704-0; pmid: 33462241
104. H. Wu *et al*., Effects of dam construction on biodiversity: A review. *J. Clean. Prod.* **221**, 480–489 (2019). doi: 10.1016/j.jclepro.2019.03.001
105. J. Best, Anthropogenic stresses on the world's big rivers. *Nat. Geosci.* **12**, 7–21 (2019). doi: 10.1038/s41561-018-0262-x
106. T. H. Soukissian *et al*., Marine renewable energy in the Mediterranean Sea: Status and perspectives. *Energies* **10**, 1512 (2017). doi: 10.3390/en10101512
107. K. Dai, A. Bergot, C. Liang, W.-N. Xiang, Z. Huang, Environmental issues associated with wind energy—A review. *Renew. Energy* **75**, 911–921 (2015). doi: 10.1016/j.renene.2014.10.074
108. A. Dhar, M. A. Naeth, P. D. Jennings, M. Gamal El-Din, Perspectives on environmental impacts and a land reclamation strategy for solar and wind energy systems. *Sci. Total Environ.* **718**, 134602 (2020). doi: 10.1016/j.scitotenv.2019.134602; pmid: 31843306
109. S. Joshi *et al*., High resolution global spatiotemporal assessment of rooftop solar photovoltaics potential for renewable electricity generation. *Nat. Commun.* **12**, 5738 (2021). doi: 10.1038/s41467-021-25720-2; pmid: 34611151
110. OECD, "Global Material Resources Outlook to 2060: Economic Drivers and Environmental Consequences" (OECD Publishing, 2018); https://dx.doi.org/10.1787/9789264307452-en.
111. P. M. Berry *et al*., Cross-sectoral interactions of adaptation and mitigation measures. *Clim. Change* **128**, 381–393 (2014). doi: 10.1007/s10584-014-1214-0
112. R. R. Hernandez *et al*., Techno-ecological synergies of solar energy for global sustainability. *Nat. Sustain.* **2**, 560–568 (2019). doi: 10.1038/s41893-019-0309-z
113. T. D. Searchinger, T. Beringer, A. Strong, Does the world have low-carbon bioenergy potential from the dedicated use of land? *Energy Policy* **110**, 434–446 (2017). doi: 10.1016/j.enpol.2017.08.016
114. A. Armstrong, T. Page, S. J. Thackeray, R. R. Hernandez, I. D. Jones, Integrating environmental understanding into freshwater floatovoltaic deployment using an effects hierarchy and decision trees. *Environ. Res. Lett.* **15**, 114055 (2020). doi: 10.1088/1748-9326/abbf7b
115. M. Agha, J. E. Lovich, J. R. Ennen, B. D. Todd, Wind, sun, and wildlife: Do wind and solar energy development "short-circuit" conservation in the western United States? *Environ. Res. Lett.* **15**, 075004 (2020). doi: 10.1088/1748-9326/ab8846
116. M. Milkoreit *et al*., Defining tipping points for social-ecological systems scholarship—An interdisciplinary literature review. *Environ. Res. Lett.* **13**, 033005 (2018). doi: 10.1088/1748-9326/aaaa75
117. D. M. Bergstrom *et al*., Combating ecosystem collapse from the tropics to the Antarctic. *Glob. Chang. Biol.* **27**, 1692–1703 (2021). doi: 10.1111/gcb.15539; pmid: 33629799
118. J. Rockström *et al*., Identifying a safe and just corridor for people and the planet. *Earths Futur.* **9**, (2021). doi: 10.1029/2020EF001866
119. U. Pascual *et al*., Governing for transformative change across the biodiversity-climate-society nexus. *Bioscience* **72**, 684–704 (2022). doi: 10.1093/biosci/biac031
120. M. Leach *et al*., Equity and sustainability in the Anthropocene: A social–ecological systems perspective on their intertwined futures. *Glob. Sustain.* **1**, e13 (2018). doi: 10.1017/sus.2018.12
121. L. R. Carrasco, J. Chan, F. L. McGrath, L. T. P. Nghiem, Biodiversity conservation in a telecoupled world. *Ecol. Soc.* **22**, art24 (2017). doi: 10.5751/ES-09448-220324
122. Y. Takahashi *et al*., Enhancing synergies in nature's contributions to people in socio-ecological production landscapes and seascapes: Lessons learnt from ten site-based projects in biodiversity hotspots. *Sustain. Sci.* **17**, 823–836 (2021). doi: 10.1007/s11625-021-00927-w
123. Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, "Methodological assessment of the diverse values and valuation of nature of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services," P. Balvanera, U. Pascual, M. Christie, B. Baptiste, D. González-Jiménez, Eds. (Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, 2022); https://doi.org/10.5281/zenodo.6522523.
124. D. Obura, Bringing the Sustainable Development Goals to life through stories. *Current Conservation* 20 June 2021; https://www.currentconservation.org/bringing-the-sustainable-development-goals-to-life-through-stories.
125. C. Folke, T. Hahn, P. Olsson, J. Norberg, Adaptive governance of social-ecological systems. *Annu. Rev. Environ. Resour.* **30**, 441–473 (2005). doi: 10.1146/annurev.energy.30.050504.144511
126. C. Wyborn, Co-productive governance: A relational framework for adaptive governance. *Glob. Environ. Change* **30**, 56–67 (2015). doi: 10.1016/j.gloenvcha.2014.10.009
127. J. Ensor, E. Hoddy, Securing the social foundation: A rights-based approach to planetary boundaries. *Earth Syst. Govern.* **7**, 100086 (2020). doi: 10.1016/j.esg.2020.100086
128. M. von Essen, E. F. Lambin, Jurisdictional approaches to sustainable resource use. *Front. Ecol. Environ.* **19**, 159–167 (2021). doi: 10.1002/fee.2299
129. X. Bai, R. R. McAllister, R. M. Beaty, B. Taylor, Urban policy and governance in a global environment: Complex systems, scale mismatches and public participation. *Curr. Opin. Environ. Sustain.* **2**, 129–135 (2010). doi: 10.1016/j.cosust.2010.05.008
130. E. G. King, D. R. Nelson, J. R. McGreevy, Advancing the integration of ecosystem services and livelihood adaptation. *Environ. Res. Lett.* **14**, 124057 (2019). doi: 10.1088/1748-9326/ab5519
131. E. Cohen-Shacham *et al*., Core principles for successfully implementing and upscaling nature-based solutions. *Environ. Sci. Policy* **98**, 20–29 (2019). doi: 10.1016/j.envsci.2019.04.014
132. M. Di Gregorio *et al*., Multi-level governance and power in climate change policy networks. *Glob. Environ. Change* **54**, 64–77 (2019). doi: 10.1016/j.gloenvcha.2018.10.003
133. I. J. Visseren-Hamakers *et al*., Transformative governance of biodiversity: Insights for sustainable development. *Curr. Opin. Environ. Sustain.* **53**, 20–28 (2021). doi: 10.1016/j.cosust.2021.06.002
134. I. M. Otto *et al*., Social tipping dynamics for stabilizing Earth's climate by 2050. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 2354–2365 (2020). doi: 10.1073/pnas.1900577117; pmid: 31964839
135. S. Sharpe, T. M. Lenton, Upward-scaling tipping cascades to meet climate goals: Plausible grounds for hope. *Clim. Policy* **21**, 421–433 (2021). doi: 10.1080/14693062.2020.1870097
136. X. Bai *et al*., Defining and advancing a systems approach for sustainable cities. *Curr. Opin. Environ. Sustain.* **23**, 69–78 (2016). doi: 10.1016/j.cosust.2016.11.010
137. X. Bai *et al*., Plausible and desirable futures in the Anthropocene: A new research agenda. *Glob. Environ. Change* **39**, 351–362 (2016). doi: 10.1016/j.gloenvcha.2015.09.017
138. C. Folke *et al*., Transnational corporations and the challenge of biosphere stewardship. *Nat. Ecol. Evol.* **3**, 1396–1403 (2019). doi: 10.1038/s41559-019-0978-z; pmid: 31527729
139. J. Timperley, Why fossil fuel subsidies are so hard to kill. *Nature* **598**, 403–405 (2021). doi: 10.1038/d41586-021-02847-2; pmid: 34671143
140. J. Alroy, Effects of habitat disturbance on tropical forest biodiversity. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 6056–6061 (2017). doi: 10.1073/pnas.1611855114; pmid: 28461482
141. R. Fuchs, C. Brown, M. Rounsevell, Europe's Green Deal offshores environmental damage to other nations. *Nature* **586**, 671–673 (2020). doi: 10.1038/d41586-020-02991-1; pmid: 33106645
142. M. Verhoeven, W. Magrath, A. Robbins, E. Kallaur, *Mobilizing and Managing Public Forestry Revenue* (World Bank, 2019); https://elibrary.worldbank.org/doi/abs/10.1596/33086.
143. U. R. Sumaila *et al*., WTO must ban harmful fisheries subsidies. *Science* **374**, 544–544 (2021). doi: 10.1126/science.abm1680; pmid: 34709891
144. J. C. V. Pezzey, The influence of lobbying on climate policies; or, why the world might fail. *Environ. Dev. Econ.* **19**, 329–332 (2014). doi: 10.1017/S1355770X14000345
145. P. Dasgupta, *Economics of Biodiversity: The Dasgupta Review* (Great Britain Treasury, 2021).
146. OECD, *Innovative Citizen Participation and New Democratic Institutions* (OECD, 2020); https://www.oecd-ilibrary.org/content/publication/339306da-en.

147. G. G. Gurney *et al*., Biodiversity needs every tool in the box: Use OECMs. *Nature* **595**, 646–649 (2021). doi: 10.1038/d41586-021-02041-4; pmid: 34312552
148. R. Chaplin-Kramer *et al*., Mapping the planet's critical natural assets. *Nat. Ecol. Evol.* **7**, 51–61 (2023). doi: 10.1038/s41559-022-01934-5; pmid: 36443466
149. C. Chaudhary, A. J. Richardson, D. S. Schoeman, M. J. Costello, Global warming is causing a more pronounced dip in marine species richness around the equator. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2015094118 (2021). doi: 10.1073/pnas.2015094118; pmid: 33876750
150. D. Storch, L. Menzel, S. Frickenhaus, H.-O. Pörtner, Climate sensitivity across marine domains of life: Limits to evolutionary adaptation shape species interactions. *Glob. Chang. Biol.* **20**, 3059–3067 (2014). doi: 10.1111/gcb.12645; pmid: 24890266
151. A. Åkesson *et al*., The importance of species interactions in eco-evolutionary community dynamics under climate change. *Nat. Commun.* **12**, 4759 (2021). doi: 10.1038/s41467-021-24977-x; pmid: 34362916
152. V. R. Burkett *et al*., Point of departure, in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, C. B. Field *et al*., Eds. (Cambridge Univ. Press, Cambridge, 2014), pp. 169–194.
153. K. M. S. Kemppinen *et al*., Global reforestation and biodiversity conservation. *Conserv. Biol.* **34**, 1221–1228 (2020). doi: 10.1111/cobi.13478; pmid: 32017194
154. FAO, ITPS, GSBI, CBD, EC, "State of knowledge of soil biodiversity - Status, challenges and potentialities, Report 2020" (FAO, 2020); https://doi.org/10.4060/cb1928en.
155. A. Wurz *et al*., Win-win opportunities combining high yields with high multi-taxa biodiversity in tropical agroforestry. *Nat. Commun.* **13**, 4127 (2022). doi: 10.1038/s41467-022-30866-8; pmid: 35882849
156. J. Couwenberg, R. Dommain, H. Joosten, Greenhouse gas fluxes from tropical peatlands in south-east Asia. *Glob. Change Biol.* **16**, 1715–1732 (2009). doi: 10.1111/j.1365-2486.2009.02016.x
157. B. W. Griscom *et al*., Natural climate solutions. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11645–11650 (2017). doi: 10.1073/pnas.1710465114; pmid: 29078344
158. A. Dobson *et al*., Savannas are vital but overlooked carbon sinks. *Science* **375**, 392–392 (2022). doi: 10.1126/science.abn4482; pmid: 35084954
159. Y. Bai, M. F. Cotrufo, Grassland soil carbon sequestration: Current understanding, challenges, and solutions. *Science* **377**, 603–608 (2022). doi: 10.1126/science.abo2380; pmid: 35926033
160. V. Enzi, B. Cameron, P. Dezsényi, D. Gedge, G. Mann, U. Pitha, "Nature-Based Solutions and Buildings – The Power of Surfaces to Help Cities Adapt to Climate Change and to Deliver Biodiversity" in *Nature-Based Solutions to Climate Change Adaptation in Urban Areas: Linkages Between Science, Policy and Practice*, Theory and Practice of Urban Sustainability Transitions, N. Kabisch, H. Korn, J. Stadler, A. Bonn, Eds. (Springer, 2017), pp. 159–183; https://doi.org/10.1007/978-3-319-56091-5_10.
161. D. Bianchi, D. A. Carozza, E. D. Galbraith, J. Guiet, T. DeVries, Estimating global biomass and biogeochemical cycling of marine fish with and without fishing. *Sci. Adv.* **7**, eabd7554 (2021). doi: 10.1126/sciadv.abd7554; pmid: 34623923
162. O. J. Schmitz *et al*., Animating the Carbon Cycle. *Ecosystems* **17**, 344–359 (2014). doi: 10.1007/s10021-013-9715-7
163. N. Zwerschke *et al*., Quantification of blue carbon pathways contributing to negative feedback on climate change following glacier retreat in West Antarctic fjords. *Glob. Chang. Biol.* **28**, 8–20 (2022). doi: 10.1111/gcb.15898; pmid: 34658117
164. C. M. Duarte, A. Bruhn, D. Krause-Jensen, A seaweed aquaculture imperative to meet global sustainability targets. *Nat. Sustain.* **5**, 185–193 (2021). doi: 10.1038/s41893-021-00773-9
165. C. M. Duarte, J. Wu, X. Xiao, A. Bruhn, D. Krause-Jensen, Can seaweed farming play a role in climate change mitigation and adaptation? *Front. Mar. Sci.* **4**, (2017). doi: 10.3389/fmars.2017.00100
166. J. A. Gephart *et al*., Environmental performance of blue foods. *Nature* **597**, 360–365 (2021). doi: 10.1038/s41586-021-03889-2; pmid: 34526707
167. N. Hilmi *et al*., The role of blue carbon in climate change mitigation and carbon stock conservation. *Front. Clim.* **3**, 710546 (2021). doi: 10.3389/fclim.2021.710546

ACKNOWLEDGMENTS

This article was initiated as a follow-up to a meeting of 62 scientists from 35 countries, cosponsored by the IPBES and the IPCC. The workshop took place virtually on 14 to 17 December 2020. Technical support was provided by the IPBES secretariat and the IPCC Working Group II TSU. We thank A. Larigauderie for her organizational leadership and Y. Otsuki Estrada, S. Langsdorf, and A. Alegria for their work on the graphics for this manuscript. We thank the following contributors to the Meeting Report, who synthesized the scientific evidence on which this article expands: J. Agard, E. Archer, X. Bai, L. Chan, W. L. Cheung, C. Donatti, N. Eisenhauer, W. Foden, M. A. Gasalla, C. Handa, T. Hickler, O. Hoegh-Guldberg, K. Ichii, U. Jacob, G. Insarov, R. Leemans, L. Levin, M. Lim, S. Maharaj, P. A. Marquet, T. Oberdorff, E. Osman, R. Pandit, A. P. F. Pires, A. Popp, V. Reyes-García, J. Settele, D. W. Sintayehu, P. Smith, N. Steiner, B. Strassburg, R. Sukumar, C. Trisos, J. Wu, E. Aldrian, C. Parmesan, R. Pichs-Madruga, D. C. Roberts, A. D. Rogers, S. Díaz, M. Fischer, S. Hashimoto, S. Lavorel, and N. Wu. Parts of this study are embedded in the Research Unit TERSANE (FOR 2332: Temperature-related stressors as a unifying principle in ancient extinctions). **Author contributions**: H.-O.P. and R.J.S. outlined and led the review; A.A., D.K. A.B., M.T.B., S.E.D., C.M.D., W.K., P.L., S.M., P.M., G.M., H.T.N., D.O., U.P., M.S., Y.J.S., and A.L.V. all contributed to the writing and development of the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Data used in Fig. 3 is from https://doi.org/10.5281/zenodo.7684120. All other data used in the figures and table are available from the cited literature.

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

Submitted 24 April 2022; accepted 13 March 2023
10.1126/science.abl4881

RESEARCH ARTICLE SUMMARY

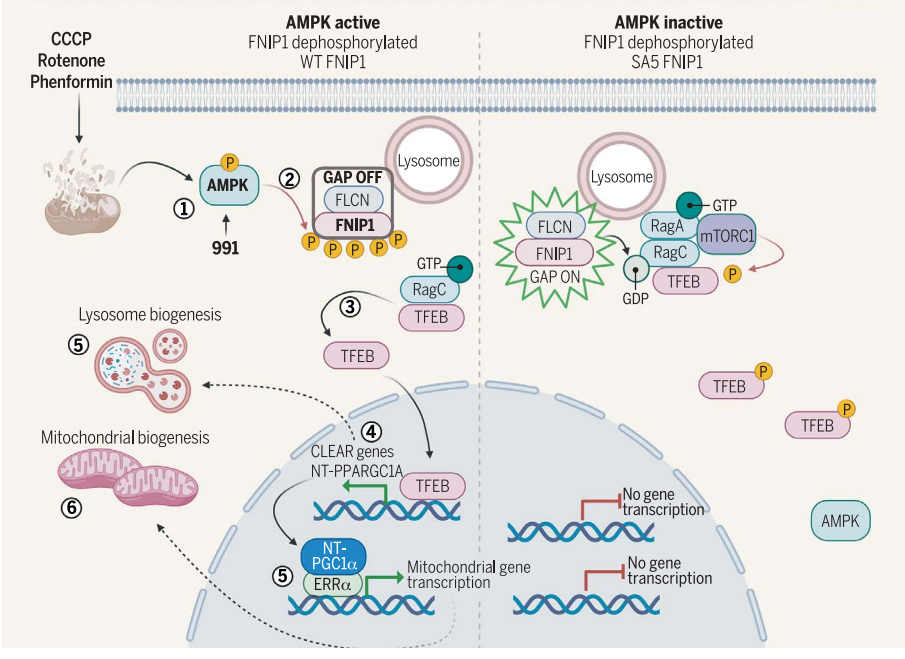
METABOLISM

Induction of lysosomal and mitochondrial biogenesis by AMPK phosphorylation of FNIP1

Nazma Malik, Bibiana I. Ferreira, Pablo E. Hollstein, Stephanie D. Curtis, Elijah Trefts, Sammy Weiser Novak, Jingting Yu, Rebecca Gilson, Kristina Hellberg, Lingjing Fang, Arlo Sheridan, Nasun Hah, Gerald S. Shadel, Uri Manor, Reuben J. Shaw*

INTRODUCTION: Eukaryotes contain a highly conserved signaling pathway that becomes rapidly activated when adenosine triphosphate (ATP) levels decrease, as happens during conditions of nutrient shortage or mitochondrial dysfunction. The adenosine monophosphate (AMP)–activated protein kinase (AMPK) is activated within minutes of energetic stress and phosphorylates a limited number of substrates to biochemically rewire metabolism from an anabolic state to a catabolic state to restore metabolic homeostasis. AMPK also promotes prolonged metabolic adaptation through transcriptional changes, decreasing biosynthetic genes while increasing expression of genes promoting lysosomal and mitochondrial biogenesis. The transcription factor EB (TFEB) is a well-appreciated effector of AMPK-dependent signals, but many of the molecular details of how AMPK controls these processes remain unknown.

RATIONALE: The requirement of AMPK and its specific downstream targets that control aspects of the transcriptional adaptation of metabolism remain largely undefined. We performed time courses examining gene expression changes after various mitochondrial stresses in wild-type (WT) or AMPK knockout cells. We hypothesized that a previously described interacting protein of AMPK, folliculin-interacting protein 1 (FNIP1), may be involved in how AMPK promotes increases in gene expression after metabolic stress. FNIP1 forms a complex with the protein folliculin (FLCN), together acting as a guanosine triphosphate (GTP)–activating protein (GAP) for RagC.



Mitochondrial damage activates AMPK to phosphorylate FNIP1, stimulating TFEB translocation to the nucleus and sequential waves of lysosomal and mitochondrial biogenesis. After mitochondrial damage, activated AMPK phosphorylates FNIP1 (1), causing inhibition of FLCN-FNIP1 GAP activity (2). This leads to accumulation of RagC in its GTP-bound form, causing dissociation of RagC, mTORC1, and TFEB from the lysosome (3). TFEB is therefore not phosphorylated and translocates to the nucleus, inducing transcription of lysosomal or autophagy genes, with parallel increases in NT-PGC1α mRNA (4), which, in concert with ERRα (5), subsequently induces mitochondrial biogenesis (6). CCCP, carbonyl cyanide m-chlorophenylhydrazone; CLEAR, coordinated lysosomal expression and regulation; GDP, guanosine diphosphate; P, phosphorylation. [Figure created using BioRender]

The FNIP1-FLCN complex has emerged as an amino acid sensor to the mechanistic target of rapamycin complex 1 (mTORC1), involved in how amino acids control TFEB activation. We therefore examined whether AMPK may regulate FNIP1 to dominantly control TFEB independently of amino acids.

RESULTS: AMPK was found to govern expression of a core set of genes after various mitochondrial stresses. Hallmark features of this response were activation of TFEB and increases in the transcription of genes specifying lysosomal and mitochondrial biogenesis. AMPK directly phosphorylated five conserved serine residues in FNIP1, suppressing the function of the FLCN-FNIP1 GAP complex, which resulted in dissociation of RagC and mTOR from the lysosome, promoting nuclear translocation of TFEB even in the presence of amino acids. FNIP1 phosphorylation was required for AMPK to activate TFEB and for subsequent increases in peroxisome proliferation–activated receptor gamma coactivator 1-alpha (PGC1α) and estrogen-related receptor alpha (ERRα) mRNAs. Cells in which the five serines in FNIP1 were mutated to alanine were unable to increase lysosomal and mitochondrial gene expression programs after treatment with mitochondrial poisons or AMPK activators despite the presence and normal regulation of all other substrates of AMPK. By contrast, neither AMPK nor its control of FNIP1 were needed for activation of TFEB after amino acid withdrawal, illustrating the specificity to energy-limited conditions.

CONCLUSION: Our data establish FNIP1 as the long-sought substrate of AMPK that controls TFEB translocation to the nucleus, defining AMPK phosphorylation of FNIP1 as a singular event required for increased lysosomal and mitochondrial gene expression programs after metabolic stresses. This study also illuminates the larger biological question of how mitochondrial damage triggers a temporal response of repair and replacement of damaged mitochondria: Within early hours, AMPK-FNIP1–activated TFEB induces a wave of lysosome and autophagy genes to promote degradation of damaged mitochondria, and a few hours later, TFEB–up-regulated PGC1α and ERRα promote expression of a second wave of genes specifying mitochondrial biogenesis. These insights open therapeutic avenues for several common diseases associated with mitochondrial dysfunction, ranging from neurodegeneration to type 2 diabetes to cancer. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: shaw@salk.edu
Cite this article as N. Malik *et al.*, *Science* **380**, eabj5559 (2023). DOI: 10.1126/science.abj5559

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abj5559

RESEARCH ARTICLE

METABOLISM

Induction of lysosomal and mitochondrial biogenesis by AMPK phosphorylation of FNIP1

Nazma Malik[1], Bibiana I. Ferreira[1], Pablo E. Hollstein[1], Stephanie D. Curtis[1], Elijah Trefts[1], Sammy Weiser Novak[2], Jingting Yu[3], Rebecca Gilson[2], Kristina Hellberg[1], Lingjing Fang[2], Arlo Sheridan[2], Nasun Hah[4], Gerald S. Shadel[1], Uri Manor[2], Reuben J. Shaw[1]*

Cells respond to mitochondrial poisons with rapid activation of the adenosine monophosphate–activated protein kinase (AMPK), causing acute metabolic changes through phosphorylation and prolonged adaptation of metabolism through transcriptional effects. Transcription factor EB (TFEB) is a major effector of AMPK that increases expression of lysosome genes in response to energetic stress, but how AMPK activates TFEB remains unresolved. We demonstrate that AMPK directly phosphorylates five conserved serine residues in folliculin-interacting protein 1 (FNIP1), suppressing the function of the folliculin (FLCN)–FNIP1 complex. FNIP1 phosphorylation is required for AMPK to induce nuclear translocation of TFEB and TFEB-dependent increases of peroxisome proliferator–activated receptor gamma coactivator 1-alpha (PGC1α) and estrogen-related receptor alpha (ERRα) messenger RNAs. Thus, mitochondrial damage triggers AMPK-FNIP1–dependent nuclear translocation of TFEB, inducing sequential waves of lysosomal and mitochondrial biogenesis.

The ability to adapt to prolonged nutrient deprivation is an essential characteristic for survival of all organisms. Eukaryotes contain a highly conserved signaling pathway that becomes rapidly activated when adenosine triphosphate (ATP) levels in the cell decrease, most often as a result of loss of mitochondrial ATP production caused by decreased oxygen or glucose concentrations or in response to mitochondrial poisons that directly interfere with oxidative phosphorylation (OXPHOS) (*1*). The adenosine monophosphate (AMP)–activated protein kinase (AMPK) becomes activated fully within minutes of OXPHOS inhibition and rapidly phosphorylates a limited number of direct substrates that regulate lipid and glucose metabolism, autophagy, and mechanistic target of rapamycin complex 1 (mTORC1) signaling (*2*, *3*). If energetic stress is prolonged, metabolism is further modified by transcriptional changes to gene expression programs governing distinct metabolic processes (*4*). The transcription factor EB (TFEB) and related TFE3 are activated in response to nutrient deprivation and energetic stress, and both are suppressed by mTORC1 signaling (*5*–*9*) and activated by AMPK signaling (*10*, *11*). mTORC1 directly phosphorylates TFEB on Ser^{122}, Ser^{142}, and Ser^{211}, resulting in its exclusion from the nucleus (*6*).

Amino acids (AAs) regulate the ability of mTORC1 to phosphorylate TFEB through a guanosine triphosphate (GTP)–activating protein (GAP) complex composed of the folliculin (FLCN) and FLCN-interacting protein 1 (FNIP1) proteins (*12*–*15*). The FLCN-FNIP1 complex dictates GTP loading of the guanosine triphosphatase (GTPase) RagC, which results in the release of TFEB and TFE3 from the lysosome and away from mTORC1, causing their nuclear translocation (*14*). In the nucleus, TFEB and TFE3 directly bind to a well-defined DNA binding element [coordinated lysosomal expression and regulation (CLEAR) motif] found conserved in the proximal promoters of >50 components of the lysosome and a number of autophagy genes (*16*, *17*). In this way, TFEB and TFE3 are held inactive under nutrient-replete conditions, but in response to specific cellular stresses, they translocate into the nucleus and promote lysosomal biogenesis and autophagy (*18*). AMPK is required for TFEB translocation to the nucleus during energetic stress (*10*, *11*, *19*). How AMPK activates TFEB remains unknown, but it is presumed to rely on AMPK-dependent suppression of mTORC1 through its established phosphorylation of the mTORC1 component Raptor and upstream regulator TSC2. We examined the possibility that TFEB and TFE3 were direct substrates of AMPK, but we did not find any evidence supporting this in vivo. Recently, AMPK has been reported to phosphorylate TFEB at three C-terminal residues, and these phosphorylation events may have a role in lysosomal gene transcription (*20*). However, all three of these tightly clustered sites poorly match the AMPK substrate consensus motif, and mutation of these sites does not disrupt the ability of AMPK to induce nuclear translocation of TFEB and TFE3 (*20*). Thus, an unknown AMPK-dependent event may govern the translocation of TFEB and TFE3 to the nucleus, without which CLEAR gene transcription cannot occur, even if TFEB and TFE3 are phosphorylated at their C terminus by AMPK-dependent signals.

The precise mechanism of how AMPK activates the transcriptional program of mitochondrial biogenesis is also unresolved. AMPK controls mitochondrial biogenesis and synthesis of peroxisome proliferator–activated receptor gamma coactivator 1-alpha (PGC1α) mRNA in response to energetic stress (*21*–*24*). Several mechanisms for how AMPK may promote accumulation of PGC1α mRNA and PGC1α function have been proposed, including control of PGC1α phosphorylation or acetylation (*25*, *26*), but the precise mechanism or mechanisms remain poorly understood. We performed time course analysis of the transcriptional response to mitochondrial OXPHOS inhibitors, revealing a temporal cascade of organellar biogenesis genetically dependent on AMPK. We identified FNIP1 as a direct AMPK substrate whose phosphorylation is critical for TFEB activation and nuclear translocation, which in turn leads to the production of PGC1α and estrogen-related receptor alpha (ERRα) mRNAs, resulting in a wave of lysosomal biogenesis followed by mitochondrial biogenesis.

Electron transport chain inhibitors require AMPK to induce mitochondrial gene transcription

To delineate the role of AMPK in the transcriptional response to mitochondrial energetic stress, we disrupted AMPKα1 and AMPKα2 in human embryonic kidney 293T (HEK293T) cells [AMPK knockout (KO)] by CRISPR-Cas9 and subjected wild-type (WT) control and AMPK KO cells to electron transport chain (ETC) inhibitors, including 100 ng/ml rotenone (complex I), 2 mM phenformin (complex I), and 5 μM CCCP (carbonyl cyanide m-chlorophenylhydrazone; a protonophore that disrupts the electrochemical gradient required for ATP production). We performed a time course of RNA sequencing (RNA-seq) of up to 16 hours. Differential expression analysis of all genes in the RNA-seq dataset revealed the gene expression patterns induced by each of the drugs. Differential expression defined by fold change (FC) ≥ 1.3 and $P \leq 0.05$ (P values generated from t tests were corrected using the Benjamini-Hochberg method) followed by hierarchical clustering uncovered a set of common genes, whose expression was induced by CCCP, rotenone, and phenformin in WT controls but not in cells lacking AMPK (fig. S1 and Fig. 1A). Approximately one-third of those genes with transcription increased similarly by the three ETC inhibitors in the first 16 hours required

[1]Molecular and Cell Biology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA, USA. [2]Biophotonics Core, The Salk Institute for Biological Studies, La Jolla, CA, USA. [3]Bioinformatics Core, The Salk Institute for Biological Studies, La Jolla, CA, USA. [4]Next Generation Sequencing Core, The Salk Institute for Biological Studies, La Jolla, CA, USA.
*Corresponding author. Email: shaw@salk.edu

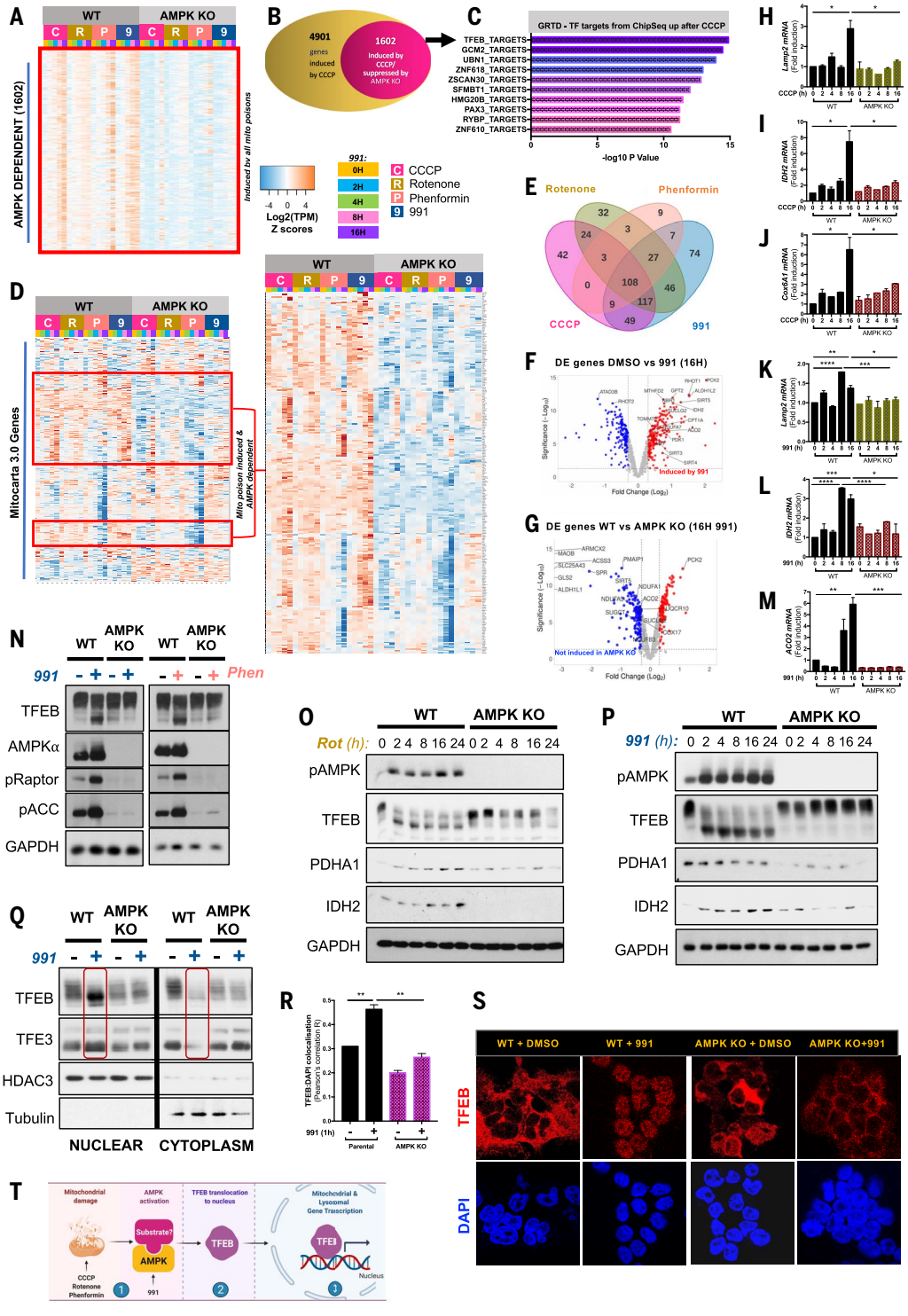


Fig. 1. Dominant role of AMPK in the transcriptional response to mitochondrial poisons through the MiT-TFE family of transcription factors. RNA-seq analysis of WT and CRISPR-Cas9–mediated AMPK KO HEK293T cells upon 0- to 16-hour treatment with the mitochondrial poisons CCCP (5 μM), rotenone (100 ng/ml), phenformin (2 mM), and the AMPK-specific activation drug 991 (50 μM). (**A**) Unbiased heatmap displaying gene expression pattern of all AMPK-dependent, differentially expressed (DE) genes (FC ≥ 1.3, $P \leq 0.05$) commonly regulated by all three mitochondrial poisons and 991. (**B**) Stacked Venn diagram showing the proportion of DE CCCP-induced genes that require AMPK. (**C**) GSEA analysis shows significantly up-regulated GTRD (ChIP-seq–based Gene Transcription Regulation Database) transcription factor targets upon CCCP treatment. (**D**) Gene clustering analysis and heatmap displaying the expression pattern of all mitochondria-specific genes as defined by the Mitocarta 3.0 inventory. Right heatmap is a zoomed-in view of the AMPK-dependent mitochondrial genes induced by the four drugs. (**E**) Overlap in regulation of AMPK-dependent mitochondrial genes in (D) by CCCP, rotenone, phenformin, and 991. (**F**) Volcano plot depicting DE mitochondrial genes from (D) after 991 compared with DMSO. Red dots represent genes significantly induced by 991 compared with DMSO. The *y* axis denotes $-\log_{10} P$ values, and the *x* axis shows $\log_2$ FC values. (**G**) Volcano plot denoting differential expression of mitochondrial genes between WT 16-hour 991-treated cells compared with AMPK KO 16-hour 991-treated cells. Blue dots represent genes significantly down-regulated by AMPK deletion compared with WT AMPK condition. The *y* axis denotes $-\log_{10} P$ values, and the *x* axis shows $\log_2$ FC values. (**H** to **J**) Quantitative RT-PCR (qRT-PCR) for lysosomal gene Lamp2 (H), mitochondrial genes IDH2 (I), and Cox6A1 (J) in WT and AMPK KO HEK293T cells after CCCP. (**K** to **M**) qRT-PCR for lysosomal gene Lamp2 (K) and mitochondrial genes IDH2 (L) and ACO2 (M) in WT and AMPK KO HEK293T cells after 991 treatment. All qRT-PCR graphs are shown as the means ± SEMs. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$; unpaired *t* test. (**N**) Analysis of AMPK signaling and TFEB protein immunoblotting of WT and AMPK KO HEK293T cells treated with DMSO, 991, or phenformin for 1 hour. (**O**) Analysis of AMPK, TFEB, and mitochondrial protein immunoblotting of WT and AMPK KO HEK293T cells treated with a rotenone (100 ng/ml) time course, ranging from 0 to 24 hours. (**P**) Analysis of AMPK, TFEB, and mitochondrial protein immunoblotting of WT and AMPK KO HEK293T cells treated with a 991 (50 μM) time course, ranging from 0 to 24 hours. (**Q**) Analysis of TFEB and TFE3 protein cytoplasm-to-nuclear shuttling by fractionation of WT and AMPK KO HEK293T cells with or without 1-hour 991 treatment, followed by immunoblotting. (**R**) Quantification of TFEB colocalization with DAPI-stained nuclei. Data are shown as the means ± SEMs of three independent experiments. *$P < 0.05$; **$P < 0.01$; unpaired *t* test. (**S**) Representative images from immunofluorescence microscopy of endogenous TFEB stained in WT and AMPK KO HEK293T cells, pretreated with 1 hour of DMSO or 991. Nuclei are stained with DAPI. (**T**) Model. After ETC poisons or the direct small-molecule activator 991, AMPK becomes activated and, through an unknown mechanism, triggers TFEB translocation to the nucleus, where TFEB induces expression of lysosomal and mitochondrial genes.

AMPK for full gene induction (Fig. 1, A and B). Analysis of this common, AMPK-dependent, and mitochondrial energetic stress–induced gene set at 2- and 4-hour time points—using the Gene Transcription Regulatory Database (GTRD) of chromatin immunoprecipitation sequencing (ChIP-seq) datasets—revealed TFEB to be the most enriched transcription factor (Fig. 1C) based on the binding of transcription factors to the −1000 to +100 bases around the transcription start site of each differentially expressed gene (*27*).

Volcano plots illustrating genes with the greatest FC and highest degree of statistical significance revealed core components of lysosomes and mitochondria as some of the most increased genes in an AMPK-dependent manner in cells exposed to CCCP for 16 hours (fig. S2, A and B); for example, *EPDR1* (ependymin related 1) is a lysosomal protein that binds gangliosides, *OXCT1* [3-oxoacid coenzyme A (CoA)–transferase 1] encodes a mitochondrial enzyme in ketone body catabolism, and *ACSS3* (acyl-CoA synthetase short chain family member 3) encodes a mitochondrial enzyme in fatty acid oxidation. Gene ontology (GO) enrichment analysis on AMPK-responsive genes in cells treated with CCCP revealed that "lysosomal lumen" and "secretory granule lumen" were among the most enriched GO organelle terms of the AMPK-dependent, CCCP-induced gene set, and "Electron Transport Chain (OXPHOS)" was the second-most enriched WikiPathways term (fig. S2C). Gene set enrichment analysis (GSEA) of the CCCP-induced RNA-seq dataset confirmed enrichment of "KEGG Lysosome" and "Hallmarks OXPHOS" gene sets in WT cells but not in cells lacking AMPK treated with CCCP for 16 hours (fig. S2D). Parallel analysis of the genes differentially expressed in response to rotenone in an AMPK-dependent manner revealed similar enrichment in GO for lysosomes and enrichment for mitochondrial processes in WikiPathways analyses, again with statistically significant enrichment for the "KEGG Lysosome" and "Hallmarks OXPHOS" gene sets (fig. S2, E and F). In addition to observing a common transcriptional response between CCCP, rotenone, and phenformin, we also found that the synthetic small-molecule, direct AMPK activator 991 (*28*) similarly enriches for the same lysosomal genes and mitochondrial targets (fig. S2, G and H). Transcription factor enrichment analysis of 991-induced genes in the WT condition identified *ESRRA* as a transcription factor, whose targets are overrepresented in the up-regulated genes (fig. S2I). The estrogen-related receptor alpha gene, *ESRRA*, which encodes ERRα, is a known key mediator of mitochondrial biogenesis (*29*, *30*).

Having observed a strong lysosomal and mitochondrial component to the common differentially expressed genes upon treatment with all four drugs, we next examined mitochondrial genes more comprehensively. We reanalyzed our data, using the MitoCarta version 3.0, a curated catalog of ~1000 genes encoding the mammalian mitochondrial proteome (*31*), to assess mitochondrial biogenesis. There was significant overlap in the mitochondrial genes increased after CCCP, rotenone, and CCCP with those increased by 991 (Fig. 1, D and E). The analysis revealed increased transcription of ~300 mitochondrial genes in WT cells treated with ETC poisons or small-molecule, AMPK activator 991 but not cells lacking AMPK (Fig. 1, D, F, and G). Despite the differences in the mechanisms by which these compounds inhibit the ETC, this analysis indicates that a large proportion of the transcriptional response mediated by ETC poisons requires AMPK activation. We validated the increased expression of core lysosomal and mitochondrial genes, including *LAMP2* (Fig. 1H), *IDH2* (Fig. 1I), *Cox6A1*, or *ACO2* (Fig. 1J), in response to CCCP and 991 (Fig. 1, K to M), all of which showed increased expression in WT cells but not those lacking AMPK when assessed by quantitative polymerase chain reaction (qPCR).

Given the prominent role for TFEB implied by the regulation of genes induced by mitochondrial poisons (Fig. 1C) and TFEB's major role in AMPK-dependent effects on transcription (*9*, *10*, *17*), we examined the regulation of TFEB protein in WT or AMPK KO cells treated with mitochondrial poisons. TFEB isolated from 991- (1 hour) or phenformin- (1 hour) treated WT cells showed a reduced mobility band-shift of endogenous TFEB, which did not occur in TFEB from AMPK KO cells (Fig. 1N). We examined TFEB electrophoretic mobility over longer time courses of treatment with rotenone and 991 and again observed a persistent downward mobility shift in WT cells between 2 and 24 hours after treatment, which was barely detectable in the AMPK KO cells (Fig. 1, O and P). We examined protein abundance of some mitochondrial targets implicated by RNA-seq or qPCR and observed up-regulation of PDHA1 and IDH2 after 991 or rotenone, which was not observed in AMPK KO cells (Fig. 1, O and P).

We assessed the localization of TFEB and TFE3 through nucleo-cytoplasmic fractionation. In WT cells, 991 treatment led to increased abundance of TFEB and TFE3 in the nuclear fraction with almost none remaining in the cytoplasmic compartment (Fig. 1Q). In cells lacking AMPK, 991 treatment did not lead to translocation of TFEB or TFE3 to the nuclear fractions. Immunofluorescence microscopy of intracellular TFEB with an antibody that detects endogenous TFEB confirmed these results. In WT cells under basal conditions, TFEB was mostly in the cytoplasm with low amounts in the nucleus (Fig. 1, R and S). Upon AMPK activation by 991, conditions under which TFEB is fully dephosphorylated, almost all TFEB translocated to the nucleus. Conversely, in AMPK KO cells, with or without 991, most TFEB remained in the cytoplasm (Fig. 1, R and S). These results indicate that TFEB is a major effector of AMPK after mitochondrial energetic stress (Fig. 1T).

FNIP1 is a conserved AMPK substrate that governs TFEB phosphorylation status and localization

A screen we performed looking for AMPK substrates that could mediate cell growth and metabolism (*32*) identified FNIP1. FNIP1 is an established interacting partner of FLCN and has been reported to coimmunoprecipitate with AMPK and be regulated by AMPK through unknown details (*33*, *34*). The FNIP1-FLCN complex has emerged as an AA sensor to mTORC1 (*11*–*14*), involved in how AAs control TFEB activation (*14*). We therefore examined whether AMPK may regulate FNIP1 to control TFEB independently of AAs. Analysis of the FNIP1 protein sequence revealed four sites that match the optimal AMPK substrate motif (Ser^{230}, Ser^{232}, Ser^{261}, and Ser^{593}) (Fig. 2A). We used mass spectrometry (MS) to examine the phosphorylation of FNIP1 in vivo. We transfected HEK293T cells with an epitope-tagged FNIP1 cDNA, and the cells treated with dimethyl sulfoxide (DMSO) or phenformin. We detected peptides spanning all four (Ser^{230}, Ser^{232}, Ser^{261}, and Ser^{593}) of these FNIP1 candidate sites (Fig. 2B), phosphorylation of which had increased in the phenformin-treated samples. MS analysis detected a fifth site, Ser^{220}. This also appeared to be highly phosphorylated after AMPK activation and largely conforms to the optimal motif (Fig. 2A), indicating that it too may be a site of AMPK phosphorylation (Fig. 2B).

To test whether FNIP1 is a direct substrate of AMPK, we undertook an in vitro ^{32}P-γ-MgATP phosphorylation assay. We immunopurified FLAG epitope–tagged WT FNIP1 or FNIP1 S-A mutant proteins (Fig. 2C) after transient transfection of cDNA into HEK293T cells. Recombinant AMPK phosphorylated WT FNIP1 and mutation of the five candidate AMPK-site serines (Ser^{230}, Ser^{232}, Ser^{261}, Ser^{593}, and Ser^{220}) showed the greatest effect, seen when all five sites were mutated (hereafter referred to as the SA5 mutant) (Fig. 2C). Notably, variability in the behavior and lack of effect on overall ^{32}P incorporation of some mutants in the in vitro phosphorylation assay leaves the possibility open that some of the sites are not directly regulated by AMPK or are redundantly phosphorylated, masking the impact of loss of one or two sites as judged by overall ^{32}P incorporation in vitro. To verify FNIP1 as an in vivo substrate of AMPK, we generated a phospho-specific antibody, targeting the pFNIP1 Ser^{220} site. To validate the Ser^{220} antibody, we used HEK293T cells lacking FNIP1 generated by

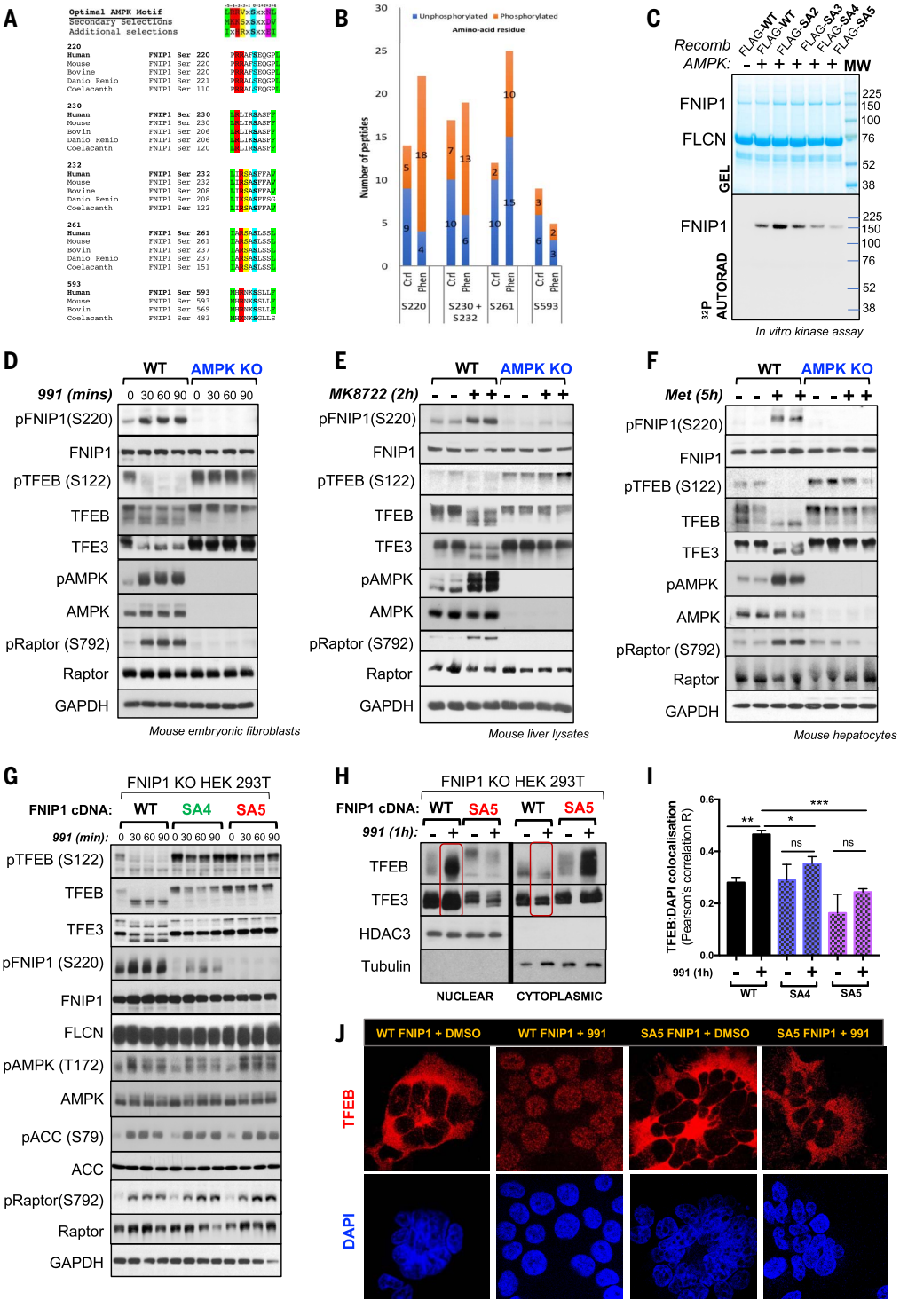


Fig. 2. FNIP1 as a conserved AMPK substrate that regulates TFEB and TFE3. (**A**) ClustalW alignment of five conserved AMPK phosphorylation sites on FNIP1 matching the AMPK substrate consensus motif. (**B**) AMPK phosphorylation sites on FNIP1 identified by MS in HEK293T cells, expressing FLAG-tagged FNIP1 cDNA, after treatment with vehicle or 1 hour of phenformin. (**C**) In vitro AMPK kinase assay. WT FNIP1 or FNIP1 mutants (SA2, SA3, SA4, SA5) were immunoprecipitated and incubated with recombinant active AMPK and [γ^{32}P]-ATP. Kinase reactions were separated by SDS-PAGE and incorporation of [γ^{32}P]-ATP was detected by autoradiography. (**D**) Immunoblot showing endogenous TFEB and FNIP1 (Ser^{220}) phosphorylation status in WT ($^{+/+}$) or AMPK DKO ($^{-/-}$) MEFs after DMSO or a 991 time course, as detected by an antibody to FNIP1 P-Ser^{220}. (**E**) Immunoblots of murine liver lysates from inducible KO of AMPKα1 and α2 (Alb-CreERT2; $Prkaa1^{fl/fl}$; $Prkaa2^{fl/fl}$) mice (AMPK KO) or control mice treated with vehicle or AMPK activator compound MK8722 for 2 hours, showing endogenous FNIP1 and TFEB phosphorylation status. (**F**) Immunoblots of lysates from primary mouse hepatocytes from WT mice or Alb-Cre ($Prkaa1^{fl/fl}$; $Prkaa2^{fl/fl}$) mice (AMPK liver DKO), treated with vehicle or metformin for 5 hours. (**G**) Immunoblots of endogenous pFNIP1 Ser^{220} and TFEB in in FNIP1 KO HEK293T cells stably reconstituted with WT, SA4, or SA5 FNIP1, treated with a 991 (50 μM) time course. (**H**) Nucleocytoplasmic fractionation of WT or SA5 FNIP1 HEK293T cells after 1-hour DMSO or 991 treatment, followed by immunoblotting. (**I**) Quantitation of TFEB colocalization with DAPI-stained nuclei in (J). Data are shown as the means ± SEMs of three independent experiments. *P < 0.05; **P < 0.01; ***P < 0.001; unpaired t test. (**J**) Representative immunofluorescence images of endogenous TFEB stained in WT FNIP1 and SA5-FNIP1 HEK293T cells, treated with 1-hour DMSO or 991. Nuclei are stained with DAPI.

CRISPR-Cas9 methodology (fig. S3A) and then transiently transfected those cells with cDNA for WT FNIP1, S220A, SA3 (S-A mutations of Ser^{230}, Ser^{232}, and Ser^{261}), SA4 (S-A mutations of Ser^{230}, Ser^{232}, Ser^{261}, and Ser^{593}), and SA5 FNIP1 mutants. These cells were treated with either DMSO or 991, and then phosphorylation status of FNIP1 was assessed with the pFNIP1 Ser^{220} antibody. The antibody detected phosphorylation of the S220 site in WT FNIP1, SA3, and SA4 FNIP1 conditions but not in the S220A or SA5 condition, in which the Ser^{220} site had been mutated to alanine (fig. S3B). A phospho-specific antibody was also generated to the Ser^{261} site, which was only effective after FNIP1 immunoprecipitation; nonetheless, we observed increased Ser^{261} phosphorylation in cells overexpressing WT FNIP1 but not SA5 FNIP1, in which Ser^{261} is mutated to alanine (fig. S3C).

To assess phosphorylation of endogenous FNIP1, we treated mouse embryonic fibroblasts (MEFs) with 991. Endogenous FNIP1 was robustly phosphorylated at the Ser^{220} site within 30 min, whereas TFEB and TFE3 were dephosphorylated under the same conditions in WT but not AMPK KO MEFs (Fig. 2D). We also

examined phosphorylation of endogenous FNIP1 in liver lysates of mice after treatment of mice with MK-8722, an orally available 991 analog that also activates AMPK (*35*). We detected Ser^{220} phosphorylation of FNIP1 in livers from WT mice treated with MK-8722 but not in livers from AMPKa1/a2 (*Prkaa1*$^{l/l}$; *Prkaa2*$^{l/l}$) liver-specific KO mice (Fig. 2E). Other AMPK substrates, such as pRaptor Ser^{792}, were also not phosphorylated in these mice. TFEB and TFE3 were dephosphorylated in livers from WT animals treated with MK-8722 but not in livers lacking AMPK (Fig. 2E). Treatment of primary hepatocytes with metformin induced phosphorylation of endogenous FNIP1 at Ser^{220} in WT hepatocytes but not hepatocytes lacking AMPK (Fig. 2F). TFEB and TFE3 became dephosphorylated in WT hepatocytes treated with metformin but remained phosphorylated in hepatocytes lacking AMPK, even after metformin administration (Fig. 2F). Thus, FNIP1 appears to be a bona fide, in vitro and in vivo substrate of AMPK, phosphorylated upon direct AMPK stimulation and in response to mitochondrial energetic stress.

Given that both mTOR and AMPK antagonistically regulate TFEB and TFE3 transcription factors and that both pathways converge on the FNIP1-FLCN complex, phosphorylation of FNIP1 by AMPK could represent the dominant mechanism through which AMPK controls these transcription factors. To test for such a role for FNIP1, we used lentivirus to stably reconstitute cells depleted of FNIP1 with full-length cDNA encoding either WT FNIP1, SA4 FNIP1, or SA5 FNIP1 and treated these cells with 991 (Fig. 2G) or phenformin (fig. S3D). In WT FNIP1 cells, AMPK activation led to dephosphorylation of TFEB and TFE3 within 30 to 60 min of 991 or phenformin addition (Fig. 2G and fig. S3D). However, in the SA4 and SA5 FNIP1 cells, TFEB and TFE3 remained in the slow-mobility hyperphosphorylated form whether or not cells were treated with 991 or phenformin (Fig. 2G and fig. S3D). This effect was further enhanced when all five AMPK sites on FNIP1 were mutated. TFEB was fully phosphorylated in FNIP1 SA5 cells, even in the presence of phenformin or 991 (Fig. 2G and fig. S3D). In our prior studies, AMPK promoted activation of TFEB in cells deprived of glucose, but the mechanism was unknown (*11*). To test whether AMPK control of FNIP1 was involved in the regulation of TFEB after glucose starvation, we examined TFEB in WT FNIP1 and SA5 FNIP1 cells deprived of glucose. Six hours of glucose deprivation activated AMPK, leading to FNIP1 Ser^{220} phosphorylation in WT FNIP1 cells and TFEB dephosphorylation, as did treatment with 991 (fig. S3E). Conversely in SA5 cells, mutation of AMPK phosphosites on FNIP1 prevented TFEB dephosphorylation induced by glucose starvation (fig. S3E). Furthermore, given the importance of Ser^{220} as the one serine different between SA4 and SA5 and the apparent in vitro phosphorylation of Ser^{593} (Fig. 2C), we examined the effect of a Ser^{220}-Ser^{593} AA mutant (SA2), which revealed incomplete regulation of TFEB, again implying a critical role for all five serines, including 230, 232, and 261 (fig. S3F).

If AMPK phosphorylation of FNIP1 is necessary for TFEB dephosphorylation, localization of TFEB and TFE3 might be expected to change upon FNIP1 phosphorylation. To test this, we isolated nuclear and cytoplasmic fractions from WT FNIP1 cells treated with 991. In contrast to the control, after 991, TFEB was enriched in the nucleus with little cytoplasmic TFEB detected (Fig. 2H). However, after 991 treatment in SA5 FNIP1 cells, most TFEB remained in the cytoplasm (Fig. 2H). Immunofluorescence microscopy to visualize endogenous TFEB and TFE3 confirmed this observation. In WT FNIP1 cells in fresh media, both endogenous TFEB (Fig. 2, I and J) and endogenous TFE3 (fig. S3, G and H) displayed a fully cytoplasmic localization, but upon 991 treatment, both translocated to the nucleus despite the presence of full AAs. Conversely, in SA5 FNIP1 cells, AMPK activation by 991 did not cause translocation of TFEB or TFE3 to the nucleus, and both were primarily cytoplasmic (Fig. 2, I and J, and fig. S3, G and H). Thus, phosphorylation of FNIP1 by AMPK appears to dominantly govern localization of TFEB (fig. S3I).

AMPK-dependent phosphorylation of FNIP1 controls mTORC1 binding to TFEB and TFE3

Although TFEB is phosphorylated by extracellular signal–regulated kinase 1 (ERK), glycogen synthase kinase-3 (GSK3), and protein kinase B (PKB or Akt), much regulation of TFEB is thought to be through phosphorylation of Ser^{122} and Ser^{211} by mTORC1 (*6*). To delineate TFEB regulation by AMPK versus mTORC1, we treated WT HEK293T, AMPK KO HEK293T, WT FNIP1, and SA5 FNIP1 HEK293T cells with 991. TFEB and TFE3 became dephosphorylated in WT cells and WT FNIP1 cells within 30 min but remained highly phosphorylated in AMPK KO and SA5 FNIP1 cells (Fig. 3, A and B). mTORC1 signaling was attenuated by 991 in WT HEK293T and WT FNIP1 HEK293T cells within 30 min of treatment, as reflected by the decreased phosphorylation of the direct mTORC1 substrates eukaryotic translation initiation factor 4E (eIF4E)–binding protein 1 (4EBP1) and ribosomal protein S6 kinase beta-1 (P70S6K), and in turn, the P70S6K substrate ribosomal protein S6 (S6). By contrast, in AMPK KO HEK293T cells, mTORC1 signaling did not change with 991 treatment, because of the absence of AMPK (Fig. 3A). However, in both WT and SA5 FNIP1 cells, canonical mTORC1 signaling to P-S6K and 4EBP1 was suppressed upon AMPK activation with 991, whereas TFEB and TFE3 remained fully phosphorylated in the SA5 cells (Fig. 3B). This revealed a disconnect where FNIP1 modification affected mTORC1-dependent phosphorylation of TFEB but not mTORC1 phosphorylation of S6K1 and 4EBP1, consistent with multiple recent studies (*15*, *36*, *37*).

We examined the effect of AA abundance on TFEB phosphorylation status and mTORC1 activity in the context of our FNIP1 phosphorylation site mutants. mTORC1 signaling, as reflected by phosphorylation of p70S6K, S6, and 4EBP1, was decreased in both WT and AMPK KO HEK293T cells deprived of AAs (Fig. 3C). In cells deprived of AAs, TFEB and TFE3 became dephosphorylated in both WT and AMPK KO cells (Fig. 3C), contrasting with cells treated with 991, where TFEB and TFE3 became dephosphorylated in WT but not AMPK KO HEK293T cells (Fig. 3A). Similarly, mTORC1 signaling was decreased after AA deprivation in both WT FNIP1 and SA5 FNIP1 cells. TFEB and TFE3 also became dephosphorylated when both WT FNIP1 and SA5 FNIP1 cells were deprived of AAs for 1 hour (fig. S4A), whereas activation of AMPK with 991 caused TFEB to be dephosphorylated in WT FNIP1 cells but not SA5 FNIP1 cells. These findings reveal that AMPK-FNIP1–mediated control of the microphthalmia–transcription factor E (MiT-TFE) family of transcription factors through mTORC1 is distinct from that caused by a shortage of AAs. However, dephosphorylation of TFEB and TFE3 dephosphorylation in the SA5 FNIP1 cells deprived of AAs was slower than that observed in WT FNIP1 cells, which demonstrates the inherent control that FNIP1 exerts on TFEB and TFE3 in general (fig. S4A).

To clarify whether TFEB phosphorylation remaining in SA5 cells after AMPK activation was from uninhibited mTORC1-dependent phosphorylation or an unknown kinase, we treated WT and SA5 FNIP1 cells with two potent and selective inhibitors of mTOR, AZD8055 and Torin1, with or without 991. Both AZD8055 and Torin1 decreased TFEB phosphorylation in cells expressing SA5 FNIP1, demonstrating that the TFEB hyperphosphorylation observed in these cells was indeed a result of mTORC1 (Fig. 3D). Although phosphorylation of other mTOR substrates, such as S6K1 or 4EBP1, were not affected by mutation of AMPK sites on FNIP1, TFEB specifically was highly regulated by the AMPK-dependent phosphorylation of FNIP1. To further understand the regulation of TFEB in this context, we generated WT FNIP1 and SA5 FNIP1 HEK293T cells that stably express TFEB tagged with green fluorescent protein (GFP) and subjected these cells to 991 treatment (Fig. 3E). GFP-TFEB immunoprecipitates from WT FNIP1 cells revealed that AMPK activation by 991 caused TFEB dissociation from mTOR and Raptor. Conversely in

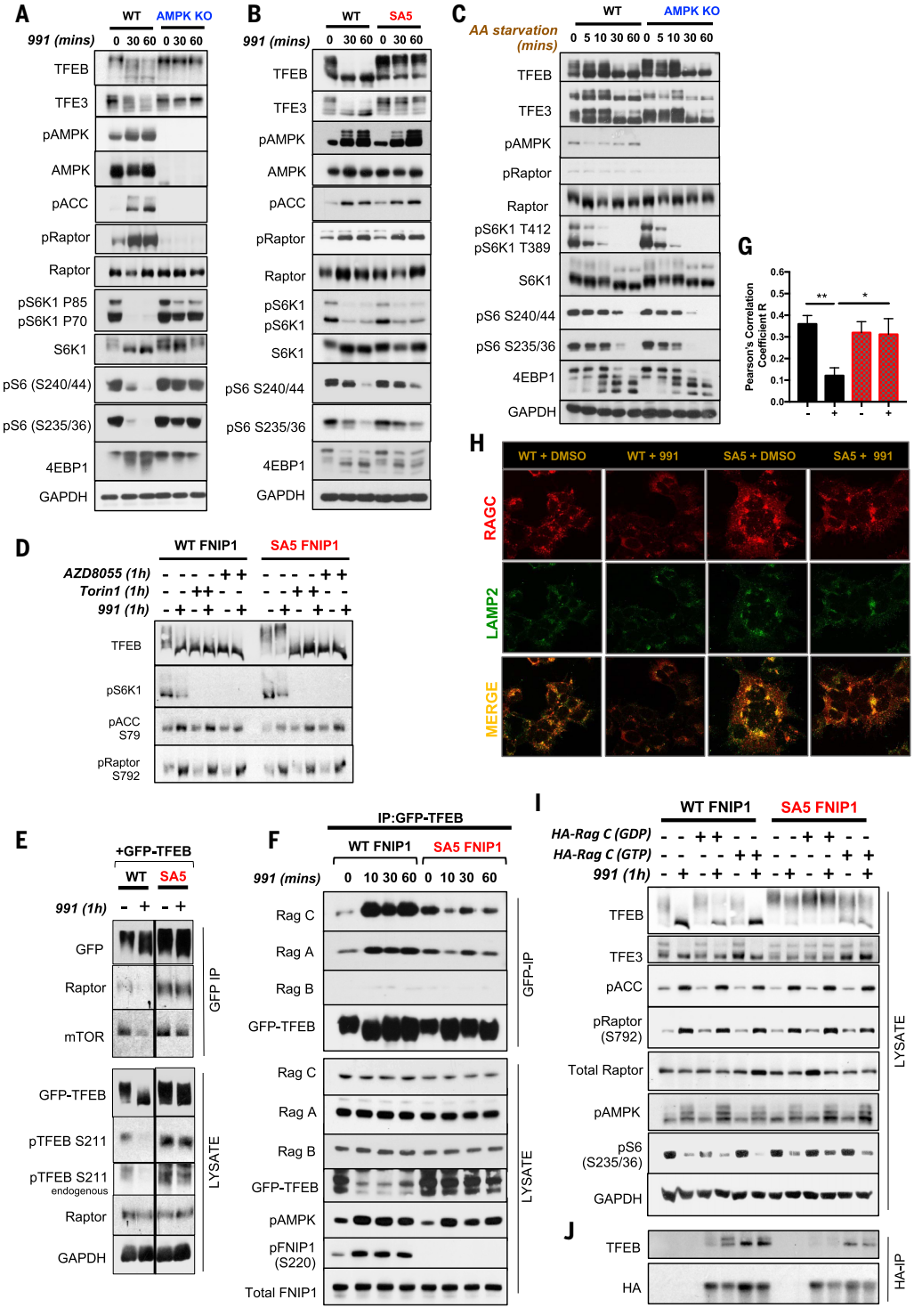


Fig. 3. AMPK phosphorylation of FNIP1 blocks FLCN-FNIP1 GAP activity to control RagC and thereby TFEB and TFE3 activity. (**A**) WT and AMPK KO HEK293T cells were subjected to a short 991 (50 μM) time course, and lysates were immunoblotted with the indicated antibodies to probe for mTORC1 signaling. (**B**) WT FNIP1 and SA5 FNIP1 HEK293T cells were treated as in (A), and lysates were immunoblotted to probe for mTORC1 signaling. (**C**) WT parental or AMPK-null HEK29T cells were AA starved for the indicated times, and lysates were immunoblotted to examine mTORC1 signaling. (**D**) WT FNIP1 or SA5 FNIP1 HEK293T cells were administered with the mTOR inhibitors AZD8055 or Torin1 either individually or in combination with 991, as indicated. Lysates were subsequently immunoblotted to examine TFEB phosphorylation status. (**E**) WT or SA5 FNIP1 cells, stably expressing GFP-TFEB cDNA were treated with or without 50-μM 991 for 1 hour. GFP-TFEB was immunoprecipitated from the lysates, and immunoprecipitates were analyzed by Western blotting. (**F**) Immunoprecipitates of GFP-TFEB, stably expressed in WT FNIP1 and SA5 FNIP1 HEK293T cells, were subjected to immunoblotting to probe interactions with the Rag GTPases. (**G**) Quantitation of RagC colocalization with Lamp2 in (H). Data are shown as the means ± SEMs of three independent experiments. *P < 0.05; **P < 0.01; unpaired t test. (**H**) Representative immunofluorescence images of endogenous RagC costained with Lamp2 in WT FNIP1 and SA5 FNIP1 HEK293T cells treated with 1 hour of DMSO or 991 (50 μM). (**I**) WT FNIP1 or SA5 FNIP1 cells were transiently transfected with HA-RagC mutants locked in either the GTP-bound state (Q120L) or the GDP-bound state (S75N) and subsequently treated with or without 1-hour 991. Lysates were immunoblotted with the indicated antibodies. (**J**) HA-RagC mutants from (I) were immunoprecipitated from lysates using HA magnetic beads and immunoprecipitates analyzed by Western blotting.

SA5-FNIP1 cells, 991 did not have this effect, and mTOR and Raptor remained associated to GFP-TFEB (Fig. 3E). Although mTORC1 signaling to S6K1 and 4EBP1 is decreased by AMPK activation in SA5-FNIP1 cells, it appears that TFEB is constitutively phosphorylated in cells expressing SA5 FNIP1 treated with 991 because of mTORC1 remaining associated with TFEB.

AMPK phosphorylation of FNIP1 inhibits FNIP1-FLCN GAP activity to control TFE transcription factors through RagC

In cells stimulated with AAs, the Rag GTPases recruit mTORC1 to the lysosomal surface, where it is activated. The Rag proteins function as heterodimers, in which the active complex consists of GTP-bound RagA or RagB in complex with guanosine diphosphate (GDP)–bound RagC or RagD (*38*, *39*). Activation of mTORC1 by intracellular AAs occurs because AAs stimulate GTP binding to RagA and RagB, promoting binding to Raptor and assembly of the activated mTORC1 complex (*40*). In the absence of AAs, the Rags take up an inactive conformation (GDP-bound RagA or B and GTP-bound RagC or D), causing inactivation and relocalization of mTORC1 to the cytosol. Concordantly,

active Rag heterodimers interact with and promote recruitment of TFEB to the lysosomes, leading to mTORC1-dependent phosphorylation and retention of TFEB in the cytosol. Depletion or inactivation of Rags prevents recruitment of TFEB to lysosomes (*5*, *7*, *9*). Furthermore, a catalytic arginine in FLCN is required for AA-dependent translocation of TFEB and TFE3, connecting control of FLCN-FNIP1 GAP activity to AA regulation of RagC (*15*). To further delineate how FNIP1 phosphorylation by AMPK controls TFEB, we treated the WT FNIP1 and SA5 FNIP1 cells stably expressing GFP-TFEB with 991 and immunoprecipitated GFP-TFEB. Within 10 min, interaction of endogenous RagA and especially endogenous RagC interaction with GFP-TFEB was enhanced and maintained for the duration of the treatment (Fig. 3F). Conversely in SA5 FNIP1 cells, RagA and RagC interactions with TFEB did not change after treatment with 991 (Fig. 3F). RagB binding to TFEB was minimal and did not change between WT and SA5 FNIP1 cells or with 991 treatment.

To determine how the FLCN-FNIP1–RagA-C complex controls TFEB after AMPK activation, we tested whether components of the TFEB regulatory machinery changed localization at the lysosome in cells exposed to 991. We used Lyso-IP, a method for the rapid isolation of mammalian lysosomes (*41*), which uses expression of lysosomal transmembrane protein 192 fused to three tandem hemagglutinin (HA) tags (HA-TMEM192) in WT FNIP1 and SA5 FNIP1 HEK293T cells (fig. S4B). As controls, we also expressed FLAG-TMEM192 in WT FNIP1 and SA5 FNIP1 HEK293T cells. All cell lines were treated with 991, and HA-TMEM192 was immunoprecipitated. The lysosome surface marker, lysosomal-associated membrane protein 1 (Lamp1), was highly enriched in the immunopurified lysosomal fraction with little remaining in the supernatant, which indicates that most of the lysosomes had been extracted with the Lyso-IP technique (fig. S4B). Organelle markers, such as Golgin97 for Golgi, were not detected in the immunoprecipitates, demonstrating that organelle contamination had not occurred in the lysosome purifications. We observed a rapid decrease in association of endogenous RagC and RagA with the lysosome in WT FNIP1-HA-TMEM192 cells within 20 min of treatment with 991 (fig. S4B). By contrast, both RagC and RagA remained in the lysosome fractions in SA5 FNIP1-HA-TMEM192 cells with or without 991. FLCN, FNIP1, and mTOR followed a similar pattern. Depletion of AAs reduced RagA and RagC in lysosome fractions in WT FNIP1 but also SA5 FNIP1 cells, reiterating that AMPK-FNIP1–mediated control of mTOR and TFEB is not required in these conditions and is distinct from TFEB regulation by AAs and mTOR (fig. S4C).

To visualize the changes in RagC and mTOR localization after AMPK activation in cells, we used immunofluorescence imaging of endogenous RagC or mTOR, with antibodies to RagC or mTOR. RagC localized with the lysosomal marker Lamp2 in WT FNIP1 cells in fresh medium (Fig. 3, G and H). If AMPK was activated with 991, RagC localization with Lamp2—measured by Pearson's correlation—was one-third of that in WT FNIP1 cells but remained localized with Lamp2 in SA5 FNIP1 cells with or without 991 (Fig. 3, G and H), confirming our biochemical observations. mTOR similarly showed decreased lysosomal localization in WT FNIP1 cells treated with 991 but not in cells expressing SA5 FNIP1 (fig. S4, D and E). Thus, AMPK activation appears to displace RagC and mTOR from lysosomes, preventing TFEB phosphorylation (fig. S4F). The enhanced interaction of RagC and TFEB in cells in which AMPK is activated with 991 indicates that RagC is not only required for recruitment of TFEB to the lysosome but also that increased binding to RagC may promote TFEB removal from the lysosome, perhaps by physically chaperoning TFEB into the cytosol.

We explored how FNIP1 phosphorylation by AMPK triggers the disassembly of the lysosomal machinery controlling TFEB. The FLCN-FNIP1 complex functions as a GAP for RagC and RagD, promoting the GDP-bound state of RagC and RagD, which is required for mTOR recruitment to the lysosome (*11*–*13*). Thus, phosphorylation of FNIP1 by AMPK might alter the GAP activity of the FLCN-FNIP1 complex. To test this, we transiently transfected either an HA-tagged, GTP-locked mutant of RagC (Q120L) or an HA-tagged, GDP-locked RagC mutant (S75N) in WT FNIP1 and SA5 FNIP1 or WT and AMPK KO cells and treated them with 991 (Fig. 3I and fig. S4G). If phosphorylation of FNIP1 by AMPK blocks the GAP activity of the FLCN-FNIP1 complex toward RagC, then GDP-loaded RagC should accumulate in cells lacking AMPK or SA5 cells. One prediction of this model is that overexpression of the GTP-locked mutant of RagC (Q120L) but not the GDP-locked RagC mutant (S75N) should restore TFEB and TFE3 dephosphorylation in AMPK KO cells and SA5 FNIP1 cells. Overexpression of GTP-locked RagC overrode the effect of SA5 FNIP1, allowing TFEB and TFE3 to be dephosphorylated in SA5 cells with or without 991 (Fig. 3I). We observed the opposite effect in WT FNIP1 cells, in which less TFEB dephosphorylation occurred after treatment with 991, when the GDP-locked RagC mutant was overexpressed compared with when GTP-locked RagC was present (Fig. 3I). Similar results were observed in the AMPK KO cells, where GTP-locked RagC overcame the absence of AMPK, enabling dephosphorylation of TFEB, whereas GDP-locked RagC prevented complete TFEB dephosphorylation in WT cells treated with 991 (fig. S4G).

Thus, AMPK phosphorylation of FNIP1 appears to inhibit FLCN-FNIP1 GAP activity, driving RagC to accumulate in its inactive GTP-bound form, which not only falls off the lysosome (Fig. 3, G and H, and fig. S4B) but also binds more tightly to TFEB (Fig. 3F). Immunoprecipitation experiments using cells expressing GTP- or GDP-locked RagC mutants, under the same conditions as shown in Fig. 3I, showed that in WT FNIP1 cells, more TFEB was bound to RagC in the GTP-locked state, bypassing the need for AMPK activation, in contrast to the GDP-locked state in which binding only occurred if cells were treated with 991 (presumably because of the presence of endogenous RagC, which would also bind TFEB in 991-treated cells) (Fig. 3J). Furthermore, in SA5 FNIP1 cells, exogenous GTP-locked RagC was associated with TFEB, overriding the effect of the SA5 mutations, whereas GDP-locked RagC did not bind TFEB to the same extent, regardless of whether cells were treated with 991 (Fig. 3J). A similar pattern was observed in the WT cells compared with cells lacking AMPK (fig. S4H). These results further corroborate and explain the results shown in Fig. 3F: Increased HA-RagC-TFEB interaction occurs in cells with active AMPK because of inactivation of the FLCN-FNIP1 GAP complex, propelling RagC to its GTP-bound state, which binds to TFEB more strongly than the GDP-bound form of RagC. Although GDP-RagC recruits TFEB to the lysosome for mTOR-dependent phosphorylation, GTP-RagC appears to be required to chaperone TFEB off the lysosome, preventing its phosphorylation.

FNIP1 phosphorylation by AMPK is required for lysosomal biogenesis

The MiT-TFE family of transcription factors, including TFEB and TFE3, are oncogenes and master regulators of lysosome biogenesis and autophagy (*16*, *17*). Given that AMPK activation leads to nuclear translocation of TFEB and TFE3, an increase in their transcriptional activity would be expected. Our RNA-seq data in WT and AMPK KO HEK293T cells displayed an AMPK-dependent lysosomal gene signature. To assess the functional role of FNIP1 phosphorylation by AMPK, we analyzed global transcription in WT and SA5 FNIP1 cells treated with 991 for 0 to 16 hours by RNA-seq. Unsupervised hierarchical clustering of differentially expressed genes ($P \leq 0.05$; FC ≥ 1.3) revealed that genes clustered according to 991 treatment or FNIP1 mutation status (WT or SA5) or both, demonstrating that FNIP1 phosphorylation by AMPK governed ~20% of the AMPK-responsive genes, whose transcription increased in cells treated with 991 (fig. S5, A and B). Enrichment analysis of differentially expressed genes found "clathrin-coated endocytic vesicle membrane" and "autolysosome" to be the most overrepresented terms in the GO cellular components category for transcripts less abundant in SA5 FNIP1 cells than in WT FNIP1

cells treated with 991 (fig. S5, C and D); GSEA analysis revealed "KEGG Lysosome" as one of the most enriched gene sets in WT FNIP1 cells treated with 991 for 16 hours compared with SA5 FNIP1 cells with the same treatment (Fig. 4A).

To more deeply examine how AMPK and FNIP1 control lysosome biology, we manually curated a list of ~1500 genes, which included experimentally validated CLEAR targets (*42*) in addition to the Molecular Signatures Database (MsigDB)–GSEA defined TFEB targets.

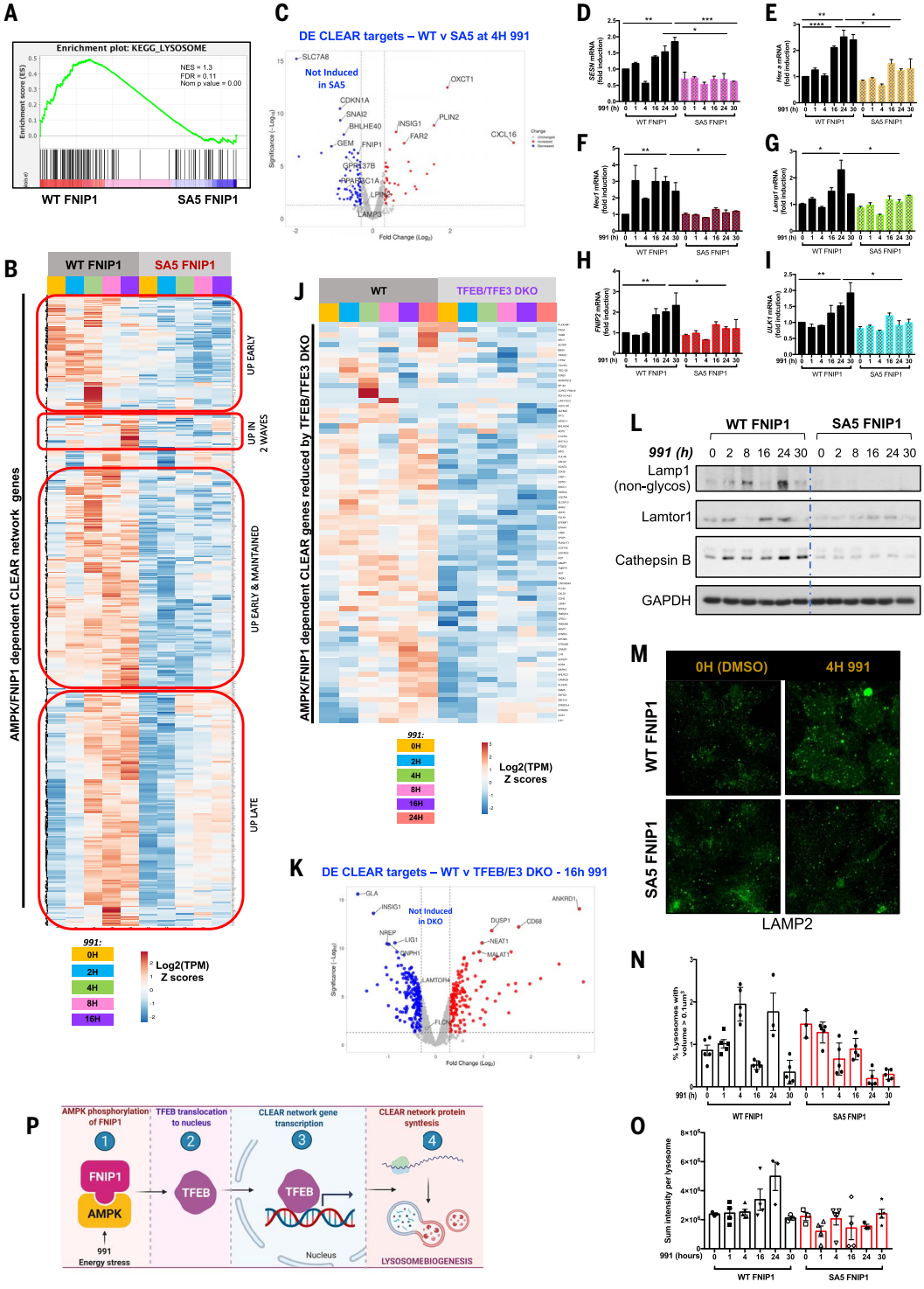


Fig. 4. Lysosomal biogenesis mediated by the MiT-TFE family of transcription factors is dependent on AMPK phosphorylation of FNIP1. (**A**) GSEA plot for the "KEGG Lysosome" gene set, which was enriched in WT FNIP1 16-hour 991-treated but not SA5 conditions. (**B**) RNA-seq analysis of WT FNIP1 and SA5 FNIP1 cells subjected to a 0- to 16-hour 991 (50 μM) time course. Clustering analysis and heatmap displays expression patterns of AMPK-FNIP1–dependent CLEAR network genes that have been previously validated or GSEA defined. (**C**) Volcano plot depicting differential expression of CLEAR network genes after 4-hour 991 in WT FNIP1 versus SA5 FNIP1 conditions. Blue dots represent genes significantly down-regulated by mutation of AMPK sites on FNIP1. The *y* axis denotes $-\log_{10} P$ values, and the *x* axis shows $\log_2$ FC values. (**D** to **I**) qRT-PCR of CLEAR network genes SESN (C), Hex A (D), Neu1 (E), Lamp1 (F), FNIP2 (G), and ULK1 (H) in WT FNIP1 and SA5 FNIP1 HEK293T cells subjected to a 0- to 30-hour 991 (50 μM) time course. Graphs are shown as means ± SEMs. $n = 3$. $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; unpaired *t* test. (**J**) RNA-seq analysis of WT parental and CRISPR-Cas9–mediated TFEB-TFE3 DKO HEK293T cells treated with a 0- to 24-hour 991 time course. Heatmap shows AMPK-FNIP1–dependent genes whose expression is reduced by loss of TFEB-TFE3. (**K**) Volcano plot denoting CLEAR network DE genes after 16 hours of 991 in WT versus TFEB-TFE3 DKO FNIP1 cells. Blue dots represent genes significantly reduced by deletion of TFEB-TFE3. The *y* axis denotes $-\log_{10} P$ values, and the *x* axis shows $\log_2$ FC values. (**L**) Immunoblotting of lysosomal proteins in WT FNIP1 and SA5 FNIP1 HEK 293 cells after a 0- to 30-hour 991 time course. (**M**) Representative immunofluorescence images of lysosome structures stained with Lamp2 antibody after DMSO or 4-hour 991 treatment of WT FNIP1 and SA5 FNIP1 cells. (**N**) Quantitation of Lamp2 lysosomal structures from (M) and at the time points indicated, showing percentage of lysosome structures with volume greater than 0.1 μm^3 after a 0- to 30-hour 991 time course in WT FNIP1 and SA5 FNIP1 cells. (**O**) Quantitation of Lamp2 sum intensity per lysosome in (M) and other time points from the same experiment. (**P**) Model. AMPK phosphorylation of FNIP1, induced by 991 or energetic stress, triggers TFEB entry into the nucleus, where it binds to CLEAR elements on lysosomal gene promoters, inducing lysosomal gene transcription, enhancing lysosomal protein expression and thereby lysosome biogenesis.

Gene expression pattern analysis showed that ~75% of these CLEAR target genes showed increased transcription after AMPK activation in WT FNIP1 conditions (Fig. 4B and fig. S5, E and F). The genes clustered into four main groups (Fig. 4B)—one group showed early transcriptional activation between 2 and 4 hours only, the second group showed two waves of transcription with one wave at an early time point (2 hours) and the next at a later time point (16 hours), the third group was activated early from 2 hours onward but remained steadily high up to 16 hours, and the fourth group responded mainly at later time points between 8 and 16 hours. A large proportion of these genes, whose expression was increased by 991 in WT FNIP1 cells, did not respond to AMPK activation in cells overexpressing SA5 FNIP1 (~600 genes) (Fig. 4, B and C, and fig. S5C).

To validate some of the targets from our RNA-seq analysis, we subjected another batch of WT FNIP1 and SA5 FNIP1 cells to prolonged treatment with 991 for 0 to 30 hours and performed qPCR with primers targeting several canonical CLEAR network members. Expression patterns were similar to those in the RNA-seq data. A rapid increase in expression of *SESN*, *Hex A*, *Neu1*, *Lamp1*, *FNIP2*, and *ULK1* mRNA, ranging from ~1.2- to 6-fold, was detected within 2 to 16 hours of 991 administration in WT FNIP1 cells (Fig. 4, D to I). With several lysosomal CLEAR target genes, such as *Neu1*, *SESN*, *HEXA*, and *Lamp1*, two separate waves of transcription were observed in WT FNIP1 cells—one at early time points between 1 and 2 hours and a second larger wave occurring at 16 to 24 hours. However, when AMPK phosphosites on FNIP1 were mutated, no increased transcription of the CLEAR genes was detected (Fig. 4, D to I). Glucose deprivation showed similar effects with increased transcription of the CLEAR network gene *GLA* in WT FNIP1 but not in SA5 FNIP1 cells (fig. S5G).

To demonstrate that the differential expression of lysosomal genes between WT and SA5 FNIP1 cells is regulated by TFEB and TFE3, we used CRISPR methods to make HEK293T cells lacking both TFEB and TFE3 (fig. S5, H and I). A cell line lacking only TFEB had minimal changes in gene expression, suggesting redundancy between TFEB and TFE3 (*11*, *43*, *44*). We treated parental WT and cells lacking both TFEB and TFE3 [TFEB-TFE3 double knockout (DKO)] with 991 for 24 hours and performed RNA-seq analysis. Specifically focusing on our manually curated list of lysosomal CLEAR genes, this analysis revealed loss of expression in a subset of AMPK-FNIP1–dependent lysosomal genes in the TFEB-TFE3 DKO condition compared with the control, including *GLA* and *LAMTOR4* (Fig. 4, J and K).

To further study the effect of AMPK and FNIP1 on lysosomal proteins, we examined protein expression of lysosomal components including Lamp1, Lamtor 1, and cathepsin B (*45*). After 991 treatment, a strong correlation in the amounts of nascent, nonglycosylated Lamp1 protein reflected the expression levels of Lamp1 mRNA, showing two separate waves of increased expression for both mRNA and protein in WT FNIP1 cells (Fig. 4L). By contrast, no change in abundance of the Lamp1 protein was detected in the cells overexpressing SA5 FNIP1, which had low amounts like those in WT FNIP cells before AMPK activation; Lamtor1 and cathepsin B followed similar expression patterns to that of Lamp1 (Fig. 4L).

We used immunofluorescence imaging to analyze lysosomal structures by staining for intracellular Lamp2. Lysosomal structures were quantitated by measurement of lysosome volume and Lamp2 fluorescence intensities. The percentage of lysosomal structures above the threshold volume of 0.1 μm^3 in WT FNIP1 cells treated with 991 displayed the same biphasic up-regulation pattern observed for Lamp1 mRNA and protein (Fig. 4, G and L to N). In 991-treated SA5 FNIP1 cells, the percentage of lysosomes varied but did not increase beyond the starting time point (Fig. 4, M and N). We also quantitated the Lamp2 sum intensity per lysosome, which did not display a biphasic pattern but did increase at 16 and 24 hours of 991 treatment in WT FNIP1 but not SA5 FNIP1 cells (Fig. 4O). Thus, FNIP1 phosphorylation by AMPK appears to contribute to lysosomal biogenesis (Fig. 4P).

AMPK regulation of PGC1α gene induction is governed by FNIP1 phosphorylation

Phenformin, rotenone, and CCCP all induced expression of PGC1α (*PPARGC1A*) mRNA in WT cells but not cells lacking AMPK in our RNA-seq datasets. The RNA-seq analysis also showed 991-induced transcription of *PPARGC1A* mRNA in WT FNIP1 but not SA5 FNIP1 cells (Fig. 4C). The *PPARGC1A* proximal promoter is a direct target of TFEB and TFE3 (*46*–*48*). Therefore, FNIP1 phosphorylation by AMPK, through control of TFEB-TFE3, might be one of the long-sought mechanisms underpinning the transcriptional regulation of mitochondrial biogenesis by AMPK.

The *PPARGC1A* gene undergoes extensive alternative splicing (*49*). qPCR assessment of total *PPARGC1A* using primers against exon 2, which is found in all PGC1α isoforms, showed an increase in transcription in WT FNIP1 but not SA5 FNIP1 cells treated with 991 (Total *PPARGC1A*, Fig. 5A). In humans, PGC1α transcription has been reported from three distinct promoters: a proximal promoter located just upstream of the canonical exon 1a; a distal alternate promoter followed by an alternative exon 1b, located ~13.7 kb upstream from exon 1a (*50*); or a much further upstream promoter termed the brain-specific promoter, ~500 kb upstream of the canonical proximal promoter (*51*, *52*). *PPARGC1A* contains two potential CLEAR elements to direct binding of TFEB and TFE3 in the proximal promoter adjacent to the canonical exon 1a (*46*, *48*). In addition to distinct promoters, alternative splicing between exons 6 and 7 of the *PPARGC1A* gene produces a transcript encoding the N-terminal isoform of PGC1α, which contains 267 AAs of classical PGC1α and three AAs from the splicing insert (*53*) (Fig. 5B). NT-PGC1α is a constitutive transcriptional coactivator because it retains the transcription activation and nuclear receptor interaction domains of full-length PGC1α but is no longer subject to phosphorylation-mediated turnover associated with the full-length protein (*53*, *54*), thus leading to a truncated and transcriptionally active form of PGC1α with a longer protein half-life.

To further investigate regulation of *PPARGC1A* in 991-treated cells, we examined endogenous PGC1α protein expression by Western blotting (Fig. 5C). We did not detect the canonical, full-length ~100-kDa isoform of PGC1α in 991-treated WT or SA5 FNIP1 cells. We did, however, detect a time-dependent increase in the expression of the much smaller ~35-kDa N-terminal isoform of the PGC1α protein in WT FNIP1 cells after 991 administration, under the same conditions in which FNIP1 is phosphorylated and TFEB and TFE3 are activated (Fig. 5C). These changes in the 35-kDa PGC1α protein were quantitated by densitometry, which showed 8- to 16-fold increases in PGC1α expression after 991 treatment in WT FNIP1 cells (Fig. 5D). Abundance of the shorter PGC1α protein remained relatively low in SA5 FNIP1 cells (Fig. 5, C and D). We detected similar changes using two commercial antibodies (Santa Cruz and EMD Millipore), both of which have been validated for detection of NT-PGC1α isoforms (*53*, *55*).

Because *PPARGC1A* is extensively alternatively spliced, we sought to further examine which *PPARGC1A* mRNA splice isoforms might accumulate in an AMPK-FNIP1–dependent manner. Analysis of the specific *PPARGC1A* mRNA splice isoforms in our RNA-seq data in WT and AMPK KO cells treated with mitochondrial poisons revealed accumulation of the *NT-PPARGC1A* short isoform (ENST no. 506055.5) that encodes the 271-AA truncated form but not full-length PGC1α isoforms (full-length, ENST no. 264867.7) after these stresses (fig. S6A). To directly compare with our initial observations of total *PPARGC1A* mRNA accumulation (targeting common exon 2, present in all splice isoforms), we carried out qPCR with primers targeting exons 5 and 7a to specifically detect expression of the *NT-PPARGC1A* alternative splice form (Fig. 5A). *NT-PPARGC1A* appeared to be more abundant in WT FNIP1 cells treated with 991 than total *PPARGC1A* mRNA; no effect was detected in cells expressing SA5 FNIP1 (Fig. 5A). qPCR from CCCP-treated WT and AMPK KO cells showed a similar pattern to what was observed with 991

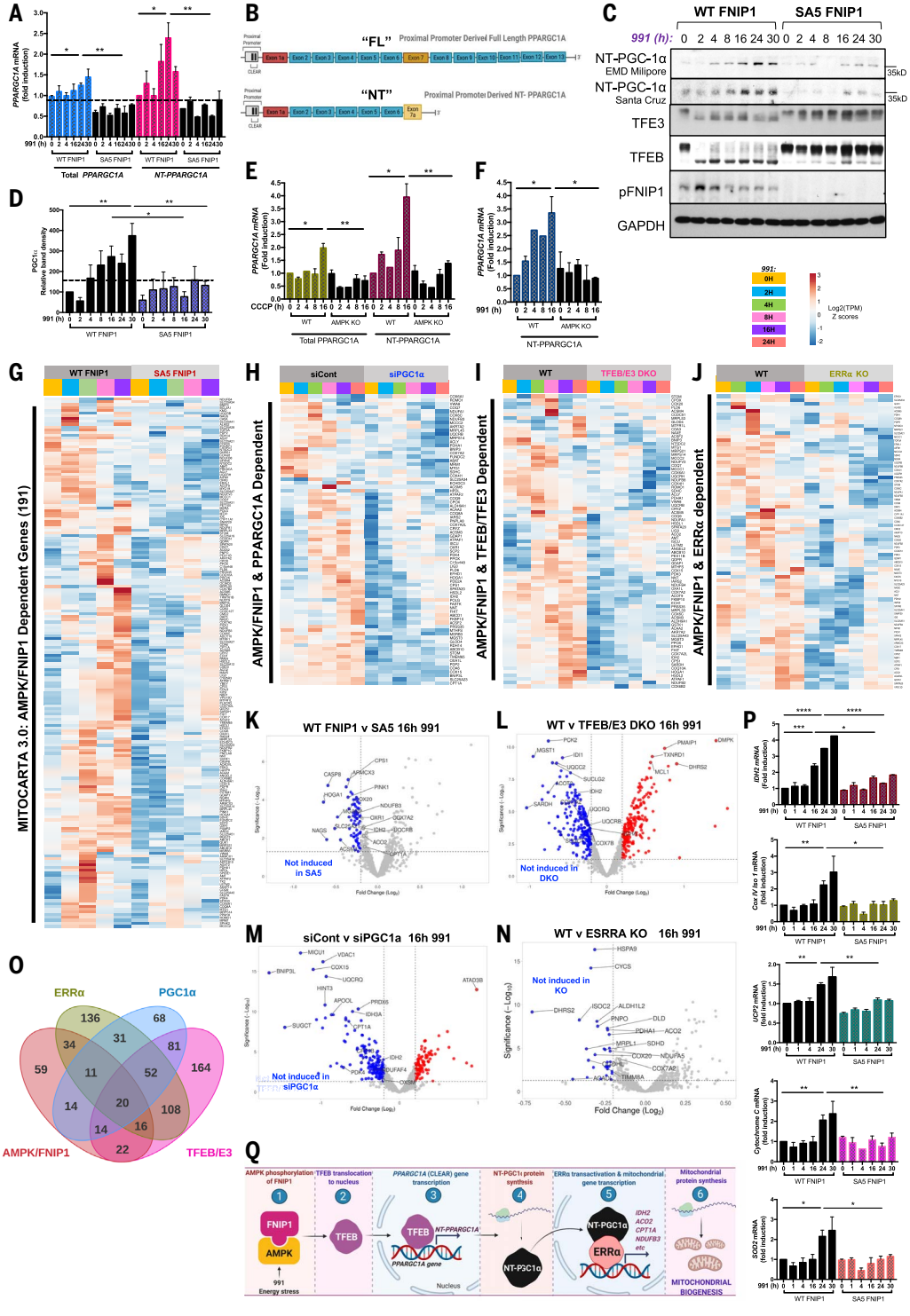


Fig. 5. FNIP1 phosphorylation by AMPK is critical for induction of the PGC1α- and ERRα-mediated mitochondrial biogenesis program through MiT-TFE transcription factors. (**A**) qRT-PCR showing expression of total *PPARGC1A* compared with expression of the shorter *NT-PPARGC1A* splice isoform. Graphs are shown as means ± SEMs. $n = 3$. $^{*}P < 0.05$; $^{**}P < 0.01$; unpaired t test. (**B**) Schematic of the two predominant splice isoforms of *PPARGC1A* induced in this cell type in our conditions. (**C**) Immunoblots reflecting changes in expression of NT-PGC1α protein levels in WT FNIP1 versus SA5 FNIP1 cells in the presence or absence of a 991 time course. Molecular weights are indicated on the right. (**D**) Densitometry analysis of NT-PGC1α immunoblots. (**E**) qRT-PCR comparing expression of total *PPARGC1A* compared with that of shorter *NT-PPARGC1A* splice isoform in WT and AMPK KO HEK293T after a 0- to 16-hour CCCP time course. (**F**) qRT-PCR comparing expression of *NT-PPARGC1A* splice isoform in WT and AMPK KO HEK293T after a 0- to 16-hour 991 time course. For (D) to (F), all values are shown as means ± SEMs. $n = 3$. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; unpaired t test. (**G**) Gene clustering analysis of RNA-seq data from WT FNIP1 and SA5 FNIP1 cells subjected to 991 time course treatments. Heatmap displays expression pattern of AMPK- and FNIP1-dependent mitochondria specific genes, as defined by the Mitocarta 3.0 inventory. (**H**) RNA-seq analysis of HEK293T cells transfected with control siRNA (siCont) or siRNA targeting PGC1α (siPGC1α) and treated with 0 to 24 hours of 991, as indicated. Heatmap displays AMPK-FNIP1–dependent Mitocarta 3.0 genes whose expression is lost with PGC1α knockdown. (**I**) RNA-seq analysis of WT and TFEB-TFE3 DKO HEK293T cells treated with 0 to 24 hours of 991. Heatmap displays expression pattern of AMPK-FNIP1–dependent mitochondrial genes, as defined by Mitocarta 3.0, that are lost upon TFEB-TFE3 deletion. (**J**) RNA-seq analysis of WT versus ERRα KO HEK293T cells subjected to a 991 time course, as indicated. Heatmap visualizes the expression pattern of AMPK-FNIP1–dependent mitochondrial genes in WT and ERRα KO HEK293T cells. (**K**) Volcano plot displaying Mitocarta 3.0 DE genes after 16-hour 991 in WT versus SA5 FNIP1 cells. Blue dots represent genes significantly down-regulated by mutation of AMPK sites on FNIP1. (**L**) Volcano plot depicting DE Mitocarta 3.0 genes in 16-hour 991-treated WT versus 16-hour 991-treated TFEB-TFE3 DKO RNA-seq dataset. Blue dots represent genes significantly down-regulated by CRISPR deletion of TFEB-TFE3. (**M**) Volcano plot depicting DE Mitocarta 3.0 genes in 16-hour 991 siCont versus 16-hour 991 siPGC1α RNA-seq dataset. Blue dots represent genes significantly down-regulated by knockdown of PGC1α. (**N**) RNA-seq analysis of WT versus ERRα KO HEK293T cells. Volcano plot displaying DE Mitocarta 3.0 genes after 16-hour 991 in WT versus ERRα null cells. Blue dots represent genes significantly down-regulated by deletion of ERRα. (**O**) Four-way Venn diagram showing overlap of gene sets controlled by AMPK-FNIP1, TFEB-TFE3, PGC1α, and ERRα. (**P**) qRT-PCR of mitochondrial genes including IDH2, Cox IV, CytoC, UCP2, and SOD2 in WT FNIP1 and SA5 FNIP1 HEK293T cells subjected to a 0- to 30-hour 991 time course. All data are shown as means ± SEMs. $n = 3$. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$; unpaired t test. (**Q**) Model. AMPK phosphorylation of FNIP1 after energy stress or 991 facilitates TFEB nuclear entry where it binds to CLEAR network gene promoters, including the *PPARGC1A* promoter, which induces expression of the short ~35-kDa transcriptional coactivator NT-PGC1α isoform. In turn NT-PGC1α transactivates the ERRα transcription factor for induction of mitochondrial genes.

treatment in the WT FNIP1 and SA5 FNIP1 cells. After treatment with CCCP, the *NT-PPARGC1A* short isoform accumulated in WT cells but not cells lacking AMPK (Fig. 5E). Similar results were also seen with the *NT-PPARGC1A* short isoform after 991 treatment in WT cells but not in cells lacking AMPK (Fig. 5F). Glucose deprivation also induced abundance of *NT-PPARGC1A* mRNA in WT FNIP1 but not SA5 FNIP1 cells (fig. S6B). Thus, the NT-PGC1α isoform is the predominant splice isoform of PGC1α expressed in HEK293T cells in response to mitochondrial poisons, glucose starvation, or 991 activation of AMPK.

To discern whether FNIP1 is a critical link between AMPK and PGC1α to promote mitochondrial biogenesis, we performed hierarchical clustering and differential expression analyses of our RNA-seq datasets. Again, using the Mitocarta 3.0 catalog, we specifically examined mitochondrial gene expression in WT and SA5 cells treated with 991. This revealed a subset of ~200 mitochondrial genes whose mRNA expression was increased by 991 at time points ranging from 2 to 16 hours, although most of these transcripts accumulated at the longest time point studied (16 hours) (Fig. 5G and fig. S6C). Notably, these transcripts did not accumulate when AMPK phosphosites in FNIP1 were mutated in the SA5 FNIP1 samples, even if cells were treated with 991 (Fig. 5, G and K). To investigate whether phosphorylation of FNIP1 by AMPK triggers mitochondrial biogenesis through PGC1α and to define the mitochondrial gene signature regulated by AMPK and FNIP1 in a PGC1α-dependent manner, we used small interfering RNA (siRNA) to deplete HEK293T cells of PGC1α (90 to 95% efficiency) (fig. S6D) and performed RNA-seq after treating cells with 991. PGC1α depletion led to decreased expression of 291 mitochondrial genes from Mitocarta 3.0 (fig. S6E); a subset of ~80 of these genes was dependent on AMPK phosphorylation of FNIP1 (Fig. 5, H and M). To test whether changes in mitochondrial gene expression were mediated by AMPK-FNIP1 control of TFEB, we also analyzed the expression of Mitocarta 3.0 genes in our WT and TFEB-TFE3 DKO RNA-seq dataset. We specifically examined the AMPK- and FNIP1-dependent mitochondrial genes and made an observation similar to that observed with the CLEAR network genes: Expression of ~50% of AMPK-FNIP1–dependent mitochondrial genes increased by 991 treatment in WT but not TFEB-TFE3 DKO samples (Fig. 5, I and L, and fig. S6F).

We investigated mitochondria gene expression in cells expressing WT FNIP1 or SA5 FNIP1 after treatment with 991, validating transcription of several PGC1α mitochondrial targets by qPCR. mRNA levels of *IDH2*, *Cox IV*, *Cyto C*, *SOD2*, and *UCP2* remained low at early time points but accumulated between 16 and 24 hours after 991 administration (Fig. 5P). Cells deprived of glucose for 6 hours also showed increased transcription of reported PGC1α target genes such as *PDHA1* in WT FNIP1 cells (fig. S6G). In cells expressing SA5 FNIP1, transcription of all mitochondrial genes tested did not increase despite prolonged stimulation of cells with 991 or glucose deprivation (Fig. 5P and fig. S6G). Although many lysosomal genes showed two waves of expression, mitochondrial gene expression generally occurred at later time points, consistent with regulation by PGC1α, mRNA of which was itself transcribed in sync with the first wave of lysosomal biogenesis (model in Fig. 5Q).

AMPK-FNIP1 induction of mitochondrial biogenesis requires ERRα

Our RNA-seq analysis detected an AMPK-dependent increase in transcripts for the nuclear receptors *ESRRA* and *ESSRG*, an effect that was abolished by deleting AMPK or *PPARGC1A* siRNA knockdown. In WT cells, mRNAs of genes encoding nuclear receptors, including *ESRRA* and *ESRRG*, were increased from 2 to 16 hours after AMPK activation, and depletion of *PPARGC1A* suppressed this effect (fig. S7A). PGC1α initiates mitochondrial biogenesis through interaction with ERRα (*30*, *56*, *57*). Transcription of the ERRα gene is also increased by PGC1α (*29*). The requirement of PGC1α for 991 to promote *ESRRA* expression (fig. S7A) suggests that at least some of the AMPK- and FNIP1-dependent mitochondrial biogenesis program mediated by PGC1α might be through transcriptional coactivation of ERRα. Furthermore, transcription factor enrichment analysis of the global transcriptome of WT cells and those lacking AMPK revealed that *ESRRA* targets were overrepresented upon 991 treatment in the WT condition and underrepresented in cells lacking AMPK (fig. S2J). To test this, we genetically deleted *ESRRA* by CRISPR (fig. S7C). RNA-seq analysis of 991-treated WT cells or cells lacking ERRα showed that many Mitocarta 3.0 genes required ERRα for basal expression or increased expression in response to 991 (fig. S7D). We focused on the AMPK- and FNIP1-dependent Mitocarta genes to assess which of these were up-regulated by 991 in WT controls but suppressed when ERRα was absent (Fig. 5, J and N). We validated the expression of several of these genes by qPCR, including *CPT1A*, *COX6A1*, *IDH2*, and *PDHA1* (fig. S7F), and observed increased transcription in cells treated with 991, with maximal increase at later times of 16 to 24 hours. 991 did not induce transcription of these genes in cells lacking ERRα.

We next performed a four-way comparison of the 991-regulated Mitocarta genes that varied between (i) WT and SA5 FNIP1 cells, (ii) WT and TFEB-TFE3 DKO cells, (iii) control and PGC1α siRNA cells, and (iv) WT and cells lacking ERRα (Fig. 5O). We observed an overlap between all groups, with 131 out of 190 AMPK- and FNIP1-dependent genes being coregulated by one or multiple of the other factors. Twenty Mitocarta genes were regulated by all four factors, and 61 genes were regulated by three of the four conditions (in an SA5-FNIP1–dependent manner), which we propose as a core minimal AMPK-FNIP1-TFEB-PGC1α-ERRα–dependent gene set involved in mitochondrial biogenesis (Fig. 5O and fig. S7E).

AMPK-FNIP1 is required for mitochondrial and lysosomal biogenesis

We performed immunoblotting of mitochondrial proteins as a measure of mitochondrial biogenesis. We detected increased expression of key mitochondrial proteins, including isocitrate dehydrogenase type 2 (IDH2), ubiquinol-cytochrome C reductase binding protein (UQCRB), and NADH:ubiquinone oxidoreductase subunit B3 (NDUFB3), after 991 treatment in WT FNIP1 cells but not cells expressing SA5 FNIP1 (Fig. 6A); these were all detected in our RNA-seq data showing similar expression patterns. In agreement, expression of mitochondrial proteins, such as UQCRB, PDHA1, and NDUFB3, was increased by 991 treatment of WT but not TFEB-TFE3 DKO cells, indicating that these proteins also require TFEB and TFE3 (Fig. 6B). Finally, we also immunoblotted for these mitochondrial proteins after siRNA knockdown of PGC1α or CRISPR deletion of ERRα, observing that PGC1α (fig. S7B) and ERRα (Fig. 6C) were both required for these mitochondrial proteins to accumulate in 991-treated cells.

Although most of the genes required for mitochondrial biogenesis are encoded by the nucleus, the mitochondrial genome encodes 13 proteins, which are components of the OXPHOS system. As another parameter to measure mitochondrial biogenesis, we quantitated relative mitochondrial DNA (mtDNA) copy number by qPCR (*58*) (i.e., the ratio between a mitochondrial gene to a reference nuclear gene) in our cell lines at the 24-hour time point when mitochondrial gene transcription and protein expression was maximal. In cells treated with 991 for 24 hours, we observed an ~30% increase in mtDNA abundance in WT FNIP1 cells but not in SA5 FNIP1 cells (Fig. 6D).

To visualize mitochondria in WT and SA5 FNIP1 cells, we performed Airyscan microscopy after staining cells with IDH2 or CoxIV antibodies. Exposure of cells to 991 for 24 hours increased endogenous IDH2 staining in mitochondria in WT but not SA5 FNIP1 cells (Fig. 6, E and F). Again using Airyscan microscopy, we observed an increase in both mitochondrial (CoxIV) and lysosomal (LAMP2) volumes after 24 hours of 991 treatment in cells expressing WT but not SA5 FNIP1 (Fig. 6, G to I, and fig. S8A). We also observed an apparent increase in mitochondria-lysosome colocalization in 991-treated WT FNIP1 cells (Fig. 6G and fig. S8B), which we further examined using serial section

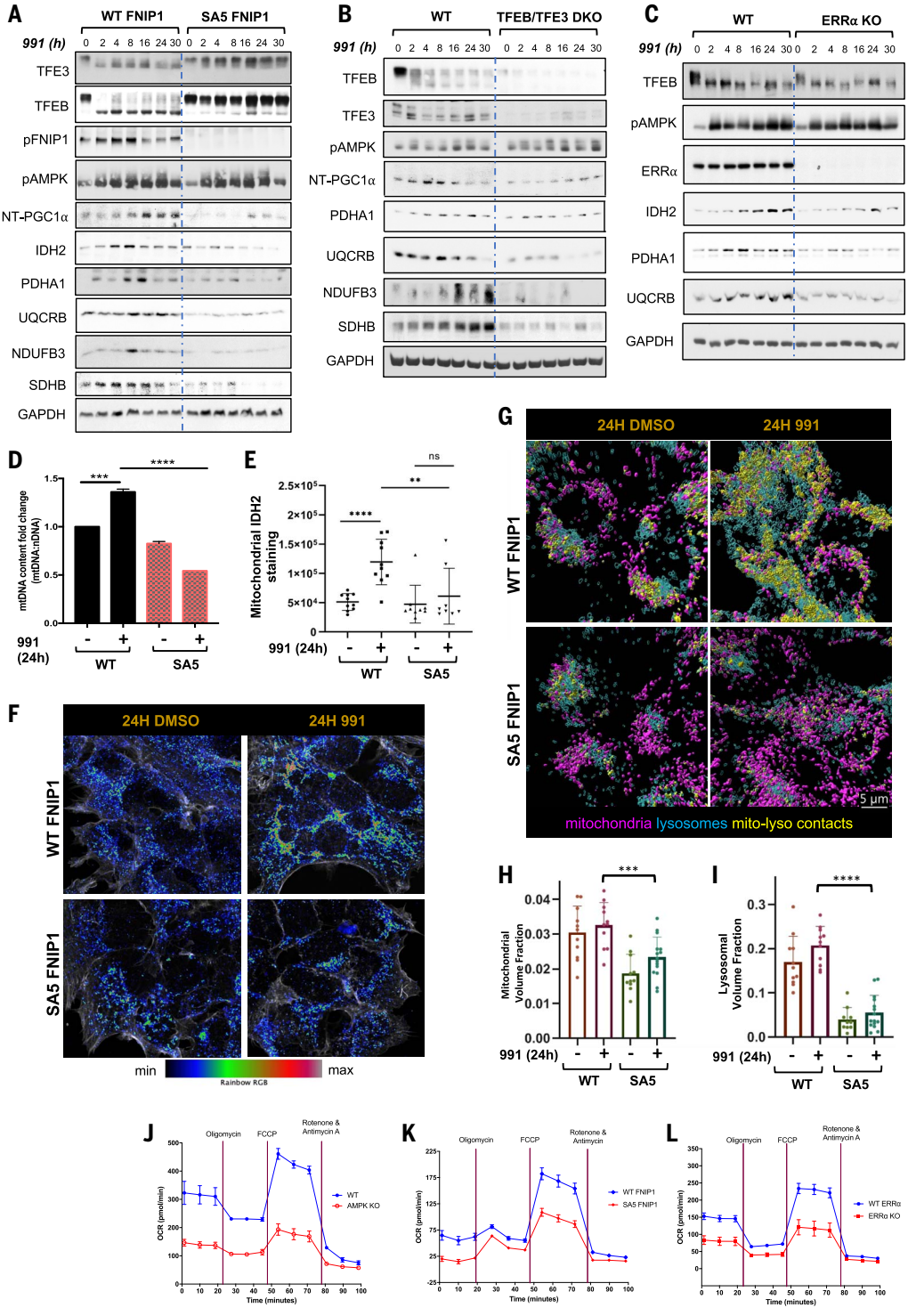


Fig. 6. AMPK-FNIP1–mediated mitochondrial biogenesis affects mitochondrial function and behavior. (**A** to **C**) Western blots probing mitochondrial protein expression after a 991 time course ranging from 0 to 30 hours in WT FNIP1 and SA5 FNIP1 HEK293T cells (A), WT and TFEB-TFE3 DKO HEK293T cells (B), and WT and ERRα KO HEK293T cells (C). (**D**) Mitochondrial DNA content analysis. The ratio of mitochondrial (16*S*) to nuclear (actin) DNA was determined by qRT-PCR after treatment for 24 hours with 991 or DMSO (vehicle), as indicated. (**E**) Quantitation of IDH2 staining in (F). (**F**) Representative Airyscan microscopy images of mitochondrial IDH2 staining in WT FNIP1 and SA5 FNIP1 HEK293T cells treated for 24 hours with 991 or DMSO, as indicated. (**G**) Representative Airyscan images of Lamp2-stained lysosomes and Cox IV–stained mitochondria in WT FNIP1 and SA5 FNIP1 HEK293T cells treated for 24 hours with DMSO or 991, as indicated. (**H**) Quantitation of mitochondrial volume in (G). (**I**) Quantitation of lysosomal volume in (G). (**J**) Seahorse assay to measure OCR in WT compared with AMPK KO HEK293T cells. (**K**) Seahorse assays displaying OCR in WT FNIP1 compared with SA5 FNIP1 HEK293T cells. (**L**) Seahorse assays measuring OCR in WT compared with ERRα KO HEK293T cells. Graphs are shown as the means ± SEMs. $n = 3$. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$; unpaired *t* test.

EM. In EM data, an increase in mitochondria to lysosome contacts was observed (as defined by <20-nm distances between mitochondrial outer membrane and lysosome) only in the WT 991-treated condition (fig. S8, C to E).

Finally, we conducted Seahorse experiments to measure oxygen consumption rate (OCR) as a measure of mitochondrial function in WT cells or cells lacking AMPK (Fig. 6J), WT FNIP1 and SA5 FNIP1 cells (Fig. 6K), and WT and cells lacking ERRα (Fig. 6L). OCR was decreased in cells lacking AMPK compared with that in WT cells. Concordantly, OCR was also reduced in the SA5 FNIP1 cell line compared with WT FNIP1 cells. Moreover, a comparison of WT cells with cells lacking ERRα revealed that OCR was also decreased when ERRα was deleted. This further supports the common role of AMPK, FNIP1, and ERRα in controlling mitochondrial biogenesis and function.

Discussion

FNIP1 was originally identified as a binding partner for the FLCN hamartoma suppressor (*34*) and in the same study was found to co-immunoprecipitate with endogenous AMPK subunits, although a functional role for FNIP1 mediating aspects of AMPK function has never been examined in the decades since. Subsequent studies have focused on how AMPK signaling is hyperactivated in FLCN-deficient

states (59–62). Our findings indicate that the physiological relationship between AMPK and FLCN-FNIP1 is that after metabolic stress, AMPK lies upstream of FLCN-FNIP1, wherein AMPK-dependent phosphorylation of FNIP1 acutely inhibits FLCN-FNIP1 GAP function, leading to TFEB activation.

Our results are consistent with the emerging model that the FNIP1-FLCN GAP complex controls RagC to retain TFEB and TFE3 at the lysosome and that TFEB and TFE3 are specific and selective substrates of mTORC1 in this regulation (15, 36, 37). A critical arginine required for GAP activity of FLCN in the FLCN-FNIP1 complex is required for AA-induced control of translocation of TFEB and TFE3 (15). Our data indicate that AMPK activation inhibits FLCN-FNIP1 GAP activity and promotes the accumulation of GTP-loaded RagC, the inactive state of the Rag complex. This mechanism allows activation of TFEB and TFE3 under low-energy conditions, even if AAs are plentiful. AMPK phosphorylation of FNIP1 promotes dissociation of RagC and mTOR from the lysosome as well as separation of mTOR from TFEB itself, consistent with loss of mTORC1-mediated TFEB phosphorylation. In cells that are AMPK deficient or just mutated in the five AMPK sites in FNIP1, even in the face of energy stress, the FNIP1-FLCN complex cannot be regulated by AMPK, and mTOR remains bound to TFEB at the lysosomal surface, rendering TFEB resistant to activation by mitochondrial poisons, glucose deprivation, or direct AMPK activators. By contrast, these cells still exhibit normal regulation of TFEB in response to AA withdrawal. Thus, AMPK maintains specific control of TFEB and TFE3 through its phosphorylation of FNIP1, independent of AA regulation of the FLCN-FNIP1 complex. Collectively, FNIP1 phosphorylation by AMPK at Ser^{220}, Ser^{230}, Ser^{232}, Ser^{261}, and Ser^{593} is a key mechanism by which AMPK controls the MiT-TFE family of transcription factors to increase lysosomal biogenesis and in parallel to increase PGC1α mRNA, which induces mitochondrial biogenesis, placing FNIP1 at the center of multiple AMPK-dependent processes for which the direct biochemical substrate of AMPK had remained elusive.

Materials and methods

Antibodies and reagents

Abcam antibodies used include Total OXPHOS Rodent WB Antibody Cocktail [catalog no. (Cat#) ab110413], monoclonal anti-UQCRB (Cat# ab190360 [EPR15591]), monoclonal anti-NDUFB3 (Cat# ab202585 [EPR15571]), monoclonal Tomm20 (Cat# ab56783), and monoclonal LAMP2 (Cat# ab25631). Cell Signaling Technology (CST) antibodies used in the study were as follows: TFEB (Cat# 4240), monoclonal TFEB (D2O7D, Cat# 37785), monoclonal Phospho-TFEB S122 (Cat# 86843), Tfe3 (Cat# 14779), FNIP1 (Cat# 36892), Phospho-FNIP1 S220 (Cat# 40812, in development), FLCN (Cat# 3697 (D14G9), RagA (D8B5, Cat# 4357), RagB (D18F3 Cat# 8150), RagC (D8H5, Cat# 9480), IDH2 (D8E3B, Cat# 56439), PDHA1 (C54G1, Cat# 3205), Tricarboxylic Acid Cycle Antibody Sampler Kit (Cat# 47767), monoclonal anti-LAMP2 (H4B4, Abcam Cat# ab25631), monoclonal anti-LAMP1 (D2D11, Cat# 9091), ERRα (Cat# 13826), AMPKα (Cat# 2532), Phospho AMPKα T172 (40H9, Cat# 2535), ACC (Cat# 3662), Phospho-ACC S79 (Cat# 3661), Raptor (Cat# 2280), Phospho-Raptor S792 (Cat# 2083), anti-Ulk1 (D8H5 Cat# 8054), Phospho-S6K T389 (Cat# 9205), 4EBP1(Cat# 9452), S6 (5G10, Cat# 2217), Phospho-S6 S235/236 (Cat# 4858), mTOR (Cat# 2972), LAMTOR1/C11orf59 (D11H6, Cat# 8975), monoclonal Hdac3 (7G6C5 Cat# 3949), Golgin-97 (D8P2K Cat# 13192), monocloncal GAPDH (D16H11, Cat# 8884), GFP (D5.1, Cat# 2956), and DYKDDDDK (FLAG) tag (Cat# 2368). A-Tubulin antibody was from Sigma-Aldrich (B-5-1-2, Cat# T5168). Monoclonal anti-PGC1α antibodies were from Santa Cruz Biotechnology (D-5, Cat# sc-518025) and Millipore (4C1.3, Cat# ST1202). Polyclonal Tfeb antibody was from Bethyl Laboratories (Cat# A303-673A). GFP-Trap Agarose was from ChromoTek (Cat# gta-20) and Pierce HA Magnetic Beads were from Thermo Fisher Scientific (Cat# 88837). 991 was purchased from Glixx Laboratories (Cat# GLXC-09267), and phenformin hydrochloride (Cat# P7045), rotenone (Cat# R8875), CCCP (Cat# C2759), and metformin (Cat# PHR1084) were from Sigma-Aldrich.

Plasmids

The cDNA encoding human FNIP1 (Uniprot Q8TF40) was generated by reverse transcription polymerase chain reaction (RT-PCR) from RNA obtained from IMR90 cells. Fusion tag FLAG-CHERRY was added to the N-terminal end of FNIP1 and subcloned into pDONR221 with BP Clonase (Invitrogen). The cDNA for human FLCN was obtained from Invitrogen (IOH12359). Mammalian expression vectors were generated, by recombining ENTR clones into DEST vectors using LR Clonase (Invitrogen). Destination vectors include: pLentiCMV/TO (Addgene no. 17293), pcDNA3 N-term FLAG DEST, pcDNA3 N-Term MYC DEST, and pQCXIB CMV/TO (Addgene no. 17400). Site-directed mutagenesis for FNIP1 was performed using QuikChange II XL (Stratagene) according to the manufacturer's instructions. Untagged human AMPKα1 cDNA was cloned into pLentiCMV/TO puro Gateway destination vector. Other plasmids used include pLJC5-Tmem192-2xFlag (41) (Addgene no. 102929), pLJC5-Tmem192-3xHA (41) (Addgene no. 102930), pEGFP-N1-TFEB (5) (Addgene no. 38119), pRK5-HA GST RagC 120L (Addgene no. 19306), and pRK5-HA GST RagC 75L (Addgene no. 19305).

Cell culture and cell lines

All cells were cultured in Dulbecco's modified essential medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS) (Hyclone, Thermo Fisher Scientific), 2 mM l-glutamine, penicillin/streptomycin (Gibco) at 37°C in 5% CO_2 and maintained under antibiotic selection for stable cell lines. Stably reexpressing WT FNIP1 and SA5 FNIP1 cells were generated by using lentivirus-mediated transduction of FNIP1 KO HEK293T cells with human Myc–tagged FLCN cDNA, in combination with either FLAG-Cherry–tagged WT FNIP1, SA4 FNIP1, or SA5 FNIP1 cDNA under double puromycin (Sigma) and hygromycin (Invitrogen) selection. GFP-TFEB cell lines were generated by stable infection of WT FNIP1 and SA5 FNIP1 HEK293T cells with lentivirus expressing GFP-TFEB cDNA and blasticidin resistance. Lyso-IP cells were generated by stable infection of WT FNIP1 or SA5 FNIP1 with lentivirus expressing pLJC5-Tmem192-3xHA or pLJC5-Tmem192-2xFlag cDNA. Mitochondrial poisons and AMPK-activating drugs were used at the following concentrations: 991 (50 μM), CCCP (5 μM), phenformin (2.5 mM), and rotenone (1 mM). For AA starvation experiments, cells were first washed with RPMI AA-free medium supplemented with 10% dialysed FBS, then the same media added for the times indicated in figures. For AA-replete conditions, RPMI AA-free medium was supplemented with 10% dialysed FBS in addition to 1x L-glutamine, 1x essential AAs, and 1x nonessential AAs for the times indicated in the figures. For glucose deprivation experiments, glucose-free DMEM media (Invitrogen) was supplemented with 10% dialysed FBS and a range of glucose concentrations including 1 mM, 2.5 mM, and 25 mM. For transient expression of proteins and packaging of virus, HEK293T cells were transfected with the plasmid of interest using Lipofectamine 2000 (Invitrogen, Carlsbad, CA) following the manufacturer's protocol.

Mouse studies

All procedures using animals were approved by the Salk Institute Institutional Animal Care and Use Committee (IACUC) protocol 11-000029. AMPKa1 and AMPKa2 floxed allele (*Prkaa1*$^{fl/fl}$; *Prkaa2*$^{fl/fl}$) mice bearing *Albumin*-creERT2 were treated with tamoxifen (1 mg per day) or vehicle (control) for 5 consecutive days. Eight weeks after tamoxifen injection, mice were fasted overnight, refed for 1 hour, then for 2 hours with vehicle or MK-8722 (30 mg/kg) before euthanizing as previously described (63). Livers were collected, and lysates were prepared in lysis buffer. In Fig. 2F, primary hepatocytes were made and treated with metformin as previously described (64).

Western blots

For biochemical analysis of cells, cells were washed with ice-cold phosphate-buffered saline (PBS) and lysed in buffer containing 20 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100, 2.5 mM pyrophosphate, 50 mM NaF, 5 mM β-glycero-phosphate, 50 nM calyculin

A, 1 mM Na_3VO_4, and protease inhibitors (Roche). Lysates were clarified by centrifugation at 16,000 ×*g* for 10 min at 4°C. Protein concentration was calculated using the BCA protein kit (Pierce). Lysates were resolved on 10 to 12% SDS–polyacrylamide gel electrophoresis (SDS-PAGE) gels, depending on molecular weight of proteins assessed and immunoblotted. Nuclear and cytoplasmic fractions were isolated using a NE-PER nuclear and cytoplasmic extraction kit (Thermo Fisher Scientific) according to manufacturer's instructions.

Immunoprecipitation

Cells were lysed in standard lysis buffer as described above and samples equilibrated. GFP-TFEB was immunoprecipitated from 2 mg of lysates using GFP-Trap beads (Chromotek) at 4°C for 2 hours under rotation. Subsequently, the beads were washed three times with lysis buffer. Protein complexes were eluted from beads using Laemmli sample buffer with 2% (v/v) beta-mercaptoethanol. Coimmunoprecipitating proteins were detected by immunoblot analysis.

Immunopurification of lysosomes (Lyso-IPs)

Lyso-IP cells were generated by infection with viruses containing Tmem192-3xHA (HA-Lyso) or Tmem192-2xFlag (Cont-Lyso). The Lyso-IP protocol was carried out as described in Abu-Remaileh *et al.* (*41*). Briefly, WT FNIP1/HA-Lyso or SA5 FNIP1/HA-Lyso cells in addition to control cells were treated with short 9*1* time courses and washed with PBS then scraped into 1-ml KPBS (136 mM KCl, 10 mM KH2PO4, pH 7.25 adjusted with KOH). Cells were gently homogenized with 20 strokes of a 2-ml homogenizer. The homogenate was then centrifuged at 1000 ×*g* for 2 min and the supernatant containing the cellular organelles including lysosomes was incubated with 100 μl beads of anti-HA magnetic beads, prewashed in KPBS, on a gentle rotator for 15 min at 4°C. Immunoprecipitates were then gently washed three times with KPBS and the lysosome fraction was eluted from the beads through incubation in lysis buffer for 10 min.

CRISPR-Cas9 techniques

Small guide RNAs (sgRNAs) targeting human TFEB, TFE3, and ERRα were selected using the CRISPR design tool Benchling program https://www.benchling.com/crispr/. Guides with high targeting scores and low probability of off-target effects were chosen (table S1). At least three independent sgRNA sequences were tested for each gene. Oligonucleotides for sgRNAs were synthesized by IDT, annealed in vitro, and subcloned into BsmbI-digested plentiCRISPRv.2-puro (Addgene no. 52961) or lentiCRISPRv.2-blast (Addgene no. 98293). Validation of guide specificity was assessed by Western blotting of low-passage cells after puromycin selection.

HEK293T AMPK KO cells were generated using the Cas9 nickase strategy. Briefly, a pair of guide RNAs (gRNAs) targeting exon 1 was designed for both human *PRKAA1* and *PRKAA2* genes using the online design tool at http://crispr.mit.edu (AMPKα1 A/B and AMPKα2 A/B, respectively) (table S1). Each gRNA duplex was cloned into pX462 vector encoding SpCas9n-2A-Puro (Addgene no. 48141). HEK293T cells were transfected with the gRNA pair to generate single AMPKα1 or AMPKα2 KO or transfected with both pairs together to generate double AMPK α1/α2 KO (DKO). After puromycin selection, single-cell cloning was performed by cell sorting into 96-well plates. Individual clones were screened by Western blot and a clone lacking both AMPK α1 and α2 protein expression was selected. HEK293T FNIP1 KO cells were made using a single gRNA (table S1) targeting exon 2 and subsequently cloned into pX459 encoding SpCas9(BB)-2A-Puro (Addgene no. 48139). Single-cell clones were isolated after selection with puromycin and screened by immunoblotting.

RNA interference studies

All siRNAs were purchased from Horizon Discovery. siRNAs against a nontargeting sequence was used as a negative control (ON-TARGET-*plus* nontargeting control siRNA, Cat# D-001810-01-05) or ON-TARGET*plus* SMARTpool siRNAs targeting PPARGC1A (L-005111-00-0005) or targeting TFEB (Cat# L-009798-00-00050) or TFE3 (Cat# L-009363-00-0005) were used. HEK293T cells were plated in a 6-well plate and allowed to adhere overnight. Cells were transfected with the nontargeting control siRNA pool (20 nM), or PPARGC1A siRNA pool (20 nM). Transfection was carried out using Lipofectamine RNAiMAX (Invitrogen) according to the manufacturer's protocol. 48 hours after transfection, cells were treated with 9*1* as indicated in the figures and then harvested for downstream applications.

Lentivirus production

Lentiviruses were produced by transfecting HEK293T cells with the cDNA construct of interest including 3xHA-Tmem192, 2xFLAG-Tmem192, FLAG-WT FNIP1, FLAG-SA4 FNIP1, FLAG-SA5 FNIP1, MYC-FLCN, or GFP-TFEB constructs, in combination with VSV-G and CMV-ΔVPR packaging plasmids. Sixteen hours later, the media was changed to DMEM with 10% FBS. The virus containing supernatant was collected the next day and 0.45-μm filtered, then frozen at −80°C or supplemented with 8 μg/ml polybrene and applied to destination cells for 24 hours. Sixteen hours later, the media was refreshed and the appropriate antibiotic was added for selection.

In vitro kinase assays

HEK293T cells transiently transfected with FLAG-FNIP1 or FLAG-tagged FNIP1 mutants (SA2, SA3, SA4, SA5) were lysed and subjected to FLAG immunoprecipitation. Immunoprecipitates were washed three times in lysis buffer, followed by three times in kinases assay buffer (50 mM Tris pH 7.5 10 mM MgCl2). Subsequently, immunoprecipitates were subjected to a kinase reaction containing 0.1 mM [γ32P]-ATP (PerkinElmer) and 2 mM dithiothreitol (DTT) with or without 0.1 μg of active recombinant 50 ng of active recombinant AMPK (Millipore no. 14-840) in the presence of kinase assay buffer. The reaction was incubated at 30°C for 30 min. Reactions were terminated with LDS sample buffer and resolved by SDS-PAGE electrophoresis. Proteins were detected with Coomassie staining. Dried gels were exposed to UltraCruz autoradiography film overnight, in an autoradiography cassette and the films were later developed using an auto-developer.

Immunofluorescence

Cells were seeded on coverslips (precoated with poly-L-lysine). Following treatments described in figure legends, cells were fixed with 4% (v/v) paraformaldehyde and permeabilized with 1% (v/v) NP-40. Cells were blocked using 5% bovine serum albumin (BSA) in PBS, then incubated for 1 hour with primary antibodies, followed by three washes in 0.2% BSA/PBS. Coverslips were incubated for 1 hour with secondary antibodies and counterstained with 4′,6-diamidino-2-phenylindole (DAPI) for 5 min, then washed three times in 0.2% BSA/PBS and once in water. Coverslips were mounted with Fluoromount-G (Southern Biotech).

Confocal imaging of TFEB, TFE3, and RagC

For visualization of TFEB, TFE3, and RagC, the following primary rabbit antibodies were used: TFEB (CST no. 37785), TFE3 (CST no.14779), and RagC (CST no. 9480) followed by anti-rabbit secondary antibodies conjugated to Alexa Fluor (AF) 594. Cells were imaged on a Zeiss LSM 700 confocal microscope using the 63X objective. Quantification was performed using imageJ software.

Airyscan confocal imaging of IDH2, Cox IV, and Lamp2

For imaging IDH2, coverslips were incubated in anti-IDH2 (CST no. 56439) antibody, followed by secondary anti-rabbit antibody conjugated to AF-568. IDH2 was costained with phalloidin 647 and Tomm20 using anti-Tomm20 (Abcam no. ab56783) antibody, followed by secondary antibody conjugated to AF-488. Cox IV and Lamp2 were costained with anti-Cox IV (CST no. 11967) and anti-Lamp2 (Abcam no. ab18528) antibodies, followed by anti-rabbit secondary conjugated to AF-594 and anti-mouse secondary conjugated to AF-488, respectively. Cells were imaged with a 63x 1.4 NA oil objective on a ZEISS 880 LSM Airyscan confocal system with an inverted stage. High-resolution Airyscan images were acquired using a pixel dwell time

of 0.66 μs and 2x Nyquist pixel size of 43 nm per pixel in SR mode (i.e., a virtual pinhole size of 0.2 Airy units), then processed using ZEISS Zen software with the Airyscan parameter determined by auto-filter settings. The zoom factor was set to 2 to obtain a large field of view. Alexa Fluor 488 was imaged with a 488-nm laser with a laser power of ~64 μW, Alexa Fluor 568 and 594 were imaged with the 561-nm laser with a laser power of ~268 μW, and Alexa Fluor 647 was imaged with the 633-nm laser with a laser power of ~98 μW.

Fluorescence image quantification

IDH2 fluorescence intensity quantification was performed with Imaris (version 9.6.0, Bitplane, Zurich, Switzerland). Mitochondrion regions were masked by surfaces generated TOM20 fluorescence using automatic threshold settings in Imaris. Mitochondrial associated IDH2 was quantified by summing up the IDH2 voxel intensity values inside the segmented TOM20 regions and normalizing to the mitochondria volume. Statistical significance was determined using an unpaired *t* test. Quantification was performed with Imaris software 9.6.0 (Bitplane). COX IV fluorescence image channels were preprocessed with a Gaussian smoothing filter with a filter size of 0.0707 μm. Background subtraction was performed with an estimation of the diameter of the largest sphere that fits into the object as 0.265 μm. Mitochondrial clusters were then defined by surfaces generated with automatic threshold settings, which makes use of an iterative selection method. Clusters smaller than 0.003 μm^3 were removed. Clusters for LAMP 2 fluorescence signal were defined similarly but without smoothing or background subtraction. Touching objects were separated using region growing with a seed point diameter of 0.25 μm. The Imaris quality filter with a lower threshold of 86.5 was then used to select positive signals. Clusters smaller than 0.065 μm^3 were removed. Cell volume was defined by first smoothing the LAMP2 signal with a Gaussian filter with a diameter of 1 μm. Surfaces generated with manual threshold of 41.5 absolute intensity. Lysosomes colocalized with mitochondria were defined as lysosomes located within 0.2 μm from its closest mitochondria cluster. The lysosome colocalized volume, total lysosome volume, and mitochondria volume were normalized to the cell volume. Statistical significance was determined using an unpaired *t* test.

Electron microscopy imaging and quantification

Cells were cultured on 10-cm dishes to reach ~70% confluence before fixation. Materials were sourced from Electron Microscopy Sciences (Hatfield, PA) unless noted otherwise. Culture media was gently poured off and 4 ml of warm 37°C fixative (3% glutaraldehyde in 0.1 M sodium cacodylate buffer with 3 mM CaCl2) was added to the dish before being replaced with ice-cold fixative after 10 s. After an hour of fixation at 4°C, cells were rinsed with 0.1 M sodium cacodylate buffer with 3 mM CaCl2 three times for 10 min and postfixed with reduced osmium tetroxide (0.1 M sodium cacodylate buffer, 0.1 M CaCl2, 1.5% osmium tetroxide, 1.5% potassium ferrocyanide) for 40 min in the dark at room temperature. Dishes were rinsed with ice-cold deionized water three times, with the final rinse of 1 ml of water left in the dish, and scraped with a sharpened piece of Teflon secured in a hemostat. The cell suspension was collected into Eppendorf tubes and left overnight at 4°C. The following day, cells were stained with aqueous 1% uranyl acetate at room temperature for an hour before serial dehydration in ice-cold ethanol solutions. Dehydrated cells were rinsed twice with anhydrous ethanol at room temperature and infiltrated with Eponate 12 resin (hard formulation) for 2 hours at 3:1 and 1:1 ethanol:resin mixtures and left in 1:3 resin to infiltrate overnight. The following day, cells were infiltrated in two changes of pure resin throughout the day and pelleted in a third change of fresh resin at 12,000 rpm in a tabletop centrifuge (Pelco) and left to polymerize in an oven at 60°C for 48 hours.

Polymerized blocks were removed from Eppendorf tubes using razor blades and trimmed for utlrathin serial sectioning as previously described (*65*) using diamond knives (Diatome) mounted on a Zeiss UC8 ultramicrotome. Series of ultrathin (70-nm) sections were collected onto silicon chips and imaged using a scanning electron microscope (Zeiss Sigma VP) equipped with a backscattered electron detector (Gatan) using Atlas5 (Fibics) control software and scan generator. Ribbons of serial sections from each condition were screened to identify a cluster of cells suitable for further analysis. For each series, the region of interest in each section was identified and captured at an overview resolution (200 nm per pixel) and a midlevel resolution (8 nm per pixel) to minimize drift during high-resolution (2 nm per pixel) acquisition. High-resolution stacks of images were aligned. Approximately 50 mitochondria from two randomly sampled subvolumes of registered 3DEM image stacks were segmented using Vast Lite (*66*). The two volumes (total 105.28 μm^3, average = 52.64 μm^3) of densely labeled mitochondria (n = 53, average = 26.5) from the WT-DMSO condition were proofread. Because of the prevalence of touching mitochondria, all labels were first eroded to ensure separated boundaries. A three-dimensional (3D) U-Net was used to detect boundaries and LSDs in a multitask learning framework. The network had an input shape of [48,284,284] and output shape of [16,196,196] (voxels, zyx). It consisted of three layers and was downsampled by a factor of [1,2,2] in the first two layers and [2,2,2] in the final layer. The reverse was done for the up-sampling path. Twelve initial feature maps were used and multiplied by a factor of 5 between layers. The number of feature maps in the last layer was increased to 14 to account for the 13 feature maps generated by the boundaries (3) and LSDs (10). A mean squared error loss and Adam optimizer were used to train the network. A single voxel neighborhood [1,1,1] was used for the boundaries. The LSDs used a sigma of 140 and were down-sampled by a factor of 2. Training was done using Tensorflow and Gunpowder. Inference, seeded watershed, and percentile agglomeration were performed in a block-wise fashion for each condition. Four subvolumes from the non–WT-DMSO conditions (1 WT-991, 1 SA5-DMSO, 2 SA5-991, 116.73 μm^3, average = 29.2 μm^3) were proofread (n = 110, average = 27.5) and used to refine training along with the original two volumes. Additionally, 11 negative samples which did not contain any mitochondria (i.e., resin, cytosol, background) were used for retraining (~450 μm^3, average = 40.9 μm^3). Each batch randomly sampled from one of the 17 total subvolumes such that the probability of choosing either a positive or negative sample was 50%. After training, segmentations were created for each condition, and small objects (i.e., oversegmented debris) were filtered out. Mitochondria were then randomly sampled from each volume (n = 100) and proofread for subsequent analysis. Iterative cycles of training and manual proofreading were used until a fairly accurate but excessively permissive segmentation of mitochondria was achieved for all conditions. One hundred mitochondria were randomly sampled from the machine-generated segments. Segmentations were meticulously proofread by human experts before inclusion in analysis. Contingency tables of sampled mitochondria with and without lysosomal contacts were assembled for pairs of experimental conditions (i.e., combinations of WT/SA5, DMSO/991 treatment). Chi square tests were performed in Python using the chi2_contingency function in the scipy.stats library.

Mass spectrometry

HEK293T cells were transfected with an epitope-tagged FNIP1 cDNA, and cells were treated with DMSO or phenformin. After immunopurification of FNIP1 protein, FNIP1 was isolated from an SDS-polyacrylamide gel. Bands on the gels were cut out and subjected to reduction with dithiothreitol, alkylation with iodoacetamide, and in-gel digestion with chymotrypsin overnight at pH 8.3, followed by reversed-phase microcapillary or liquid chromatography with

tandem mass spectrometry (LC-MS/MS). LC-MS/MS was performed using an Easy-nLC nanoflow HPLC (Proxeon Biosciences) with a self-packed 75 μm id × 15 cm C18 column coupled to a LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific) in the data-dependent acquisition and positive ion mode at 300 nL/min. Peptide ions from predicted phosphorylation sites were also targeted in MS/MS mode for quantitative analyses. MS/MS spectra collected through collision-induced dissociation in the ion trap were searched against the concatenated target and decoy (reversed) single entry and full Swiss-Prot protein databases using Sequest (Proteomics Browser Software, Thermo Fisher Scientific) with differential modifications for Ser/Thr/Tyr phosphorylation (+79.97) and the sample processing artifacts Met oxidation (+15.99), deamidation of Asn and Gln (+0.984) and Cys alkylation (+57.02). Phosphorylated and unphosphorylated peptide sequences were identified if they initially passed the following Sequest scoring thresholds against the target dtabase: 1+7 ions, Xcorr ≥ 2.0 Sf ≥ 0.4, P ≥ 5; 2+ ions, Xcorr ≥ 2.0, Sf ≥ 0.4, P ≥ 5; 3+ ions, Xcorr ≥ 2.60, Sf ≥ 0.4, P ≥ 5 against the target protein database. Passing MS/MS spectra were manually inspected to ensure that all b- and y-fragment ions aligned with the assigned sequence and modification sites. Determination of the exact phosphorylation sites was aided using FuzzyIons and GraphMod and phosphorylation site maps were created using ProteinReport software (Proteomics Browser Software suite, Thermo Fisher Scientific). False discovery rates (FDRs) of peptide hits (phosphorylated and unphosphorylated) were estimated below 1.5% based on reversed database hits.

Seahorse assays

OCR of cells were measured using the Seahorse XF96 Cell Mito Stress Test Kit (Seahorse Biosciences) with an XF96 Extracellular Flux Analyzer (Seahorse Bioscience) in accordance with the manufacturer's instructions. Seahorse XF Cell Mito Stress Test Kit (Cat# 103015–100) and Seahorse xFe96 FluxPak (Cat# 102416-100) were purchased from Agilent Technologies. The XF96 cell culture microplates were polylysine coated before seeding HEK293T cells in XF base media supplemented with 1 mM pyruvate, 2 mM glutamine, and 10 mM glucose. Each condition was seeded in six replicate wells. The Seahorse sensor cartridge was hydrated overnight in a non-CO_2 incubator at 37°C. On the day of the assay, cells were incubated in a CO_2-free incubator at 37°C for 1 hour to allow for temperature and pH equilibration before loading into the XF96 apparatus. Mitochondrial stress tests were conducted following Seahorse guidelines (Agilent Technologies). Inhibitors were used at the following concentrations: 1 μM oligomycin, 1 μM FCCP, and 0.5 μM antimycin A + 0.5 μM rotenone. Analyses were conducted using Wave software (Agilent Technologies).

mRNA extraction and qPCR

RNA extracted from cells using QIAGEN rNeasy Plus mini kit (Cat# 74134) and 1 μg of RNA was reverse transcribed using Iscript cDNA Synthesis Kit (Bio-Rad, Cat# 1708891). qPCR used diluted cDNA, relevant primers, and SYBR Green PCR master mix (Thermo Fisher Scientific, Cat# 4309155) in a C1000 Thermal Cycler (BioRad). Relative mRNA levels calculated using the $\Delta\Delta C_T$ method, with β-actin serving as the internal control. All mRNA measurements performed in triplicate. All primers are listed in table S1.

Mitochondrial DNA content analysis

Mitochondrial DNA was quantified by determining the mtDNA/nDNA ratio. Total genomic DNA was extracted from cells using the Qiagen dNeasy Blood & Tissue Kit (Cat# 69504). Extracted DNA was diluted to a final concentration of 10 ng of DNA per microliter, and mitochondrial DNA was quantified relative to the nuclear DNA specific gene β-actin by qPCR. Primers used for β-actin and mitochondrial 16*S* rRNA genes are listed in table S1. Reactions used SYBR Green PCR Master Mix (Thermo Fisher Scientific), using a C1000 Thermal Cycler (BioRad). The relative mtDNA copy number was calculated using the $\Delta\Delta C_T$ method.

RNA-seq

Cells were administered with 991, CCCP, rotenone, or phenformin time courses, ranging from 0 to 24 hours. Each time point and condition was prepared in triplicate. Total RNA was isolated using the QIAGEN rNeasy Plus mini kit. The quality of the isolated total RNA was assessed using Agilent TapeStation 4200 and RNA-seq libraries were prepared with 500 ng of total RNA using the TruSeq stranded mRNA Sample Preparation Kit according to the manufacturer's protocol (Illumina). Libraries were quantified, multiplexed, and pooled for sequencing at paired-end 75 base pairs using the Illumina NextSeq500 or NovsSeq6000 platform at the Salk Next-Generation Sequencing Core. Raw sequencing data were demultiplexed and converted in the FASTQ files using CASAVA (version 1.8.2). Libraries were sequenced at an average depth of 12 to 40 million reads per sample.

Bioinformatic analysis of RNA-seq data

Raw RNA-seq reads in FASTQ files were quality-tested using FASTQC (v0.11.8) (Andrews 2010) and mapped to the human reference genome (GRCh38) with STAR (v2.5.3a) aligner with default parameters (*67*). Raw or TPM (transcripts per million) gene expression levels were quantified across all the exons of the top isoform in RefSeq with analyzeRepeats.pl in HOMER (v4.11.1) (*68*). Differential expression analysis was performed on the Biojupies platform (https://amp.pharm.mssm.edu/biojupies/) (*69*) with raw gene counts, comparing gene expression levels between control and experimental groups using the limma R package (*70*). Replicates were used to compute within-group dispersion and correction for batch effects. Volcano plots were generated to display the results of differential gene expression analysis using VolcanoseR (*71*). Gene fold changes were $\log_2$ transformed and displayed on the x axis; P values were corrected using the Benjamini-Hochberg method, transformed using $-\log_{10}$ and displayed on the y axis. Differentially expressed genes were defined as having a P value ≤ 0.05; FC ≥ 1.3. Red points indicate significantly up-regulated genes, and blue points indicate significantly down-regulated genes. Raw gene counts were first normalized to TPM (transcripts per million). TPM counts were log transformed, filtered, scaled, and centered before gene clustering and heatmap generation using ClustVis (*72*), Heatmapper (*73*), or the R package (pheatmap). ClustVis uses code from BoxPlotR; several R packages are used internally, including shiny, ggplot2, pheatmap, gridSVG, rColorBrewer, FactoMineR, pcaMethods, gProfileR, shinyBS, shinyjs, and others. The source code of ClustVis is available in GitHub. For enrichment analyses, the up-regulated and down-regulated gene sets were generated by extracting genes with the respectively highest and lowest values from the gene expression signature. The gene sets were subsequently submitted to Enrichr (*74*), which is freely available at http://amp.pharm.mssm.edu/Enrichr/. The following libraries were used for the analysis: GO_Cellular_Component_2018, WikiPathways_2016, and ChEA_2016. Significant terms were determined by using a cut-off of $P < 0.1$ after applying Benjamini-Hochberg correction. GSEA was carried out with GSEA desktop v4.0.3 using preranked lists generated from FDR values, setting gene set permutations to 1000, using the Hallmark or the c2 collections in mSigDB v7.2 or "KEGG Lysosome" gene sets. Four-way Venn diagrams were generated using Venny 2.1 (https://bioinfogp.cnb.csic.es/tools/venny/).

Statistical analysis

Statistical parameters including the exact value of n, measures (means ± SEMs), and statistical significance are reported in the figures and figure legends. Data are judged to be statistically significant when $P < 0.05$ by two-tailed Student's t test. In figures, asterisks denote statistical significance as calculated by Student's t test (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$). Statistical analyses were performed using Graph Pad Prism 7. Analysis of RNA-seq data has been described in the section "Bioinformatic analysis of RNA-seq data."

REFERENCES AND NOTES

1. A. González, M. N. Hall, S. C. Lin, D. G. Hardie, AMPK and TOR: The Yin and Yang of Cellular Nutrient Sensing and Growth Control. *Cell Metab.* **31**, 472–492 (2020). doi: 10.1016/j.cmet.2020.01.015; pmid: 32130880
2. D. G. Hardie, Keeping the home fires burning: AMP-activated protein kinase. *J. R. Soc. Interface* **15**, 20170774 (2018). doi: 10.1098/rsif.2017.0774; pmid: 29343628
3. S. Herzig, R. J. Shaw, AMPK: Guardian of metabolism and mitochondrial homeostasis. *Nat. Rev. Mol. Cell Biol.* **19**, 121–135 (2018). doi: 10.1038/nrm.2017.95; pmid: 28974774
4. D. P. Kelly, R. C. Scarpulla, Transcriptional regulatory circuits controlling mitochondrial biogenesis and function. *Genes Dev.* **18**, 357–368 (2004). doi: 10.1101/gad.1177604; pmid: 15004004
5. A. Roczniak-Ferguson *et al*., The transcription factor TFEB links mTORC1 signaling to transcriptional control of lysosome homeostasis. *Sci. Signal.* **5**, ra42 (2012). doi: 10.1126/scisignal.2002790; pmid: 22692423
6. R. Puertollano, S. M. Ferguson, J. Brugarolas, A. Ballabio, The complex relationship between TFEB transcription factor phosphorylation and subcellular localization. *EMBO J.* **37**, e98804 (2018). doi: 10.15252/embj.201798804; pmid: 29764979
7. C. Settembre *et al*., A lysosome-to-nucleus signalling mechanism senses and regulates the lysosome via mTOR and TFEB. *EMBO J.* **31**, 1095–1108 (2012). doi: 10.1038/emboj.2012.32; pmid: 22343943
8. S. Vega-Rubin-de-Celis, S. Peña-Llopis, M. Konda, J. Brugarolas, Multistep regulation of TFEB by MTORC1. *Autophagy* **13**, 464–472 (2017). doi: 10.1080/15548627.2016.1271514; pmid: 28055300
9. J. A. Martina, Y. Chen, M. Gucek, R. Puertollano, MTORC1 functions as a transcriptional regulator of autophagy by preventing nuclear transport of TFEB. *Autophagy* **8**, 903–914 (2012). doi: 10.4161/auto.19653; pmid: 22576015
10. N. P. Young *et al*., AMPK governs lineage specification through Tfeb-dependent regulation of lysosomes. *Genes Dev.* **30**, 535–552 (2016). doi: 10.1101/gad.274142.115; pmid: 26944679
11. L. J. Eichner *et al*., Genetic Analysis Reveals AMPK Is Required to Support Tumor Growth in Murine Kras-Dependent Lung Cancer Models. *Cell Metab.* **29**, 285–302.e7 (2019). doi: 10.1016/j.cmet.2018.10.005; pmid: 30415923
12. Z. Y. Tsun *et al*., The folliculin tumor suppressor is a GAP for the RagC/D GTPases that signal amino acid levels to mTORC1. *Mol. Cell* **52**, 495–505 (2013). doi: 10.1016/j.molcel.2013.09.016; pmid: 24095279
13. K. Shen *et al*., Cryo-EM Structure of the Human FLCN-FNIP2-Rag-Ragulator Complex. *Cell* **179**, 1319–1329.e8 (2019). doi: 10.1016/j.cell.2019.10.036; pmid: 31704029
14. C. S. Petit, A. Roczniak-Ferguson, S. M. Ferguson, Recruitment of folliculin to lysosomes supports the amino acid-dependent activation of Rag GTPases. *J. Cell Biol.* **202**, 1107–1122 (2013). doi: 10.1083/jcb.201307084; pmid: 24081491
15. R. E. Lawrence *et al*., Structural mechanism of a Rag GTPase activation checkpoint by the lysosomal folliculin complex. *Science* **366**, 971–977 (2019). doi: 10.1126/science.aax0364; pmid: 31672913
16. C. Settembre *et al*., TFEB links autophagy to lysosomal biogenesis. *Science* **332**, 1429–1433 (2011). doi: 10.1126/science.1204592; pmid: 21617040
17. M. Sardiello *et al*., A gene network regulating lysosomal biogenesis and function. *Science* **325**, 473–477 (2009). doi: 10.1126/science.1174447; pmid: 19556463
18. R. M. Perera, C. Di Malta, A. Ballabio, MiT/TFE Family of Transcription Factors, Lysosomes, and Cancer. *Annu. Rev. Cancer Biol.* **3**, 203–222 (2019). doi: 10.1146/annurev-cancerbio-030518-055835; pmid: 31650096
19. C. Collodet *et al*., AMPK promotes induction of the tumor suppressor FLCN through activation of TFEB independently of mTOR. *FASEB J.* **33**, 12374–12391 (2019). doi: 10.1096/fj.201900841R; pmid: 31404503
20. M. Paquette *et al*., AMPK-dependent phosphorylation is required for transcriptional activation of TFEB and TFE3. *Autophagy* **17**, 3957–3975 (2021). doi: 10.1080/15548627.2021.1898748; pmid: 33734022
21. W. W. Winder *et al*., Activation of AMP-activated protein kinase increases mitochondrial enzymes in skeletal muscle. *J. Appl. Physiol.* **88**, 2219–2226 (2000). doi: 10.1152/jappl.2000.88.6.2219; pmid: 10846039
22. R. Bergeron *et al*., Chronic activation of AMP kinase results in NRF-1 activation and mitochondrial biogenesis. *Am. J. Physiol. Endocrinol. Metab.* **281**, E1340–E1346 (2001). doi: 10.1152/ajpendo.2001.281.6.E1340; pmid: 11701451
23. H. Zong *et al*., AMP kinase is required for mitochondrial biogenesis in skeletal muscle in response to chronic energy deprivation. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 15983–15987 (2002). doi: 10.1073/pnas.252625599; pmid: 12444247
24. R. M. Reznick, G. I. Shulman, The role of AMP-activated protein kinase in mitochondrial biogenesis. *J. Physiol.* **574**, 33–39 (2006). doi: 10.1113/jphysiol.2006.109512; pmid: 16709637
25. S. Jäger, C. Handschin, J. St.-Pierre, B. M. Spiegelman, AMP-activated protein kinase (AMPK) action in skeletal muscle via direct phosphorylation of PGC-1α. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 12017–12022 (2007). doi: 10.1073/pnas.0705070104; pmid: 17609368
26. C. Cantó, J. Auwerx, PGC-1α, SIRT1 and AMPK, an energy sensing network that controls energy expenditure. *Curr. Opin. Lipidol.* **20**, 98–105 (2009). doi: 10.1097/MOL.0b013e328328d0a4; pmid: 19276888
27. S. Kolmykov *et al*., GTRD: An integrated view of transcription regulation. *Nucleic Acids Res.* **49**, D104–D111 (2021). doi: 10.1093/nar/gkaa1057; pmid: 33231677
28. G. R. Steinberg, D. Carling, AMP-activated protein kinase: The current landscape for drug development. *Nat. Rev. Drug Discov.* **18**, 527–551 (2019). doi: 10.1038/s41573-019-0019-2; pmid: 30867601
29. V. K. Mootha *et al*., Errα and Gabpa/b specify PGC-1α-dependent oxidative phosphorylation gene expression that is altered in diabetic muscle. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 6570–6575 (2004). doi: 10.1073/pnas.0401401101; pmid: 15100410
30. S. N. Schreiber *et al*., The estrogen-related receptor α (ERRα) functions in PPARγ coactivator 1α (PGC-1α)-induced mitochondrial biogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 6472–6477 (2004). doi: 10.1073/pnas.0308686101; pmid: 15087503
31. S. Rath *et al*., MitoCarta3.0: An updated mitochondrial proteome now with sub-organelle localization and pathway annotations. *Nucleic Acids Res.* **49**, D1541–D1547 (2021). doi: 10.1093/nar/gkaa1011; pmid: 33174596
32. D. M. Gwinn *et al*., AMPK phosphorylation of raptor mediates a metabolic checkpoint. *Mol. Cell* **30**, 214–226 (2008). doi: 10.1016/j.molcel.2008.03.003; pmid: 18439900
33. C. Behrends, M. E. Sowa, S. P. Gygi, J. W. Harper, Network organization of the human autophagy system. *Nature* **466**, 68–76 (2010). doi: 10.1038/nature09204; pmid: 20562859
34. M. Baba *et al*., Folliculin encoded by the BHD gene interacts with a binding protein, FNIP1, and AMPK, and is involved in AMPK and mTOR signaling. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 15552–15557 (2006). doi: 10.1073/pnas.0603781103; pmid: 17028174
35. R. W. Myers *et al*., Systemic pan-AMPK activator MK-8722 improves glucose homeostasis but induces cardiac hypertrophy. *Science* **357**, 507–511 (2017). doi: 10.1126/science.aah5582; pmid: 28705990
36. G. Napolitano *et al*., A substrate-specific mTORC1 pathway underlies Birt-Hogg-Dubé syndrome. *Nature* **585**, 597–602 (2020). doi: 10.1038/s41586-020-2444-0; pmid: 32612235
37. B. S. Gosis *et al*., Inhibition of nonalcoholic fatty liver disease in mice by selective inhibition of mTORC1. *Science* **376**, eabf8271 (2022). doi: 10.1126/science.abf8271; pmid: 35420934
38. T. Sekiguchi, E. Hirose, N. Nakashima, M. Ii, T. Nishimoto, Novel G proteins, Rag C and Rag D, interact with GTP-binding proteins, Rag A and Rag B. *J. Biol. Chem.* **276**, 7246–7257 (2001). doi: 10.1074/jbc.M004389200; pmid: 11073942
39. M. Gao, C. A. Kaiser, A conserved GTPase-containing complex is required for intracellular sorting of the general amino-acid permease in yeast. *Nat. Cell Biol.* **8**, 657–667 (2006). doi: 10.1038/ncb1419; pmid: 16732272
40. Y. Sancak *et al*., The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. *Science* **320**, 1496–1501 (2008). doi: 10.1126/science.1157535; pmid: 18497260
41. M. Abu-Remaileh *et al*., Lysosomal metabolomics reveals V-ATPase- and mTOR-dependent regulation of amino acid efflux from lysosomes. *Science* **358**, 807–813 (2017). doi: 10.1126/science.aan6298; pmid: 29074583
42. M. Palmieri *et al*., Characterization of the CLEAR network reveals an integrated control of cellular clearance pathways. *Hum. Mol. Genet.* **20**, 3852–3866 (2011). doi: 10.1093/hmg/ddr306; pmid: 21752829
43. N. Pastore *et al*., TFE3 regulates whole-body energy metabolism in cooperation with TFEB. *EMBO Mol. Med.* **9**, 605–621 (2017). doi: 10.15252/emmm.201607204; pmid: 28283651
44. N. Pastore *et al*., Nutrient-sensitive transcription factors TFEB and TFE3 couple autophagy and metabolism to the peripheral clock. *EMBO J.* **38**, e101347 (2019). doi: 10.15252/embj.2018101347; pmid: 31126958
45. S. R. Bonam, F. Wang, S. Muller, Lysosomes as a therapeutic target. *Nat. Rev. Drug Discov.* **18**, 923–948 (2019). doi: 10.1038/s41573-019-0036-1; pmid: 31477883
46. C. Settembre *et al*., TFEB controls cellular lipid metabolism through a starvation-induced autoregulatory loop. *Nat. Cell Biol.* **15**, 647–658 (2013). doi: 10.1038/ncb2718; pmid: 23604321
47. G. Mansueto *et al*., Transcription Factor EB Controls Metabolic Flexibility during Exercise. *Cell Metab.* **25**, 182–196 (2017). doi: 10.1016/j.cmet.2016.11.003; pmid: 28011087
48. N. Salma, J. S. Song, Z. Arany, D. E. Fisher, Transcription Factor Tfe3 Directly Regulates Pgc-1alpha in Muscle. *J. Cell. Physiol.* **230**, 2330–2336 (2015). doi: 10.1002/jcp.24978; pmid: 25736533
49. P. R. Jannig, P. A. Dumesic, B. M. Spiegelman, J. L. Ruas, SnapShot: Regulation and biology of PGC-1α. *Cell* **185**, 1444–1444.e1 (2022). doi: 10.1016/j.cell.2022.03.027; pmid: 35427500
50. S. Miura, Y. Kai, Y. Kamei, O. Ezaki, Isoform-specific increases in murine skeletal muscle peroxisome proliferator-activated receptor-γ coactivator-1α (PGC-1α) mRNA in response to β2-adrenergic receptor activation and exercise. *Endocrinology* **149**, 4527–4533 (2008). doi: 10.1210/en.2008-0466; pmid: 18511502
51. V. Martínez-Redondo, A. T. Pettersson, J. L. Ruas, The hitchhiker's guide to PGC-1α isoform structure and biological functions. *Diabetologia* **58**, 1969–1977 (2015). doi: 10.1007/s00125-015-3671-z; pmid: 26109214
52. S. M. Soyal *et al*., A greatly extended *PPARGC1A* genomic locus encodes several new brain-specific isoforms and influences Huntington disease age of onset. *Hum. Mol. Genet.* **21**, 3461–3473 (2012). doi: 10.1093/hmg/dds177; pmid: 22589246
53. Y. Zhang *et al*., Alternative mRNA splicing produces a novel biologically active short isoform of PGC-1α. *J. Biol. Chem.* **284**, 32813–32826 (2009). doi: 10.1074/jbc.M109.037556; pmid: 19773550
54. B. L. Olson *et al*., SCF^{Cdc4} acts antagonistically to the PGC-1α transcriptional coactivator by targeting it for ubiquitin-mediated proteolysis. *Genes Dev.* **22**, 252–264 (2008). doi: 10.1101/gad.1624208; pmid: 18198341
55. J. S. Chang, S. Ghosh, S. Newman, J. M. Salbaum, A map of the PGC-1α- and NT-PGC-1α-regulated transcriptional network in brown adipose tissue. *Sci. Rep.* **8**, 7876 (2018). doi: 10.1038/s41598-018-26244-4; pmid: 29777200
56. J. M. Huss, I. P. Torra, B. Staels, V. Giguère, D. P. Kelly, Estrogen-related receptor α directs peroxisome proliferator-activated receptor α signaling in the transcriptional control of energy metabolism in cardiac and skeletal muscle. *Mol. Cell. Biol.* **24**, 9079–9091 (2004). doi: 10.1128/MCB.24.20.9079-9091.2004; pmid: 15456881
57. C. Luo *et al*., ERRα Maintains Mitochondrial Oxidative Metabolism and Constitutes an Actionable Target in PGC1α-Elevated Melanomas. *Mol. Cancer Res.* **15**, 1366–1375 (2017). doi: 10.1158/1541-7786.MCR-17-0143; pmid: 28596418
58. J. P. Rooney *et al*., in *Mitochondrial Regulation*, C. Palmeira, A. Rolo, eds., vol. 1241 of *Methods in Molecular Biology* (Humana Press, 2015), pp. 23–38.
59. O. M. Siggs *et al*., Mutation of Fnip1 is associated with B-cell deficiency, cardiomyopathy, and elevated AMPK activity. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E3706–E3715 (2016). doi: 10.1073/pnas.1607592113; pmid: 27303042
60. N. L. Reyes *et al*., Fnip1 regulates skeletal muscle fiber type specification, fatigue resistance, and susceptibility to muscular dystrophy. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 424–429 (2015). doi: 10.1073/pnas.1413021112; pmid: 25548157
61. M. Yan *et al*., The tumor suppressor folliculin regulates AMPK-dependent metabolic transformation. *J. Clin. Invest.* **124**, 2640–2650 (2014). doi: 10.1172/JCI71749; pmid: 24762438
62. E. Possik *et al*., Folliculin regulates Ampk-dependent autophagy and metabolic stress survival. *PLOS Genet.* **10**, e1004273 (2014). doi: 10.1371/journal.pgen.1004273; pmid: 24763318
63. C.-M. Hung *et al*., AMPK/ULK1-mediated phosphorylation of Parkin ACT domain mediates an early step in mitophagy. *Sci. Adv.* **7**, eabg4544 (2021). doi: 10.1073/pnas.1413021112; pmid: 25548157

64. J. J. Howell *et al*., Metformin Inhibits Hepatic mTORC1 Signaling via Dose-Dependent Mechanisms Involving AMPK and the TSC Complex. *Cell Metab.* **25**, 463–471 (2017). doi: 10.1172/JCI71749; pmid: 24762438
65. K. M. Harris *et al*., Uniform serial sectioning for transmission electron microscopy. *J. Neurosci.* **26**, 12101–12103 (2006). doi: 10.1523/JNEUROSCI.3994-06.2006; pmid: 17122034
66. D. R. Berger, H. S. Seung, J. W. Lichtman, VAST (Volume Annotation and Segmentation Tool): Efficient Manual and Semi-Automatic Labeling of Large 3D Image Stacks. *Front. Neural Circuits* **12**, 88 (2018). doi: 10.3389/fncir.2018.00088; pmid: 30386216
67. A. Dobin *et al*., STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013). doi: 10.1093/bioinformatics/bts635; pmid: 23104886
68. S. Heinz *et al*., Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010). doi: 10.1016/j.molcel.2010.05.004; pmid: 20513432
69. D. Torre, A. Lachmann, A. Ma'ayan, BioJupies: Automated Generation of Interactive Notebooks for RNA-Seq Data Analysis in the Cloud. *Cell Syst.* **7**, 556–561.e3 (2018). doi: 10.1016/j.cels.2018.10.007; pmid: 30447998
70. M. E. Ritchie *et al*., limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015). doi: 10.1093/nar/gkv007; pmid: 25605792
71. J. Goedhart, M. S. Luijsterburg, VolcaNoseR is a web app for creating, exploring, labeling and sharing volcano plots. *Sci. Rep.* **10**, 20560 (2020). doi: 10.1038/s41598-020-76603-3; pmid: 33239692
72. T. Metsalu, J. Vilo, ClustVis: A web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* **43**, W566–W570 (2015). doi: 10.1093/nar/gkv468; pmid: 25969447
73. S. Babicki *et al*., Heatmapper: Web-enabled heat mapping for all. *Nucleic Acids Res.* **44**, W147–W153 (2016). doi: 10.1093/nar/gkw419; pmid: 27190236
74. M. V. Kuleshov *et al*., Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016). doi: 10.1093/nar/gkw377; pmid: 27141961

ACKNOWLEDGMENTS

Funding: This study was supported by grants to R.J.S. from the National Institutes of Health (NIH) (R35CA220538, P01CA120964, and R01DK080425) and the Leona M. and Harry B. Helmsley Charitable Trust grant no. 2012-PG-MED002. The study received support to R.J.S. and G.S.S. by an AHA–Allen Initiative in Brain Health and Cognitive Impairment award made jointly through the American Heart Association and the Paul G. Allen Frontiers Group (19PABH134610000). The Salk NCI Cancer Center CCSG P30 CA014195 supported the Functional Genomics Core, the Bioinformatics Core, the Mass Spectrometry Core, the Flow Cytometry Core, and the Biophotonics Core. This research was also supported by the Nathan Shock Center for Aging Research at the Salk Institute P30 AG068635. U.M., L.F., R.G., A.S., and S.W.N. are supported by the Waitt Foundation and NCI CCSG P30 CA014195. U.M. is a Chan-Zuckerberg Initiative Imaging Scientist and is supported by NSF NeuroNex award 2014862 and NIH R21 DC018237. N.M. was supported by a Glenn Foundation Fellowship. We thank G. Kasof at Cell Signaling Technology for partnership in developing the pSer220 FNIP1 antibodies. We thank J. Moresco and J. Deitrich for their assistance in the Mass Spectrometry Core for Fig. 2B. All RNA-seq datasets in this study available through the Gene Expression Omnibus (GEO) under the accession no. GSE218827. **Author contributions:** N.M. and R.J.S. designed all the experiments and wrote the manuscript with input from all authors. N.M. executed all experiments except as stated here: B.I.F. originated the study of FNIP1 in the laboratory of R.J.S. and generated the CRISPR-mediated FNIP1 KO HEK293T cells and SA4 and other FNIP1 constructs. P.E.H. performed mass spectrometry to identify all phosphorylation sites in FNIP1 and generated SA5 constructs. N.M. analyzed RNA-seq datasets with involvement from J.Y. and N.H. S.D.C., E.T., and K.H. contributed key murine cell and tissue lysates for some panels in Fig. 2. L.F., R.G., and A.S. analyzed imaging data generated by S.W.N. and U.M. with supervision by U.M. for part of Fig. 6 and fig. S8. G.S.S. provided scientific insights during the study and aided in experimental design for part of Fig. 6 and fig. S8. **Competing interests:** All authors declare that they have no competing interests. **Data and materials availability:** All RNA-seq datasets generated for this study are available on GEO under the accession no. GSE218827. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj5559
Supplementary Text
Figs. S1 to S8
Table S1
References (*75*–*86*)
MDAR Reproducibility Checklist

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

Submitted 20 May 2021; resubmitted 27 June 2022
Accepted 22 March 2023
10.1126/science.abj5559

RESEARCH ARTICLE SUMMARY

CANCER

Oncogenic *CDK13* mutations impede nuclear RNA surveillance

Megan L. Insco*, Brian J. Abraham, Sara J. Dubbury, Ines H. Kaltheuner, Sofia Dust, Constance Wu, Kevin Y. Chen, David Liu, Stanislav Bellaousov, Anna M. Cox, Benjamin J. E. Martin, Tongwu Zhang, Calvin G. Ludwig, Tania Fabo, Rodsy Modhurima, Dakarai E. Esgdaille, Telmo Henriques, Kevin M. Brown, Stephen J. Chanock, Matthias Geyer, Karen Adelman, Phillip A. Sharp, Richard A. Young, Paul L. Boutz, Leonard I. Zon*

INTRODUCTION: Despite recent therapeutic advances, new treatments are needed for melanoma patients. Difficult-to-treat melanoma subsets rely on oncogenic gene expression for growth and therapy resistance. In normal cells, gene expression is tightly regulated by RNA surveillance pathways. Newly described mechanisms detect and degrade unnecessary RNA species in the nucleus to ensure gene expression quality.

RATIONALE: Transcriptional cyclin-dependent kinases (CDKs) are a family of kinases that have roles in directly controlling gene expression. We examined the transcriptional CDK loci in publicly available melanoma patient data to investigate whether any of these kinases could be a therapeutic target for melanoma. Unexpectedly, we observed an enrichment of *CDK13* kinase domain mutations in melanoma, suggesting that CDK13 is a dominant-negative tumor suppressor. Therefore, we investigated the mechanism of mutant CDK13-mediated oncogenesis.

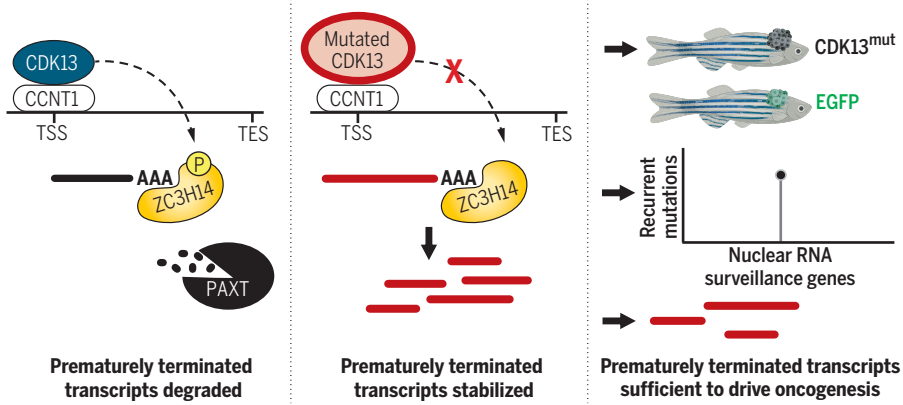
RESULTS: *CDK13* is mutated in 3.9% of cutaneous melanomas, and these mutations are selected for as measured by a computational tool that considers mutational load and severity. Kinase domain mutations were enriched compared with mutations in the rest of the protein in melanoma (2.2-fold) and other cancers (1.8-fold), suggesting dominant-negative activity. The *CDK13* kinase domain mutations largely overlap with mutations that cause a *CDK13*-related developmental disorder. Expression of kinase-mutated CDK13 expedited melanoma onset in a zebrafish melanoma model and caused human melanoma cells to be more proliferative, demonstrating *CDK13*'s dominant-negative tumor-suppressive function.

Because CDK13 is related to known transcriptional kinases, we tested for a global gene expression phenotype by quantifying differential exon usage. We found that CDK13 mutant expression or loss of function resulted in increased expression of first compared with last exons, indicating the accumulation of short RNAs. Using a specialized sequencing technique, we defined the 3′ end of RNAs, wh[illegible] showed that mutant CDK13 zebrafish m[illegible] nomas and human melanoma cells had increased prematurely terminated RNAs (ptRNAs) ending in introns. We found that ptRNAs accumulate posttranscriptionally through lack of degradation using digital droplet PCR and nascent RNA sequencing. Because ptRNAs are capped, spliced, and polyadenylated, we performed whole-cell proteomics of mutant CDK13 versus control zebrafish melanomas and found translation of some ptRNAs, including their intronic regions, which could be a source of neoantigens.

To elucidate the mechanism of mutant CDK13 oncogenesis, we immunoprecipitated CDK13 and discovered binding to proteins associated with the polyA tail exosome targeting (PAXT) complex, which targets ptRNAs for degradation in the nucleus. We immunoprecipitated ZC3H14 from $CDK13^{WT}$ and $CDK13^{mut}$ human melanoma cells and found that ZC3H14 lacks phosphorylation at S475 in $CDK13^{mut}$ cells, suggesting that CDK13 directly phosphorylates ZC3H14 S475. We rescued PAXT recruitment and activity with expression of a ZC3H14 phosphomimetic mutant in $CDK13^{mut}$ cells. We also showed that the expression of a nonphosphorylatable ZC3H14 decreased PAXT recruitment and activation in $CDK13^{WT}$ cells. CDK13 activated by CCNT1 was able to in vitro phosphorylate $ZC3H14^{WT}$, but not $ZC3H14^{S475A}$, showing that CDK13 can directly phosphorylate ZC3H14 on the relevant residue.

We found that ptRNAs accumulated in other $CDK13^{mut}$ cancers. To determine whether ptRNA expression is sufficient to expedite oncogenesis, we expressed two human ptRNAs in the zebrafish model and found that they both expedited melanoma onset. Finally, we found that additional nuclear RNA surveillance components are recurrently mutated in cancer. These experiments show that ptRNAs can promote cancer phenotypes and that components surveilling aberrant RNAs are mutated in cancer.

CONCLUSION: Our work shows that CDK13 has properties consistent with a tumor suppressor, and that mutant CDK13 is oncogenic because of deficient RNA surveillance. Our finding that recurrent mutations occur in additional PAXT components suggests a broad, previously unrecognized tumor-suppressive role for nuclear RNA surveillance. ■



Deficient nuclear RNA surveillance is oncogenic. CDK13 phosphorylates ZC3H14 to activate nuclear RNA surveillance on protein-coding genes. Mutant CDK13 fails to phosphorylate ZC3H14, and aberrant RNAs are stabilized. CDK13 has properties consistent with a tumor suppressor, as shown from patient data and zebrafish melanoma models. Recurrent mutations in nuclear RNA surveillance genes were identified in cancer. The expression of ptRNAs is sufficient to drive melanomagenesis.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: leonard.zon@enders.tch.harvard.edu (L.I.Z.); megan_insco@dfci.harvard.edu (M.L.I.)
Cite this article as M. L. Insco *et al.*, *Science* **380**, eabn7625 (2023). DOI: 10.1126/science.abn7625

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abn7625

RESEARCH ARTICLE

CANCER

Oncogenic *CDK13* mutations impede nuclear RNA surveillance

Megan L. Insco[1,2]*, Brian J. Abraham[3], Sara J. Dubbury[4,5], Ines H. Kaltheuner[6], Sofia Dust[6], Constance Wu[1], Kevin Y. Chen[1], David Liu[2], Stanislav Bellaousov[7], Anna M. Cox[1], Benjamin J. E. Martin[8], Tongwu Zhang[9], Calvin G. Ludwig[1], Tania Fabo[1], Rodsy Modhurima[1], Dakarai E. Esgdaille[7], Telmo Henriques[8], Kevin M. Brown[9], Stephen J. Chanock[9], Matthias Geyer[6], Karen Adelman[8], Phillip A. Sharp[4,5], Richard A. Young[5,10], Paul L. Boutz[7,11,12], Leonard I. Zon[1]*

RNA surveillance pathways detect and degrade defective transcripts to ensure RNA fidelity. We found that disrupted nuclear RNA surveillance is oncogenic. Cyclin-dependent kinase 13 (*CDK13*) is mutated in melanoma, and patient-mutated *CDK13* accelerates zebrafish melanoma. *CDK13* mutation causes aberrant RNA stabilization. CDK13 is required for ZC3H14 phosphorylation, which is necessary and sufficient to promote nuclear RNA degradation. Mutant CDK13 fails to activate nuclear RNA surveillance, causing aberrant protein-coding transcripts to be stabilized and translated. Forced aberrant RNA expression accelerates melanoma in zebrafish. We found recurrent mutations in genes encoding nuclear RNA surveillance components in many malignancies, establishing nuclear RNA surveillance as a tumor-suppressive pathway. Activating nuclear RNA surveillance is crucial to avoid accumulation of aberrant RNAs and their ensuing consequences in development and disease.

Transcriptional cyclin-dependent kinases (CDKs) are a family of kinases activated by a cyclin-binding partner (*1*, *2*) that have roles in controlling transcriptional subprocesses including initiation, elongation, co-transcriptional RNA processing, and termination (*3*–*9*). Transcriptional CDKs, including CDK13, are being investigated as therapeutic targets for many difficult-to-treat cancers (*10*–*14*). Mutations affecting CDK13 cause syndromic developmental disorders that affect neural crest–derived tissues (congenital heart defects, dysmorphic facial features, and intellectual development) (*15*–*19*). The role of CDK13 in transcription and RNA processing remains poorly understood.

[1]Stem Cell Program and Division of Hematology/Oncology, Boston Children's Hospital, Howard Hughes Medical Institute, Boston, MA 02115, USA. [2]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA. [3]Department of Computational Biology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA. [4]The David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [5]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [6]Institute of Structural Biology, University of Bonn, 53127 Bonn, Germany. [7]University of Rochester School of Medicine and Dentistry, Rochester, NY 14642, USA. [8]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA. [9]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD 20850, USA. [10]Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology, Cambridge, MA 02142, USA. [11]Center for RNA Biology, University of Rochester, Rochester, NY 14642, USA. [12]Center for Biomedical Informatics, University of Rochester, Rochester, NY 14642, USA.
*Corresponding author. Email: leonard.zon@enders.tch.harvard.edu (L.I.Z.); megan_insco@dfci.harvard.edu (M.L.I.)

RNA polymerase II (RNAPII) is exquisitely regulated to ensure that transcriptional initiation, elongation, co-transcriptional processing, and termination are precisely executed on protein-coding genes. During elongation, a rapid cascade of RNA-processing steps takes place (*20*). RNA-processing errors can result in aberrant transcripts that must be degraded by RNA surveillance pathways. Messages that prematurely terminate are predicted to evade nonsense-mediated decay because of the lack of downstream splice events (*21*), and require a complementary surveillance pathway to avoid accumulation and translation of pathogenic transcripts.

Aberrant or unstable nuclear RNAs are recognized by adaptor complexes, which recruit the MTR4 helicase and the nuclear RNA exosome degradation complex to specific RNAs. The recently reported polyA exosome targeting (PAXT) complex targets prematurely terminated RNAs (ptRNAs) for degradation (*22*, *23*). How ptRNAs (and not mRNAs) are specifically targeted for PAXT degradation is unknown.

In this study, we found that patient-specific mutations in CDK13 accelerate oncogenesis in human cells and in a zebrafish melanoma model, and that CDK13 is responsible for activating nuclear RNA surveillance. When CDK13 is mutated or lost, phosphorylation of the PAXT component ZC3H14 is compromised, and ptRNAs fail to be targeted for nuclear degradation. Consequently, stabilized ptRNAs are translated into truncated proteins. Overexpression of ptRNAs leads to accelerated melanoma in zebrafish. We found that PAXT subunits recruited by CDK13 phosphorylation are mutated in 17% of melanomas, and that recurrent mutations in PAXT subunits ZFC3H1 and ZC3H18 occur in many cancer types. This work demonstrates that CDK13 normally activates nuclear RNA surveillance machinery to degrade oncogenic ptRNAs.

Results

CDK13 has properties consistent with a tumor suppressor

We initially examined the transcriptional CDK loci as defined by (*2*) in melanoma patient samples from The Cancer Genome Atlas (TCGA) (*24*) and found that CDK13 was mutated in 4.6% of patients and had a cluster of mutations near the ATP-binding site (table S1). We then analyzed a larger set of publicly available patient data (*24*–*33*) specifically for CDK13 mutations and found that *CDK13* is somatically mutated in 3.9% of cutaneous melanomas (52 of 1347 patients) (fig. S1A and table S2). To determine whether *CDK13* somatic mutations are selected for in melanoma, OncodriveFM (*34*) was used because it accounts for background mutational load and predicted mutational severity. This analysis suggested that *CDK13* is a significant melanoma driver gene ($q = 0.042$, $P = 0.0049$). We observed 2.2-fold enrichment of deleterious mutations in the kinase domain versus the remainder of the protein (Fisher's exact test, $P = 0.03$, odds ratio = 5.4) (Fig. 1A, red and orange; fig. S1A; table S2; and data S1). To assess *CDK13* mutational clonality, we identified samples that were copy neutral for *CDK13* and compared the variant allelic fraction of *CDK13* mutations against known driver mutations from the same samples (*24*). Of six evaluable tumors, four were clonal (P746T, R860Q, E1248K, and P881L), one was subclonal (R1206M), and one was potentially subclonal (E1448X) with a variant allelic fraction similar to *NF1* (fig. S1B and data S1). In non-copy-neutral *CDK13* mutant samples, we observed additional samples harboring clonal or potentially clonal *CDK13* mutations (9 of 12). Of the seven samples with clonal *CDK13* mutations, all were heterozygous. *CDK13* is mutated in many other malignancies (fig. S1C), which also display an enrichment in kinase domain mutations (1.8×, $n = 13269$ from 56 studies, $P = 2.47 \times 10^{-13}$) (data S1). Loss of heterozygosity is infrequently observed in melanoma. Of the 1347 melanoma patient samples, we assessed those with copy number and mutational information for loss of heterozygosity and found that 0.5% had biallelic loss ($n = 832$, one biallelic deletion, and three with a mutation and heterozygous deletion). These data show that *CDK13* is a mutated driver gene with a modest enrichment in kinase domain mutations, suggesting that kinase domain mutations have additional selection

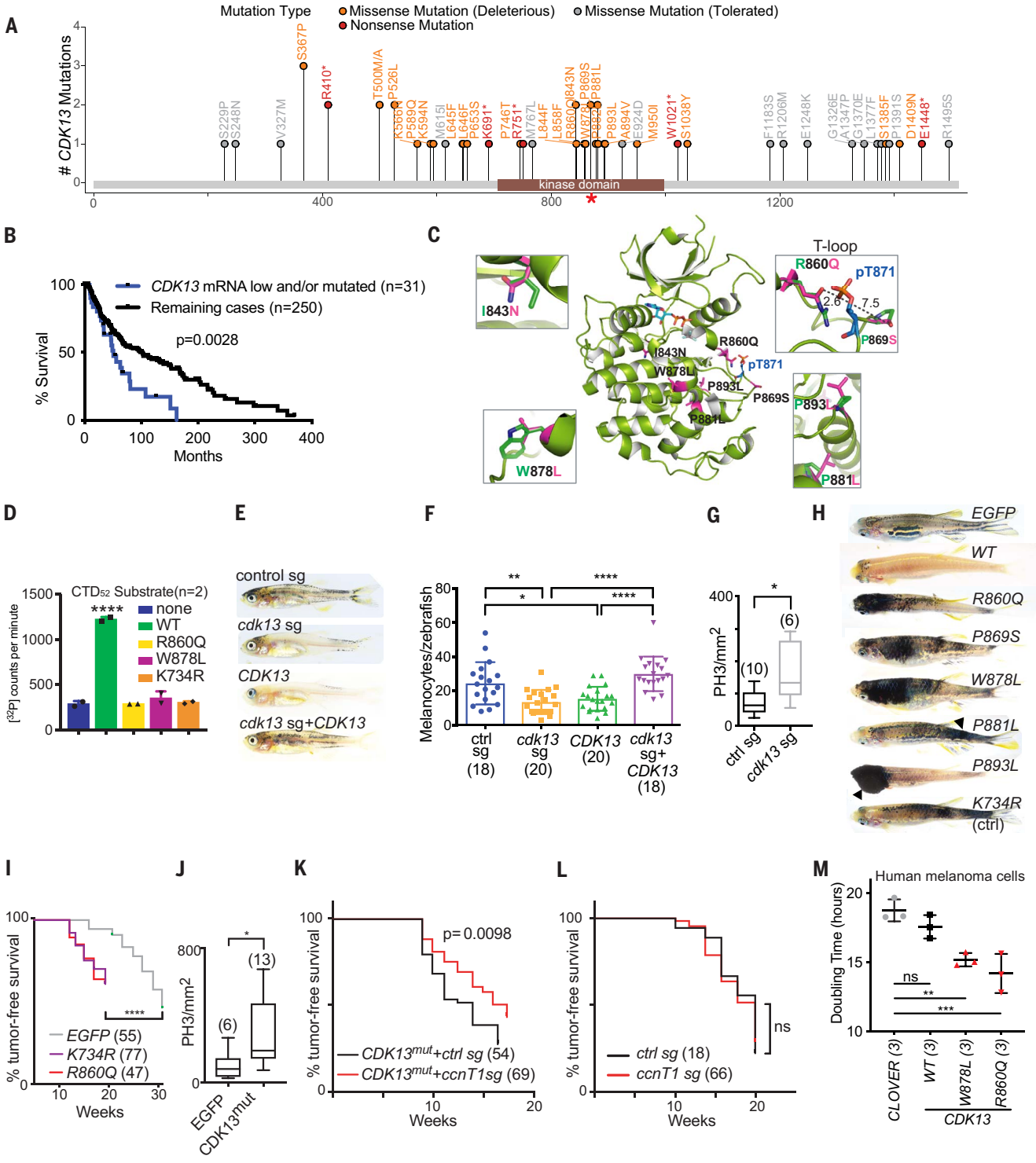


Fig. 1. CDK13 has properties consistent with a tumor suppressor. (**A**) *CDK13* melanoma mutation plot. Red asterisk indicates the phosphorylation site. (**B**) Patient survival plot with *CDK13* mutation or down-regulation versus remaining patients. $P = 0.0028$, log-rank test. n = patients. (**C**) Patient kinase-domain mutations mapped on the CDK13 crystal structure. pT871 is the phosphorylation site. (**D**) In vitro kinase assay of wild-type (WT) and patient-mutated CDK13 using full-length CTD_{52} as the substrate. One-way ANOVA with no kinase versus all conditions was used. WT CDK13, $^{****}q = 0.0001$; all other comparisons were nonsignificant. Data are shown as means ± SD. (**E**) Representative photos of *mitfa:BRAF*^{V600E}*;p53*$^{-/-}$*;mitfa*$^{-/-}$ (Triples melanoma model zebrafish) with control guide RNA (gRNA), *cdk13* gRNA, human CDK13^{WT}, or *cdk13* gRNA and human CDK13^{WT} at 4 weeks after fertilization. (**F**) Quantification of melanocytes at 3 days after fertilization for Triples zebrafish injected with control gRNA, *cdk13* gRNA, human CDK13^{WT}, or *cdk13* gRNA and human CDK13^{WT}. One-way ANOVA, multiple comparisons. Data are shown as means ± SD. $^{*}q = 0.0186$, $^{**}q = 0.0030$, $^{****}q < 0.0001$. n = number of zebrafish. (**G**) PH3 antibody staining per square millimeter of melanoma from Triples zebrafish injected with *cdk13* gRNA compared with control gRNA. $P = 0.014$ (Mann-Whitney two-tailed t test). n = number of melanomas. (**H**) Nine-week photos of melanocyte-specific expression of EGFP, human CDK13^{WT},

patient-mutant CDK13, or CDK13^{K734R} (catalytically dead control) in Triples zebrafish. Arrows indicate melanomas. (**I**) Percentage melanoma-free survival of Triples zebrafish with melanocyte-specific expression of EGFP, CDK13^{R860Q} (patient mutation), and CDK13^{K734R} (catalytically dead). ****P < 0.0001 (log-rank). n = number of zebrafish. (**J**) PH3 antibody staining/square millimeter of melanoma from Triples zebrafish expressing EGFP versus CDK13 mutant (W878L or P893L). P = 0.0125 (Mann-Whitney test, two-tailed). n = number of melanomas. (**K**) Melanoma-free survival of zebrafish with melanocyte-specific expression of CDK13^{W878L} with melanocyte-specific CRISPR of either a control gene or *ccnT1*. P = 0.0098 (log-rank). n = number of zebrafish. (**L**) Melanoma-free survival of zebrafish with melanocyte-specific CRISPR of a control gene or *ccnT1* alone. ns, nonsignificant (log-rank). n = number of zebrafish. (**M**) Doubling time for human melanoma cells expressing CLOVER, CDK13^{WT}, CDK13^{W878L}, or CDK13^{R860Q}. One-way ANOVA with CLOVER versus all conditions. CDK13^{W878L}, **q = 0.0044; CDK13^{R860Q}, ***q = 0.0009. ns, nonsignificant. Data are shown as means ± SD. n = 3 biologic replicates. For all box plots, the box is the 25th to 75th percentile range, the solid line represents the median, and whiskers show maximum to mininum.

pressure beyond loss of function, such as dominant-negative activity.

De novo heterozygous CDK13 kinase-domain mutation has been reported to cause a CDK13-associated developmental syndrome (80%, 32 of 40 patients) (*15*–*19*), suggesting the functional importance of these mutations. All but two of the CDK13-developmental kinase domain variants are found as somatic mutations in cancer. The notable difference is the N842 variant, the most common variant associated with the developmental disorder, which was not found as a somatic mutation in melanoma or nonmelanoma cancer (data S1). Although *CDK13* was not considered a driver gene in prior studies because of its low mutational rate, these data show that *CDK13* is a broadly used driver gene with an enrichment in deleterious kinase-domain mutations, suggesting a function beyond haploinsufficiency.

If mutant CDK13 functions by interfering with WT CDK13, in other words has dominant-negative activity, then mutant CDK13 would be expected to result in a phenotype similar to *CDK13* down-regulation. Survival analysis using human melanoma patient (*24*) data from TCGA revealed that patients with *CDK13* down-regulation (z score ≤–1.0, fold change 0.48) or somatic mutation (*CDK13* altered) combined had decreased overall survival (Fig. 1B and fig. S1, D and E; log-rank P = 0.0028). *CDK13* alteration remained associated with poor survival after adjustment in a multivariate model (homologous recombination repair = 1.81; 95% confidence interval, 1.13 to 2.90; Cox P = 0.014; table S3). Patients initially staged with 0/1/2 melanoma with *CDK13* alterations exhibited reduced overall survival compared with remaining stage 0/1/2 melanoma patients (fig. S1F; log-rank P = 0.0012). The percentages of *CDK13*-altered (mutated or down-regulated) melanomas with *BRAF*^{V600} mutations or *NRAS*^{Q61} were compared with the remaining cases in the TCGA melanoma cohort (*24*), and no significant differences were found (chi-square test, P = 0.287) (Table 1). These data suggest that *CDK13* mutation or down-regulation is associated with poor overall survival in melanoma patients.

CDK13 melanoma mutations that occur in the kinase domain were projected onto a previously solved crystal structure (Fig. 1C) (*35*). The P869 residue is located two residues away from the T871 phosphorylation site, and the R860 residue normally coordinates with phospho-T871, so these mutations could disrupt kinase activation. The W878 mutation may change the substrate-binding pocket, and mutants P881L, P893L, and I843N might disrupt the integrity of the kinase domain structure. In vitro, CDK13^{WT} activated by its canonical cyclin partner, cyclin K (CCNK), had robust kinase activity, whereas the R860Q, W878L, and K734R mutants failed to phosphorylate full-length RNAPII C-terminal domain (CTD_{52}) and a second substrate (Fig. 1D and fig. S1G), but maintained cyclin binding (fig. S1G, right). The CDK13^{K734R} mutation replaces the lysine required for catalysis (*36*) and occurs in the CDK13-associated developmental syndrome (*15*). These data show that CDK13 mutations observed in melanoma abrogate kinase activity.

To examine the role of CDK13 in melanoma, the MAZERATI rapid modeling system was used with *mitfa:BRAF*^{V600E}*;p53*$^{-/-}$*;mitfa*$^{-/-}$ zebrafish, hereafter referred to as the "Triples" zebrafish melanoma model. These zebrafish lack Mitfa, the master regulator of melanocyte development, and also lack melanocytes (*37*–*39*). Melanocytes were rescued with a vector that expresses Mitfa, allowing cell-autonomous melanocyte-specific gene expression or CRISPR/Cas9-mediated inactivation using the *mitfa* promoter. Zebrafish with melanocyte-specific CRISPR-deletion of zebrafish *cdk13* compared with that of a control gene showed significantly decreased melanocyte numbers (one-way ANOVA q = 0.0030), indicating that *cdk13* is required for melanocyte development (Fig. 1, E and F, and fig. S1, H and I). Overexpression of human wild-type CDK13 (CDK13^{WT}) also resulted in fewer melanocytes during development (Fig. 1, E and F, and fig. S1J) but was able to rescue melanocyte number loss caused by *cdk13* deletion (Fig. 1, E and F, and fig. S1K). Rarely, melanomas arose in zebrafish with *cdk13* melanocyte-specific CRISPR-deletion, and phospho-histone 3 (PH3) immunohistochemical staining revealed that *cdk13* CRISPR-deleted versus control melanomas were significantly more proliferative (Mann-Whitney two-tailed t test, P = 0.014) (Fig. 1G and fig. S1, L and M). Proper Cdk13 levels are required for melanocyte development and human CDK13 complements the loss of zebrafish *cdk13*.

To test whether CDK13 melanoma patient mutations cause more aggressive melanoma in vivo, mutant CDK13 was expressed in melanocytes in the Triples zebrafish melanoma model. Control enhanced green fluorescent protein (EGFP) expression resulted in expected rescue of mosaic stripes, whereas human CDK13^{WT} expression resulted in few melanocytes at 9 weeks after fertilization, so CDK13^{WT}-expressing animals were not followed for tumor onset. By contrast, overexpression of CDK13^{R860Q}, CDK13^{P869S}, CDK13^{W878L}, CDK13^{P881L}, or CDK13^{P893L} caused the appearance of black patches at 9 weeks after fertilization (Fig. 1H and fig. S1N) and expedited tumor onset (Fig. 1I and fig. S1, O and P). Because all tested *CDK13* mutations promoted melanoma to a similar degree, we have used them interchangeably in most assays (hereafter called *CDK13*^{mut}). CDK13^{mut}-expressing melanomas had more PH3-positive cells, as shown by immunohistochemistry, than controls (Fig. 1J and fig. S1Q). Because both melanocyte-specific *cdk13* CRISPR deletion and forced human CDK13^{mut} expression caused more proliferative melanomas, our data support that CDK13^{mut} acts through a dominant-negative or antimorphic mechanism in which human

Table 1. *CDK13* alterations occur in all genetic subtypes in melanoma TCGA patients.

	CDK13* mutated and down-regulated	**All other cases†**
BRAF V600 mutations	36.4% (n = 5 mutated, n = 7 down-regulated)	50.8% (n = 129)
NRAS Q61 mutations	36.4% (n = 5 mutated, n = 7 down-regulated)	26.8% (n = 68)
Other	27.3% (n = 2 mutated, n = 7 down-regulated)	22.4% (n = 57)

*Percentage of cases in N = 33 of 287. †Percentage of cases in N = 254 of 287 patients with DNA mutational and RNA expression data available.

CDK13^{mut} adversely affects intact zebrafish Cdk13^{WT} activity. Our data are consistent with existing human developmental disorder genetics showing that CDK13 mutations work through a dominant-negative mechanism (*15*–*19*).

We hypothesized that cyclin binding is required for CDK13^{mut} dominant-negative activity. To test whether CCNK (*35*, *40*–*42*) or cyclin T1 (CCNT1) are required for CDK13^{mut} dominant-negative melanomagenesis in vivo, we coinjected vectors that express melanocyte-specific CDK13^{mut} and that delete *ccnK* or *ccnT1* in the Triples zebrafish melanoma model. *ccnK* melanocyte-specific inactivation expedited melanoma in the presence of CDK13^{W878L}, but not alone (fig. S1, R and S). *ccnT1* melanocyte-specific inactivation suppressed CDK13^{W878L} and CDK13^{R860Q} oncogenesis but had no effect alone (Fig. 1, K and L, and fig. S1, T and U). These data indicate that the oncogenic mechanism of CDK13^{mut} requires CCNT1.

To determine the effect of mutant CDK13 expression in human melanoma cells, CLOVER fluorescent protein, CDK13^{WT}, CDK13^{W878L}, or CDK13^{R860Q} was expressed in human A375 melanoma cells. Cells expressing either CDK13 mutant had a decreased doubling time (increased growth rate) compared with CDK13^{WT} or CLOVER (Fig. 1M and fig. S1V). These data show that mutant CDK13 expression causes human melanoma cells to proliferate more quickly, indicating that mutant CDK13's proproliferative function is conserved.

CDK13 mutation results in accumulation of RNAs that prematurely terminate in introns

Because CDK13 is phylogenetically related to known transcriptional kinases, we tested for a global RNA phenotype by quantifying differential exon usage. Differential expression analysis of first (F), alternative first (AF), internal (I), last (L), or alternative last (AL) exons in each gene was performed in zebrafish melanomas. We observed significantly increased read coverage in the first exon compared with the last exon in CDK13^{mut}-expressing versus control melanomas (two-sided Wilcoxon rank-sum test $P < 2.2 \times 10^{-16}$) (Fig. 2A and fig. S2A). These data suggest that ptRNAs are present in CDK13^{mut} zebrafish melanomas.

Analysis of poly-A–selected RNA-sequencing (RNA-seq) from patient melanomas (patient characteristics in table S4) showed that patients with CDK13^{mut}-expression versus CDK13^{WT} expression had increased read coverage in the first exon versus the last exon (Fig. 2B), indicating the presence of ptRNAs. ptRNA evidence was of greater magnitude in the patient RNA-seq data that were poly-A selected compared with non–polyA selected, indicating that ptRNAs are polyadenylated (compare Fig. 2, B versus A). We generated mouse embryonic stem cells (mESCs) that were *Cdk13*$^{-/-}$ and carried a doxycycline-inducible mouse *Cdk13* rescue transgene. PolyA-selected RNA-seq was done at time zero (CDK13 expressed) and at 48 and 72 hours after doxycycline removal (CDK13 depleted) (fig. S2, B to D). mESCs that were CDK13 depleted for 48 hours showed evidence of accumulated ptRNAs, with this effect intensifying at 72 hours after depletion (Fig. 2C). Accumulated ptRNAs in CDK13-depleted mESCs were largely nonoverlapping from CDK12-depleted mESCs generated analogously (*7*) (fig. S2E and data S1). These data show that CDK13's role in preventing the accumulation of polyadenylated ptRNAs is conserved.

3′ RNA-seq was used to determine where ptRNAs terminate in CDK13^{mut} melanomas. 3′ sequencing was executed on Triples zebrafish melanomas with melanocyte-specific expression of EGFP ($n = 3$) or mutant CDK13^{mut} ($n = 3$) and human melanoma A375 cells expressing CLOVER ($n = 2$) or CDK13^{mut} ($n = 2$). A total of 83,660 overall recurrent termination events were identified in zebrafish melanomas and 47,787 overall termination events were identified in human melanoma cells. We used DEXSeq (*43*) to quantify differential usage of polyA termination sites between CDK13^{mut} and eGFP condition in zebrafish melanomas, which identified 802 significantly different termination events ($q < 0.05$). Parallel analysis from human melanoma cells revealed 1678 significantly different termination events ($q < 0.05$) in CDK13^{mut} compared with control CLOVER cells. To determine whether conserved genes or pathways were affected in zebrafish and human melanomas, the affected zebrafish genes were transferred to the nearest ortholog and the gene lists were overlapped. Only 3.4% of genes were found in both lists, and gene ontology analysis showed no enriched pathways using the "biologic process complete" annotation set. Given the lack of shared functional pathways, these data suggest that the proliferative phenotype observed in CDK13 mutant patient melanomas, zebrafish melanomas, and human melanoma cells may be promoted by cellular stress from accumulated truncated RNAs, as opposed to truncation of a specific conserved target gene(s).

A significant increase in intronic cleavage sites was observed in CDK13^{mut} zebrafish melanomas (Fig. 2D) and human melanoma cells (Fig. 2E). In zebrafish melanomas, the significantly changed sites were enriched in introns (median fold change 2.54) and depleted from untranslated regions (UTRs) (median fold change −1.51) ($P < 2.2 \times 10^{-16}$, two-sided Kolmogorov-Smirnov test); and in human melanoma cells, the significantly changed sites were also enriched in introns (median fold change 0.38) and depleted from UTRs (median fold change −1.67) ($P = 9.99 \times 10^{-16}$, two-sided Kolmogorov-Smirnov test). In CDK13^{mut} human melanoma cells, *SUV39H1* and *CBFB* each showed an up-regulated ptRNA that terminated in the third intron (Fig. 2F and fig. S2F), whereas *TP53* had an up-regulated ptRNA that terminated in the first intron (fig. S2G). For many up-regulated intronic polyadenylation sites, a canonical RNA polyadenylation site motif was observed (Fig. 2F and fig. S2, F and G). These data show that CDK13^{mut} causes accumulation of ptRNAs generated from existing cleavage and polyadenylation sites found in intronic DNA in zebrafish and humans.

ptRNAs could accumulate from increased production or decreased clearance in CDK13 mutant cells. If transcription elongation is impaired, then ptRNAs should be produced at a higher rate at the expense of full-length mRNA synthesis. In this case, the absolute amount of first exon expression in CDK13 mutant cells should be unchanged compared with control, whereas the last exon expression should be decreased (Fig. 2G, model 1). Alternatively, if ptRNA clearance is decreased in CDK13^{mut} cells, then ptRNA concentration should rise without affecting full-length mRNA expression. The first exon expression should increase, whereas the last exon expression should remain stable (Fig. 2G, model 2). We hypothesized that absolute quantification of RNA species could help to distinguish between these two models.

To measure concentration of ptRNAs and full-length mRNAs, digital droplet polymerase chain reaction (ddPCR) was performed for four genes with up-regulated ptRNAs and two control genes, which were identified from the 3′ sequencing described above. The signal from first exon probes represented both ptRNAs and full-length mRNAs, whereas the signal from the last exon probes represented full-length mRNAs alone (Fig. 2H and fig. S2H). The fold change between CDK13^{mut} and control cells showed that all genes with predicted ptRNAs had an increase in first exon expression, consistent with ptRNA accumulation in CDK13^{mut} cells (*CBFB*, *LATS2*, *SUV39H1*, and *TP53*). Neither control gene showed evidence of accumulated ptRNAs (*C1orf35* and *AGPAT1*). All six genes showed unchanged last exon expression in CDK13^{mut} cells compared with control cells (Fig. 2H). These data verify increased ptRNA levels in human CDK13^{mut} human melanoma cells. Last exon expression was intact for all genes tested, showing that transcription is intact and suggesting that ptRNAs accumulate in CDK13^{mut} cells through lack of clearance (Fig. 2G, model 2).

ptRNAs accumulate posttranscriptionally and are translated, including intronic sequences

To directly examine transcription in CDK13^{mut}-expressing human melanoma cells, genome-wide nascent RNA production was measured with transient transcriptome sequencing (TT-seq)

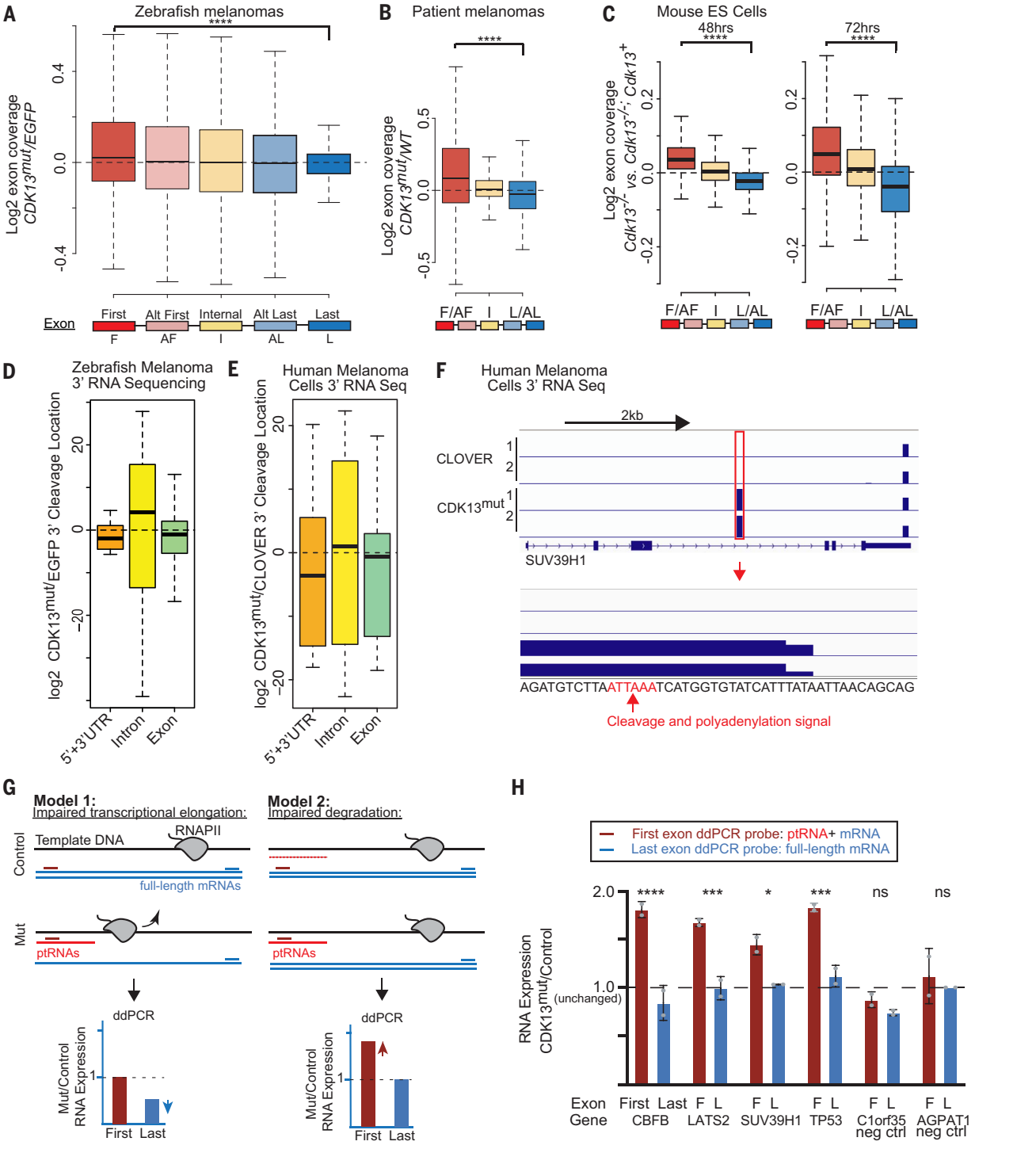


Fig. 2. CDK13 mutation results in accumulation of RNAs that prematurely terminate in introns. (**A**) Log2-fold normalized exon expression in CDK13^{R860Q}-expressing (n = 5) versus EGFP-expressing (n = 4) zebrafish melanomas. ****$P < 2.2 \times 10^{-16}$, first versus last exon. (**B**) Log2-fold normalized exon expression for CDK13^{mut} (n = 3) versus CDK13^{WT} (n = 5) matched control patient melanomas. ****$P < 2.2 \times 10^{-16}$, F/AF versus L/AL exon. (**C**) Log2-fold normalized exon expression in Cdk13$^{-/-}$ versus Cdk13$^{-/-}$;Cdk13+ (control) mESCs. 48 and 72 hours, hours of Cdk13 depletion (n = 4 for each). ****$P < 2.2 \times 10^{-16}$, F/AF versus L/AL exon. (**D** and **E**) Log2-fold CDK13^{mut}/control cleavage site utilization (3′ seq) in UTRs, introns, and exons in zebrafish melanomas (D) and human melanoma cells (E). (**F**) Integrative Genomics Viewer (IGV) plot of 3′ sequencing showing the increased *SUV39H1* ptRNA in CDK13^{mut}-expressing versus control human melanoma cells. Red box indicates the ptRNA 3′ end. (**G**) Models for ptRNA accumulation in CDK13^{mut} cells. (**H**) RNA expression in CDK13^{mut} and control human melanoma cells measured by digital droplet PCR for four genes with increased ptRNAs by 3′ sequencing and two control genes. F, first exon; L, last exon. Two-way ANOVA, adjusted P values are shown from left to right: ****$P < 0.0001$; ***$P = 0.0006$; *$P = 0.0280$; ***$P = 0.0004$; ns, 0.8626; and ns, 0.9158 ± SD. In box plots, the black horizontal line indicates the median and whiskers extend to 1.5× the interquartile range (IQR). P values in (A) to (C) are from two-sided Wilcoxon rank-sum tests.

alongside paired standard RNA-seq (Fig. 3A, TT-seq schematic) (*44*). If ptRNAs are generated through impaired transcriptional elongation, then there should be less 4-thiouridine (4sU) incorporation at the 3′ ends of genes (Fig. 3A, model 1). If ptRNAs accumulate through a posttranscriptional process such as impaired degradation, then 4sU incorporation should be intact at the 3′ ends of genes (Fig. 3A, model 2). RNA-seq differential exon analysis confirmed increased 5′ coverage in CDK13^{mut} cells consistent with accumulated ptRNAs (fig. S3A). TT-seq metagenes showed no evidence of an increase in premature termination by RNAPII; instead, we observed a slight increase in RNA synthesis across the gene body in CDK13^{mut} cells (Fig. 3B). In both TT-seq and RNA-seq, the metagene profiles indicate intact cleavage and polyadenylation in CDK13^{mut} and CLOVER control cells (fig. S3, B and C). Nascent RNA production upstream of the cleavage and polyadenylation site is intact in CDK13^{mut}-expressing human melanoma cells further demonstrating that elongation is functional (fig. S3B). These data show that CDK13^{mut} cells maintain productive RNAPII elongation and that ptRNAs accumulate through a posttranscriptional mechanism such as loss of ptRNA degradation.

To measure production and steady-state levels of ptRNAs at intronic polyadenylation (IPA) sites in CDK13^{mut} cells, TT-seq and RNA-seq read coverage were quantified across IPA sites regulated by CDK13 as measured from 3′ seq from Fig. 2E ($q < 0.1$). To allow pure intronic read quantification, IPA sites were included if they were >400 base pairs from the nearest exon ($n = 134$). The upstream/downstream coverage ratio for IPA sites from CDK13^{mut} versus control human melanoma cells was compared with the median ratio in the RNA-seq and the TT-seq (Fig. 3C). The upstream/downstream coverage ratio in the TT-seq (nascent RNA) was unaffected in CDK13^{mut} cells (Fig. 3D, TT-seq, left), showing that there is similar low-level intronic transcriptional termination. By contrast, the RNA-seq upstream/downstream coverage ratio was significantly increased in CDK13^{mut} cells ($P = 0.01$) (Fig. 3D, RNA-seq, right; and fig. S3, D and E), reflecting higher steady-state RNA levels upstream of IPA sites. These data show that CDK13^{mut} cells stabilize ptRNAs posttranscriptionally.

To further investigate the transcriptional effects of chronic CDK13^{mut} expression, we performed chromatin immunoprecipitation (IP) sequencing with antibodies to RNAPII (8WG hypophosphorylated and Ser2 CTD) in CDK13^{mut} and control human melanoma cells and zebrafish melanomas. In both systems, increased RNAPII occupancy was observed in gene bodies in the CDK13^{mut} condition (fig. S3, F to J). These data support the ddPCR and TT-seq results showing that transcriptional elongation is intact and that ptRNA accumulation occurs in CDK13^{mut} through a highly conserved posttranscriptional mechanism.

ptRNAs have been reported to be exported (*23*, *45*) and are predicted to be translated (*21*). We used tandem mass spectrometry (MS) to analyze global protein expression in melanomas isolated from age-matched Triples zebrafish with melanocyte-specific expression of CDK13^{mut} ($n = 3$) or EGFP ($n = 3$). Differential protein expression analysis revealed that 174 proteins were up-regulated, including 32 proteins related to vesicle-mediated transport with many implicated in lysosomes/autophagy (biologic complete, $q = 2.3 \times 10^{-13}$) (fig. S3K). CDK13^{mut} melanomas have increased lysosomal and autophagic protein expression, suggesting that CDK13^{mut} cells have proteomic stress that could be caused by translation of truncated RNAs into truncated proteins.

If ptRNAs are translated, then there should be increased peptide levels at the beginning of proteins and fewer at the end of proteins (Fig. 3E). To test whether ptRNAs are translated, individual peptide measurements filtered for changes between CDK13^{mut} and control conditions were analyzed ($P < 0.1$). Log2-fold change CDK13^{mut} versus EGFP control peptide measurements were binned and plotted as a percentage of canonical protein sequence length. This analysis revealed that early (N-terminal) protein peptide measurements were increased in CDK13^{mut} melanomas (Fig. 3F). The negative slope is consistent with an increase in short proteins arising from the translation of prematurely terminated transcripts, analogous to increased 5′ coverage in RNA-seq seen in CDK13^{mut} melanomas. This analysis identified 263 proteins with evidence of truncation. Any protein with more than three peptide measurements was tested for evidence of truncation, which identified 103 truncated proteins, including 56 newly identified proteins (fig. S3L). Several identified truncated proteins are predicted to lose carboxyterminal enzymatic domains, including Ilk, Crk, Ikbkb. The melanoma tumor suppressor Idh2 was also affected (Fig. 3G and fig. S3M) (*46*–*48*). These data show that, rather than being degraded by the nuclear exosome, ptRNAs are being exported and translated into truncated proteins in CDK13^{mut}-expressing cells.

We wondered whether intronic sequences at the end of ptRNAs are translated (schematized in Fig. 3E). ptRNA translation products were predicted by assuming in-frame translation from the upstream exon until the first stop codon before ptRNA termination as determined from 3′ seq. ptRNA translation products were added to the canonical proteome for tandem MS data search. As predicted, intronic peptides were enriched specifically in the CDK13^{mut} zebrafish melanomas (Fig. 3H). The abundance of some intronic peptides approached that of medium- to lowly expressed full-length proteins (fig. S3, N to P). To verify truncated protein expression, we selected two predicted truncated proteins with available N-terminal antibodies for immunoblotting. We identified truncated protein products from CRBN and CDK13 itself from CDK13-mutant expressing versus control human melanoma cells (fig. S3, Q and R). These data indicate that some ptRNAs from the CDK13 mutant melanomas are translated, including the intronic sequence preceding IPA sites, which could result in the production of tumor-specific neoantigens.

Mutant CDK13 disrupts the polyA RNA exosome

To elucidate how CDK13^{mut} expression results in ptRNA accumulation through a posttranscriptional mechanism, CDK13^{WT} was tagged, transiently expressed, and immunoprecipitated (IPed) from human melanoma cell nuclear extracts, and co-IPed proteins were identified by MS (Fig. 4A). CDK13^{WT} IP-MS identified 37 proteins enriched in the CDK13^{WT} IP versus control (Fig. 4B). The most enriched ontology among these 37 proteins was "mRNA 3′ end processing" (reactome, $q = 7.15 \times 10^{-13}$) (*49*). CDK13 bound CCNT1 (Fig. 4, B and C), again implicating CCNT1 as an important binding partner for CDK13 in melanoma as our functional data suggested (Fig. 1, K and L, and fig. S1U). The canonical cyclin-binding partner of CDK13, CCNK, was detected at levels above background but below thresholding. No other cyclins were detected above background. Native CDK13 IP identified CCNT1 and CCNK by immunoblot (fig. S4A). This unbiased approach suggests that CDK13 binds CCNT1 in addition to CCNK in melanoma.

Of the enriched proteins in the CDK13^{WT} IP (Fig. 4C), both PABPN1 and ZC3H14 interact with the polyA tail eXosome Targeting (PAXT) complex (*22*), which is responsible for targeting ptRNAs for degradation (*23*). CDK13 binding to ZC3H14 was verified by immunoblot (fig. S4B). Because ptRNAs accumulate in CDK13^{mut} melanoma through a posttranscriptional mechanism and are degraded by the PAXT complex (*50*), we hypothesized that CDK13 normally works to activate PAXT. We hypothesized that loss of CDK13 kinase activity would fail to activate PAXT, leading to accumulated ptRNAs in melanoma.

Because PABPN1 and ZC3H14 are genetically antagonistic (*51*, *52*) and because we observed more ZC3H14 in our CDK13^{WT} IP, we chose to characterize ZC3H14 phosphorylation and binding partners in the presence and absence of CDK13 kinase activity. The nuclear isoform of ZC3H14 and a control protein were tagged, transiently expressed, and IPed from CDK13^{WT} and CDK13^{mut} human melanoma cells (fig. S4C, white, black, and blue). IPed ZC3H14 from CDK13^{WT} cells had four phosphorylation sites, whereas in CDK13^{mut} cells, ZC3H14 lost

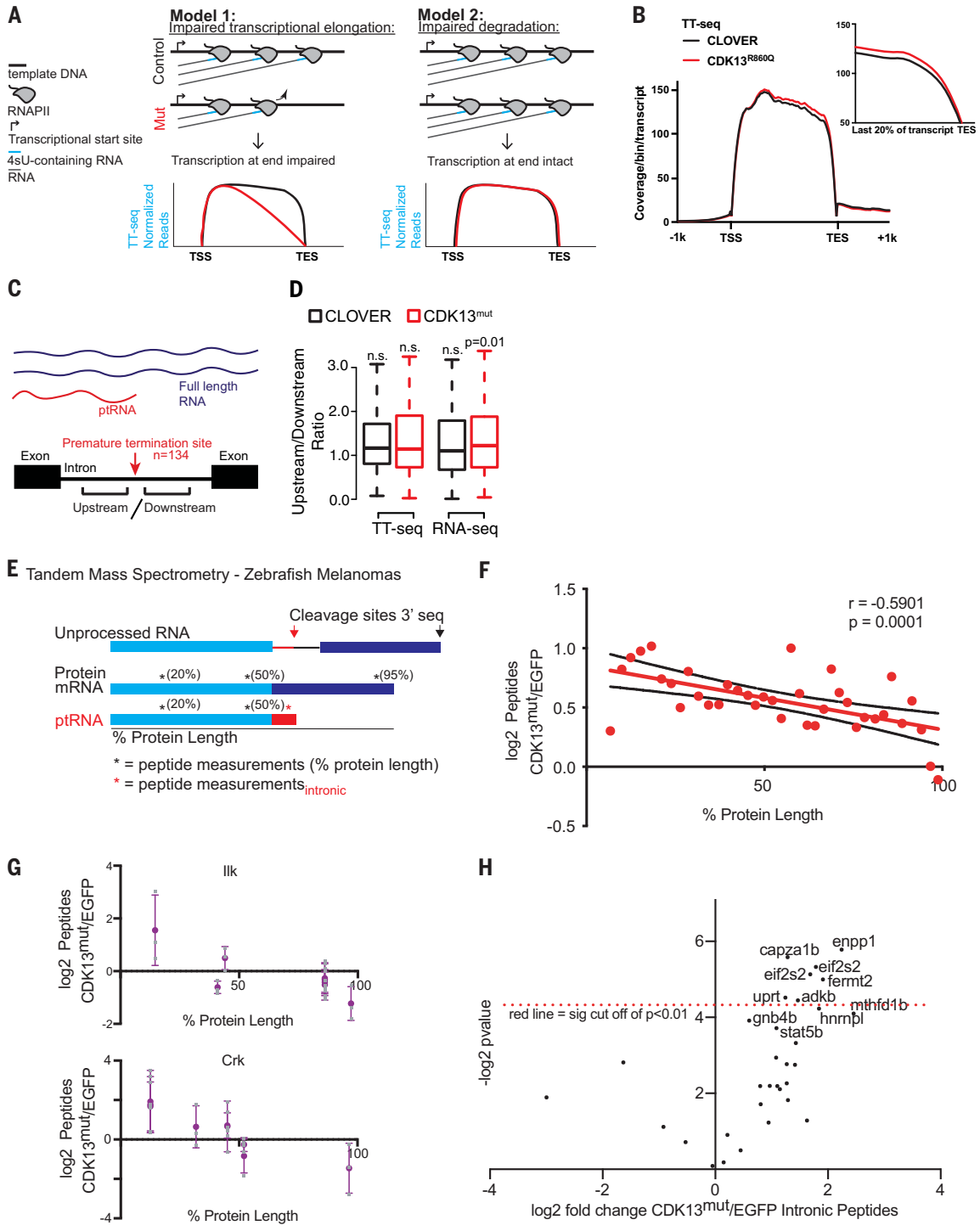


Fig. 3. PtRNAs accumulate posttranscriptionally and are translated - including intronic sequences. (**A**) Models for ptRNA accumulation as measured by TT-seq. TSS, transcriptional start site; TES, transcriptional end site. (**B**) Metagene plots are shown for TT-seq coverage over exons within nonoverlapping transcripts from expressed genes (N = 7452) from CDK13^{mut} versus CLOVER-expressing human melanoma cells. TSS to TES regions are shown, flanked ± 1-kb genomic sequence. (**C**) Schematic for analysis of TT-seq and RNA-seq coverage around intronic polyadenylation sites. (**D**) Box plots of the ratio of upstream [−300 to −1 nucleotides (nt)] to downstream (+1 to +300 nt) read coverage at the 3′ cleavage locations as schematized in (C). P values are from Wilcoxon signed-rank test compared with the median ratio from all samples. (**E**) Schema of tandem mass spectrometry data analysis from CDK13^{mut}- versus EGFP-expressing zebrafish melanomas. (**F**) Log2 peptide summed signal to noise (SSN) for CDK13^{mut}-expressing versus EGFP-expressing zebrafish melanomas plotted across a normalized protein length. n = 3 zebrafish melanomas for each condition. Black lines indicate 95% confidence intervals. (**G**) Log2 CDK13^{mut}-expressing versus EGFP-expressing zebrafish melanoma peptide measurements plotted by percentage protein length for significantly affected proteins for Ilk (upper) and Crk (lower). Gray squares are individual measurements. Error bars indicate SD. (**H**) Log2 intronic peptides SSN from CDK13^{mut}-expressing versus EGFP-expressing zebrafish melanomas as predicted from 3′ sequencing on the horizontal axis plotted by −log P value on the vertical axis from two-tailed t test. Labels are enriched intronic peptides.

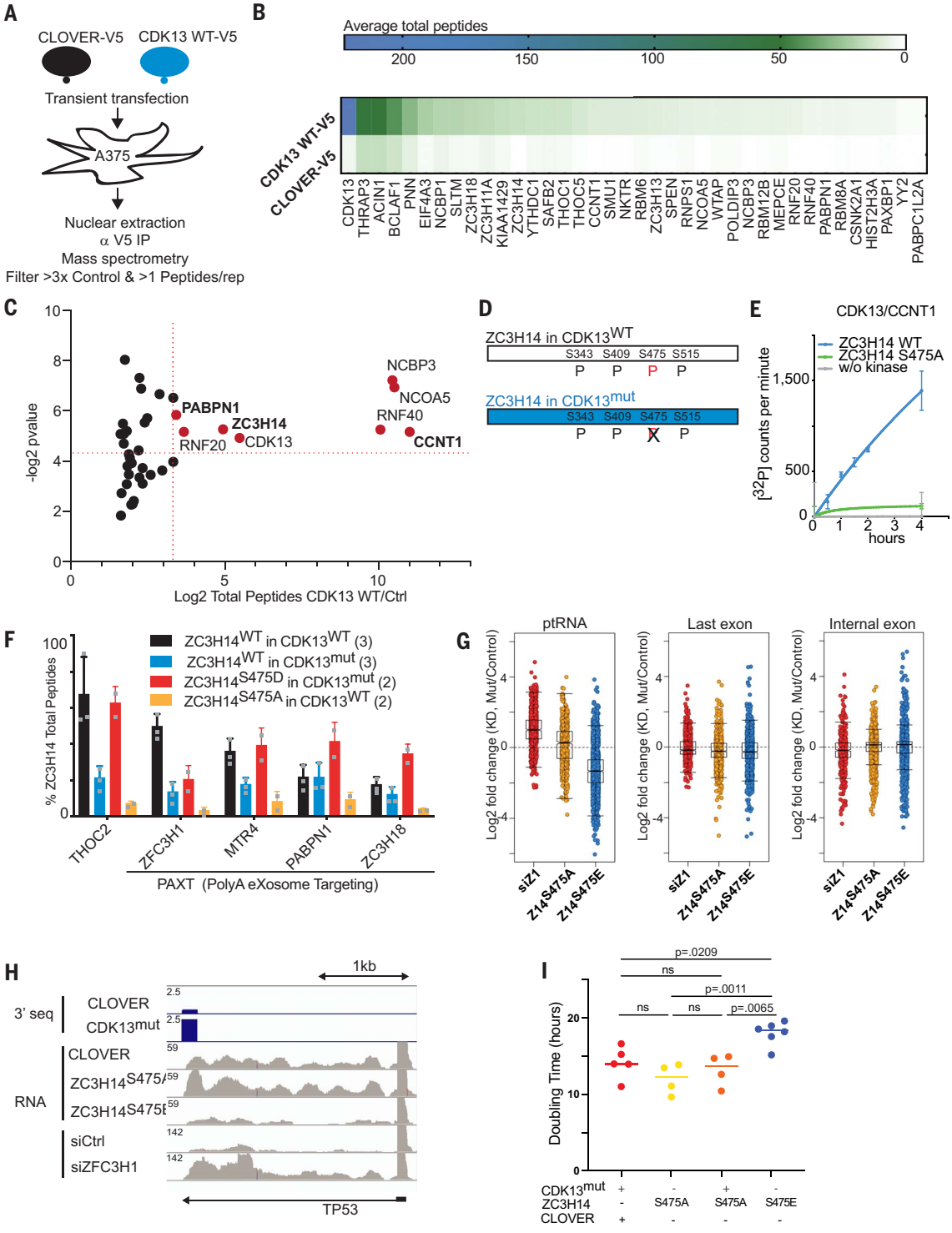


Fig. 4. Mutant CDK13 disrupts the polyA RNA exosome. (**A**) CDK13 IP-MS schema. (**B**) Heatmap of average total peptides from anti-V5 IP-MS of CDK13 WT-V5 (n = 3) or CLOVER-V5 (n = 2). (**C**) Log2-fold change CDK13^{WT}/control total peptides by −log P value (two-tailed t test). Upper right quadrant indicates proteins with a −log P value < 0.05 and a log-fold change enrichment over control >3.3. (**D**) Graphic depicting ZC3H14 phosphorylation sites identified from ZC3H14 IPed from either CDK13^{WT}-expressing or CDK13^{mut}-expressing human melanoma cells. (**E**) In vitro kinase assay of WT CDK13/CCNT1 using ZC3H14 full-length WT and S475A substrate. P = 0.00720098, two-tailed t test WT versus S475A, n = 3 for each condition. (**F**) Total peptides as a percentage of bait (ZC3H14 total peptides) for proteins with binding most regulated by CDK13 kinase activity. Data are shown as means ± SD. n = number of biologic replicates. Gray dots are individual values. (**G**) Box plot showing log2-fold change of siZFC3H1, ZC3H14^{S475A}, or ZC3H14^{S475E} compared with the relevant control for ptRNAs (left), last exons (middle), or internal exons (right). Significantly changed genes (q < 0.1) plotted for all RNA types in each experiment. Z1, ZFC3H1; Z14, ZC3H14. ptRNA, RNAs generated by usage of intronic polyadenylation sites. Internal exon is the constitutive internal exon (nonalternatively spliced) isoform. In box plots, the black horizontal line indicates the median and whiskers extend to 1.5× the IQR. Statistics in table S5 (medians and two-sided Wilcoxon rank-sum P values). (**H**) IGV plot of *TP53* ptRNA that is regulated by CDK13, ZFC3H1, and ZC3H14 S475 phosphorylation. Scale shows normalized RNA expression (note different scales). (**I**) Doubling time of human melanoma cells expressing CDK13^{mut}, ZC3H14^{S475A}, ZC3H14^{S475E}, or CDK13^{mut};ZC3H14^{S475A} from one-way ANOVA multiple-comparisons test.

phosphorylation only at S475 (Fig. 4D and fig. S4, D and E). Thus, CDK13 kinase activity is required for ZC3H14 S475 phosphorylation. In vitro kinase reactions [radioactive kinase assay in Fig. 4E and labeled ATP immunoblot (*53*) in fig. S4F] demonstrated that CDK13/CCNT1 was able to phosphorylate full-length WT ZC3H14 but not ZC3H14 S475A.

We looked at ZC3H14's binding partners in the presence and absence of CDK13 kinase activity. When ZC3H14 was IPed from CDK13^{mut} cells, fewer binding partners were identified

(ZC3H14 peptides were not statistically different) (figs. S4G, rows 2 and 3, and S4H, left columns). Total IPed peptides were normalized to total bait (ZC3H14) peptides. Differential binding was calculated between CDK13^{WT} and CDK13^{mut} conditions (*t* test $P < 0.05$), which identified 18 proteins that required intact CDK13 kinase activity to promote binding to ZC3H14. The three most abundant proteins that required CDK13 kinase activity to promote binding to ZC3H14 were THOC2, ZFC3H1, and MTR4 (Fig. 4F, black versus blue bars) (THOC2 $q < 0.0001$, ZFC3H1 $q < 0.0001$, and MTR4 $q = 0.018$). THOC2 functions in RNA export and binds ZC3H14 (*54*), ZFC3H1 is a linker between the PAXT and the nuclear RNA degradation machinery, and MTR4 is a helicase required for nuclear RNA degradation (*22*). We also observed that ZC3H14 bound PAXT proteins PABPN1 and ZC3H18; however, binding of these proteins to ZC3H14 was independent of CDK13 kinase activity (Fig. 4F). These data show that ZC3H14 binds to multiple PAXT members, and that CDK13 kinase activity promotes ZC3H14 binding to THOC2 and two key PAXT components, ZFC3H1 and MTR4.

To test whether the ZC3H14 S475 phosphorylation was sufficient to recruit PAXT binding, phosphomimetic ZC3H14^{S475D} was tagged, transiently expressed, and IPed from cells lacking CDK13 kinase activity. ZC3H14^{S475D} was sufficient to rescue binding of PAXT components to ZC3H14 (Fig. 4F, red bars, and fig. S4I, fourth row, and I and J), even in cells lacking CDK13 kinase activity. To test whether ZC3H14 S475 was necessary for PAXT recruitment, nonphosphorylatable ZC3H14^{S475A} was IPed from CDK13^{WT} cells. ZC3H14^{S475A} failed to recruit PAXT components (Fig. 4F, yellow bars, and fig. S4G, last row, and K and L), even in cells with intact CDK13 kinase activity. The amount of ZC3H14 IPed was statistically unchanged (fig. S4H, right columns). Together, these data show that ZC3H14 S475 phosphorylation promotes PAXT recruitment to ZC3H14.

To determine whether ZC3H14 S475 phosphorylation is also necessary and sufficient to activate ptRNA degradation, a two-pronged approach was undertaken. First, short interfering RNAs were used to decrease levels of ZFC3H1 ($n = 3$) or ZC3H14 ($n = 3$), and differential RNA expression was assessed using 3′ seq and RNA-seq (fig. S4M) compared with a scrambled control ($n = 3$). Second, stable human melanoma cell lines expressing nonphosphorylatable ZC3H14^{S475A} ($n = 2$), phosphomimetic ZC3H14^{S475E} ($n = 3$), or a control protein ($n = 3$) were subjected to RNA-seq (fig. S4N). ZC3H14^{S475E} was used because we were unable to make stable lines with ZC3H14^{S475D}. 3′ sequencing from siZFC3H1, siZC3H14, siControl, and prior CDK13^{mut} and control samples were used to build an expression map for ptRNAs and dominant isoforms in human melanoma cells. This ptRNA isoform expression map was used to calculate differential RNA expression using DEXseq in all conditions compared with appropriate controls. A boxplot of significantly changed RNAs ($q < 0.1$) from ZFC3H1 versus control knockdown in human melanoma cells demonstrated an increase in ptRNA isoforms, but minimally changed last exons and constitutive internal exons as expected (Fig. 4G, red). siZC3H14 visualized in the same manner showed very few significant expression changes (fig. S4O), which may be caused by the activity of residual protein or redundancy with another protein. Expression of ZC3H14^{S475A} versus CLOVER ($q < 0.1$) showed a modest increase in ptRNA expression, whereas the last and constitutive internal exons were minimally changed (Fig. 4G, yellow). Expression of ZC3H14^{S475E} versus CLOVER ($q < 0.1$) caused a decrease in ptRNAs while not affecting last or internal exons (Fig. 4G, blue) (also see medians and Wilcoxon rank-sum comparing ptRNAs, internal exons, and last exons in table S5). The *TP53* ptRNA, which is up-regulated in CDK13^{mut} cells, was up-regulated in ZFC3H1 knockdown, and upon nonphosphorylatable ZC3H14^{S475A} expression, consistent with loss of PAXT nuclear RNA degradation. By contrast, phosphomimetic ZC3H14^{S475E} caused lower expression of the *TP53* ptRNA, consistent with hyperactivation of PAXT RNA degradation (Fig. 4H). Together, these data show that ZC3H14 S475 phosphorylation is necessary and sufficient to activate PAXT to degrade ptRNAs.

To test whether expression of ZC3H14 phosphomimetic and nonphosphorylatable mutants affected human melanoma cell growth rate, doubling time was measured. ZC3H14^{S475A} expressing-cells proliferated at a similar rate to CDK13^{mut}-expressing cells, whereas ZC3H14^{S475E}-expressing cells had an increased doubling time (i.e., grew more slowly) (Fig. 4I, columns 1, 2, and 4). These data are consistent with the hypothesis that cells with higher ptRNA levels have an increased growth rate, whereas cells with lower ptRNA expression have a slower growth rate. To determine whether effects of CDK13^{mut} and ZC3H14^{S475} expression function in the same or a parallel ptRNA surveillance pathway, ZC3H14^{S475} phosphomimetic and nonphosphorylatable mutants were expressed in CDK13^{mut} human melanoma cells. We were unable to recover CDK13^{mut} cells that expressed ZC3H14^{S475E}, indicating that CDK13^{mut} cells are addicted to ptRNA expression or that carrying both mutations is deleterious to cell viability. CDK13^{mut};ZC3H14^{S475A} cells had a similar growth rate to either mutant alone (Fig. 4I, column 3) and affected the same ptRNAs to a similar magnitude as CDK13^{mut} (fig. S4, P and Q). Our data demonstrate that mutant CDK13 expression caused inefficient PAXT recruitment and ptRNA stabilization, supporting an oncogenic role for ptRNAs (fig. S4R).

ptRNA accumulation is oncogenic

To determine whether *CDK13*^{mut} causes accumulation of ptRNAs in other cancers, publicly available RNA-seq from patient tumor samples from multiple tumor types were examined (bladder, colorectal, lung, melanoma, lung, and uterus). RNA-seq from tumors with CDK13 kinase-domain or nonsense mutations were compared against CDK13^{WT} tumors with matched tumor and patient characteristics (table S6) using the PAXT-regulated ptRNA isoform map developed for Fig. 4. PAXT-regulated ptRNA isoforms are significantly up-regulated in CDK13^{mut} cancers compared with internal exons (Fig. 5A; $P < 2.2 \times 10^{-16}$). Last exons are also up-regulated relative to internal exons, suggesting a general increase in the stability of polyadenylated RNAs in CDK13 mutant tumors (Fig. 5A), which may have been revealed because of the high RNA degradation levels typically observed in patient samples. The CDK13/PAXT target *TP53* ptRNA was more highly expressed in RNA-seq coverage profiles of CDK13 mutant tumors compared with matched CDK13^{WT} controls (Fig. 5B). These data show that patient tumors with CDK13 mutations from many cancer types exhibit an accumulation of PAXT-sensitive ptRNAs.

To test whether ptRNA expression is oncogenic, human *TP53* ptRNA or *SUV39H1* ptRNA versus control EGFP were expressed in melanocytes in the Triples zebrafish melanoma model. Human *TP53* ptRNA expression caused increased black patches at 7 weeks, consistent with early melanoma ($P < 0.0001$, two-sided chi-square test) (Fig. 5, C and D, and fig. S5A) and expedited melanoma onset ($P = 0.0391$) (Fig. 5E). Expression of human *SUV39H1* ptRNA also caused increased black patches at 7 weeks ($P = 0.0051$) (Fig. 5, F and G, and fig. S5B) and expedited melanoma onset ($P < 0001$) (Fig. 5H). Quantitative PCR confirmed ptRNA expression (fig. S5, C and D). Because the *TP53* ptRNA is oncogenic and derives from the *TP53* locus, we performed p53 immunoblots and found that CDK13^{mut} human melanoma cells have intact p53 full-length protein (fig. S5E). These data show that human ptRNAs are sufficient to expedite melanoma onset in zebrafish, and more generally that ptRNAs can be oncogenic.

Because the loss of PAXT activity may represent a more widespread mechanism of oncogenesis, we probed public databases and found that CDK13-regulated PAXT members are deleted or mutated in 17% of melanomas (Table 2). Analysis with OncodriveFM in melanoma samples found that *ZC3H18* is a marginally significant genome-wide driver gene ($q = 0.057$; $P = 0.0074$), whereas *ZFC3H1* was not found to be a likely driver ($q = 0.22$;

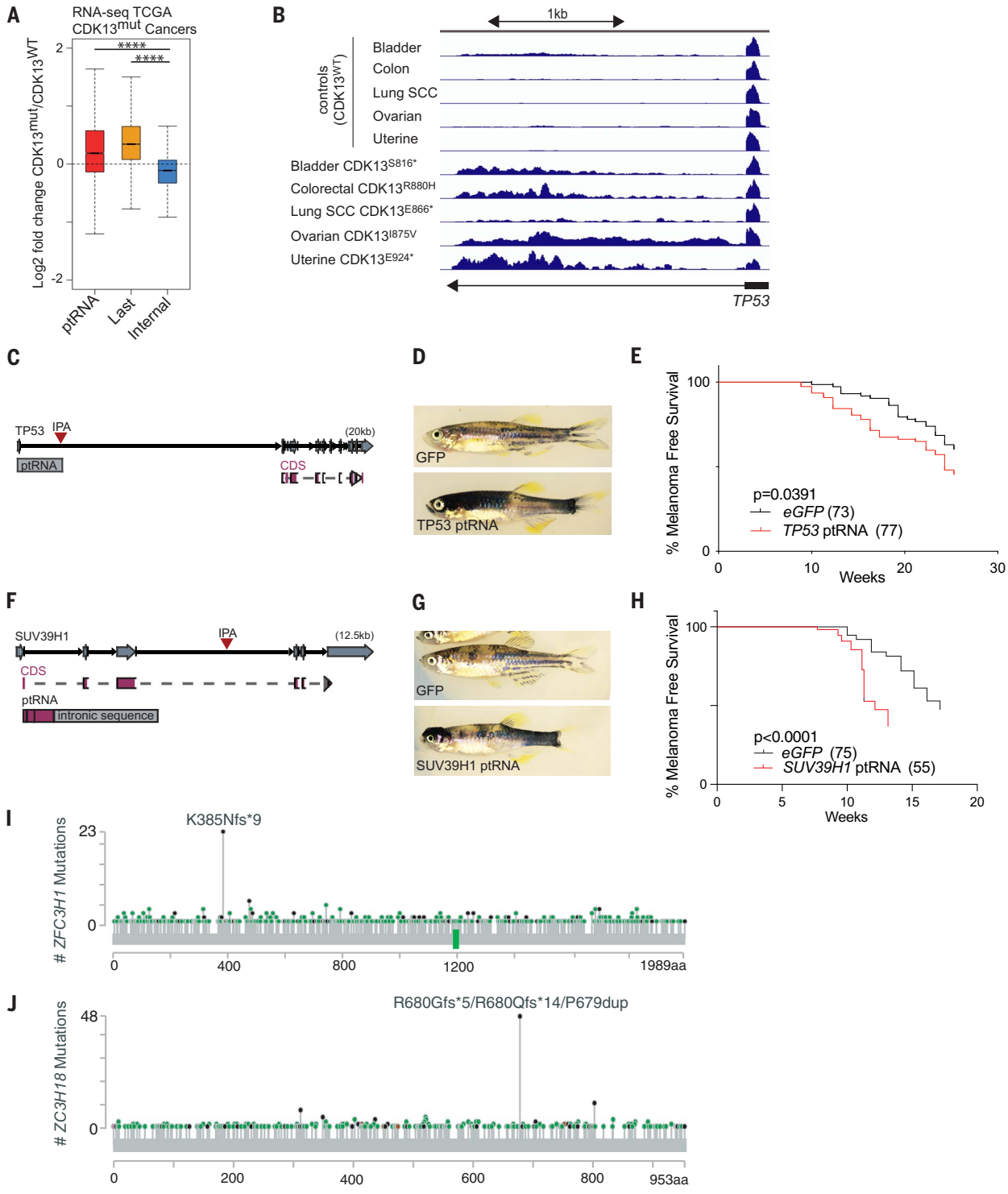


Fig. 5. ptRNA accumulation is oncogenic. (**A**) ptRNA quantification from somatic CDK13^{mut} (n = 14) compared with matched CDK13^{WT} (n = 14) cancers (many types). ****$P < 2.2 \times 10^{-16}$ (two-sided Wilcoxon rank-sum test). The black horizontal line indicates the median and whiskers extend to 1.5× the IQR. (**B**) IGV RNA-seq coverage plot of PAXT-target *TP53* ptRNA from sample subset from (A). (**C**) *TP53* ptRNA in *TP53* locus. IPA = intronic polyadenylation site. CDS = coding sequence. (**D** and **E**) Triples melanoma model zebrafish with melanocyte-specific expression of EGFP versus human *TP53* ptRNA. (D) 7-week photos. (E) Percentage melanoma-free survival (log-rank). n = number of zebrafish. (**F**) *SUV39H1* ptRNA in *SUV39H1* locus. (**G** and **H**) Triples melanoma model zebrafish with melanocyte-specific expression of EGFP versus human *SUV39H1* ptRNA. (G) 7-week photos. (H) Percentage melanoma-free survival (log-rank). (**I** and **J**) Lollipop plots of *ZFC3H1* (I) and *ZC3H18* (J) mutations in nonredundant publicly available sequencing data from all cancers.

Table 2. CDK13-regulated PAXT subunits are mutated or deleted in melanoma TCGA patients.

Gene	Mutation*
THOC2	7% (*n* = 19)
ZFC3H1	5% (*n* = 14)
ZC3H14	3% (*n* = 8)
MTR4	2% (*n* = 6)

*Percentages of cases out of 287.

P = 0.051) (fig. S5, F and G). The PAXT adaptor protein ZFC3H1 is recurrently mutated (K385Nfs*9) in 23 different patient tumors representing a wide variety of cancers (Fig. 5I) and has nonrecurrent mutations in 5 to 11% of multiple cancer types, including nonmelanoma skin cancer, endometrial cancer, colon adenocarcinoma, and small-cell lung cancer (fig. S5H). *ZC3H18* is recurrently mutated (R680Q/Gfs*) in 48 individual samples from a broad variety of cancers (Fig. 5J) and has nonrecurrent mutations in many cancers, including nonmelanoma skin cancers, vaginal cancers, and endometrial cancers (fig. S5I). Together, these data suggest that nuclear RNA quality control through the PAXT complex is critical to avoid the accumulation of ptRNAs and their aberrant protein products in multiple cancer types, and support deficient nuclear RNA surveillance as a general oncogenic mechanism.

Discussion

Our data show that *CDK13* has properties consistent with a tumor suppressor in which mutations lead to accumulation of prematurely terminated transcripts by preventing their normal degradation. CDK13 normally phosphorylates ZC3H14 S475, which is necessary and sufficient to promote PAXT recruitment and ptRNA degradation. When CDK13 is mutated, failure to recruit the PAXT complex results in the accumulation of truncated oncogenic transcripts, which are exported to the cytoplasm and translated into truncated proteins. The expression of ptRNAs observed in CDK13 mutant melanoma can cause more aggressive cancer. Our data show that mutant CDK13 disrupts a deeply conserved pathway to cause ptRNA accumulation, and that ptRNA accumulation is oncogenic.

In the CDK13-associated developmental disorder, de novo CDK13 mutations are thought to work through a dominant-negative mechanism. Homozygous CDK13 loss of function causes lethal heart malformations in mice (*55*), and likely causes embryonic lethality in humans because homozygous deletion is not observed in phenotypically normal people (*n* = 54,980) (*56*) or in patients with a developmental disorder (*n* = 43,173) (*57*). Heterozygous *CDK13* deletions have been identified in normal humans in two genomic aggregate databases (gnomAD, Database of Genomic Variants) (*19*). By contrast, heterozygous *CDK13* kinase domain mutations cause a syndromic developmental disorder that affects heart development as well as other neural crest–derived tissues. Our work raises the possibility that individuals with a *CDK13*-related disorder could be predisposed to developing cancer, which is supported by a recent report of a 9-year-old patient with the *CDK13*-related disorder developing leukemia (*58*). The strong mutant *CDK13*-associated disorder phenotypes are more severe than the loss of one *CDK13* allele, which supports that CDK13 kinase-domain mutations have dominant-negative activity.

Our initial insight for this work came from the observation of that CDK13 mutations in melanoma were enriched in the kinase domain, which suggested a genetic function beyond haploinsufficiency. Our data support a model in which CDK13 heterozygous loss of function is selected for (50% activity loss), whereas heterozygous kinase domain mutations are additionally selected for. This is because nonphosphorylatable CDK13 acts in a dominant-negative manner, that is, it blocks $CDK13^{WT}$ from activating its targets (>50% activity loss). Loss of both alleles is infrequently observed, possibly because of reduced cellular fitness, as was seen for a cancer cell line (*59*). We supposed that CDK13 kinase-domain mutations cause a more severe PAXT deficiency and thus are more selected for than mutations that cause CDK13 haploinsufficiency. Our zebrafish genetics and the lack of CDK13 mutations in melanoma cyclin-binding residues suggested that $CDK13^{mut}$ dominant-negative activity requires cyclin binding. Expression of $CDK13^{mut}$ in a $CDK13^{WT}$ background recapitulated the RNA cancer phenotypes observed in melanoma patients with heterozygous CDK13 mutations. In multiple CDK13 experimental systems, we found that mutant CDK13 causes the same RNA and cancer phenotype, namely accumulated ptRNAs and more aggressive melanoma. We hypothesized that only by modeling the human genetics and chronically expressing kinase-mutant CDK13 in the presence of $CDK13^{WT}$ would we be able to uncover the role of CDK13 in RNA surveillance. Our work shows that CDK13's function in ptRNA degradation is used broadly.

ptRNAs in $CDK13^{mut}$ cells are exported and translated into truncated proteins. Transcripts that terminate through intronic polyadenylation sites are predicted to be exported and translated if they escape nuclear decay (*21*, *45*). We are not aware of any studies documenting proteome-wide measurements of ptRNA translation, including intronic sequences. Large-scale truncated protein expression would be predicted to cause protein stress in cells, which is consistent with the observed up-regulation of autophagic and lysosomal proteins in CDK13 mutated melanomas. Of potentially great importance for predicting immunotherapy responses in patients, we found evidence that the truncated proteins can end in translated intronic sequences. These sequences have the potential to produce neoantigens that would be absent from $CDK13^{WT}$ cells and thus would be predicted to elicit a highly tumor-specific immune response.

ptRNAs derived from intronic polyadenylation sites are enriched in cancers (*60*). We hypothesize that different cancers will accumulate different ptRNAs depending on the RNAs expressed and the mechanism of ptRNA accumulation. Loss-of-function mutations in CDK12 in metastatic castration-resistant prostate cancer and serous ovarian carcinomas cause increased production of truncated RNAs in DNA repair genes (*7*, *8*), contributing to CDK12's tumor-suppressive properties. The IPA sites up-regulated by CDK12 loss appear to be mostly but not fully distinct from those seen when CDK13 is lost (fig. S2E and data S1), consistent with CDK12 and CDK13 having divergent mechanisms of ptRNA regulation (Table 3).

Widespread ptRNA formation has also been reported to take place when the U1 small nuclear RNA is prevented from pairing with 5′ splice sites in a phenomenon called "telescripting" (*61*). The U1 spliceosomal RNA is recurrently mutated in multiple cancers at base 3, which pairs with the 5′ splice site (*62*). U1 mutations could cause increased generation of ptRNAs and could thus represent a third mechanism by which ptRNA expression promotes oncogenesis. We envision mechanisms that govern 3′ end formation, such as telescripting, and those that govern ptRNA stability, such as CDK13, as being complementary ptRNA control mechanisms. It will be important to investigate the expression of ptRNAs and other aberrant RNAs in different cancers with different driver mutations because aberrant RNA accumulation could represent a final common pathway.

Germline mutations in the PAXT-associated proteins ZC3H14 and THOC2 cause a neurodevelopmental disorder (*51*, *63*), suggesting that the regulation of ptRNAs in the nucleus has broad implications in development and disease. We propose that loss of nuclear RNA surveillance through CDK13 and PAXT is a general oncogenic mechanism. PAXT members regulated by CDK13 are mutated in 17% of melanomas in addition to the 3.9% of melanomas harboring CDK13 mutations, raising the possibility that loss of nuclear surveillance is a contributing factor in up to 21% of melanoma. We showed that patients with other malignancies harboring CDK13 mutations also

Table 3. CDK13 has a related but distinct biologic role and mechanism from CDK12.

	CDK13	CDK12
Mutated in cancer	Melanoma (3.9%), cancer cell lines (8%)	Prostate, ovarian (*64*, *65*)
Mutated in developmental disorders	De novo heterozygous kinase domain mutations and truncating mutations cause developmental disorders that affect neural crest-derived tissues (congenital heart defects, dysmorphic facial features, and intellectual development disorder) (CHDFIDD) (*15*–*19*)	Not detected
RNA signature	No loss of DNA repair gene expression	Loss of DNA repair gene expression (*7*, *8*, *40*, *66*)
RNA mechanism	Required to activate PAXT degradation of ptRNAs	Required to suppress usage of intronic polyadenylation sites (*7*, *8*)

accumulate ptRNAs. We also found that recurrent mutations in the key PAXT members *ZFC3H1* and *ZC3H18* occur in many cancers, implying that these mutations may potentially be selected for. Further work will be required to support this hypothesis. The finding that mutant *CDK13* is oncogenic through deficient RNA surveillance and that recurrent mutations occur in multiple PAXT components suggests a broad, previously unrecognized tumor-suppressive role for nuclear RNA surveillance.

Methods summary

Human patient data analysis of *CDK13* was done by downloading publicly available melanoma DNA-sequencing data, and then these data were used for OncodriveFM analysis and kinase domain enrichment analysis in melanoma patients. TCGA melanoma data were used for patient survival analyses and clonality analyses. Publicly available DNA tumor sequence data from all cancers were used to elucidate the kinase domain enrichment of mutations in cancers other than melanoma. In vitro kinase assays were executed with purified CDK13 activated by a cyclin on purified protein substrates. Zebrafish melanocyte and melanoma modeling were done using a rapid F0 melanocyte-specific modeling system in Triples ($p53^{-/-}$;$mitfa$:$BRAF^{V600E}$;$mitfa^{-/-}$) zebrafish. Stable human melanoma cell lines were made with lentiviral transduction in A375 human melanoma cells. 5′ to 3′ RNA-seq analysis measured differential exon expression in four systems: (i) TCGA melanoma publicly available $CDK13^{mut}$ melanomas, (ii) zebrafish $CDK13^{mut}$-expressing melanomas, (iii) mESCs depleted for CDK13, and (iv) $CDK13^{mut}$-expressing human melanoma cells. 3′ sequencing was executed (Lexogen REV) and RNA cleavage site location was measured in $CDK13^{mut}$ versus control zebrafish melanoma and human melanoma cell systems. ddPCR measured first and last exon expression on four affected genes and two control genes in $CDK13^{mut}$ and control human melanoma cells. Nascent RNA-seq (TT-seq) was done using a 5-min 4sU pulse in $CDK13^{mut}$ and control human melanoma cells. Chromatin immunoprecipitation was done using published antibodies to hypophosphorylated RNAPII (initiating RNAPII) and RNAPII Ser2 CTD phosphorylated (elongating RNAPII) in the zebrafish and human melanoma cell systems. Tandem mass spectrometry was done on $CDK13^{mut}$ versus control zebrafish melanomas. Peptide measurements were plotted along normalized percent protein length. Translation of intronic sequence was predicted assuming in-frame translation from the upstream exon. IP-MS was performed by IPing $CDK13^{WT}$ in A375 cells and filtering for repeatability and enrichment over control IP. ZC3H14 IP-MS was performed in $CDK13^{WT}$ and $CDK13^{mut}$ human melanoma cells, and ZC3H14 phosphorylation and binding partners were assessed in both conditions. ZC3H14 S475 phosphomimetic protein was IPed from $CDK13^{mut}$ cells, and ZC3H14 unphosphorylatable S475 was IPed from $CDK13^{WT}$ cells. Proliferation and ptRNA levels were quantified from stable human melanoma cell lines expressing $ZC3H14^{S475}$ mutants with or without $CDK13^{mut}$. TCGA RNA-seq data were used for ptRNA isoform quantification in other, nonmelanoma $CDK13^{mut}$ cancers. Human ptRNAs were expressed in the above zebrafish system. Recurrent mutations in other PAXT members were identified from publicly available DNA-sequencing data from all cancers. Please see the full materials and methods in the supplementary materials for more detailed information.

REFERENCES AND NOTES

1. P. Hydbring, M. Malumbres, P. Sicinski, Non-canonical functions of cell cycle cyclins and cyclin-dependent kinases. *Nat. Rev. Mol. Cell Biol.* **17**, 280–292 (2016). doi: 10.1038/nrm.2016.27; pmid: 27033256
2. M. Malumbres, Cyclin-dependent kinases. *Genome Biol.* **15**, 122 (2014). doi: 10.1186/gb4184; pmid: 25180339
3. K. Adelman, J. T. Lis, Promoter-proximal pausing of RNA polymerase II: Emerging roles in metazoans. *Nat. Rev. Genet.* **13**, 720–731 (2012). doi: 10.1038/nrg3293; pmid: 22986266
4. S. McCracken *et al*., 5′-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.* **11**, 3306–3318 (1997). doi: 10.1101/gad.11.24.3306; pmid: 9407024
5. B. E. Schwartz, S. Larochelle, B. Suter, J. T. Lis, Cdk7 is required for full activation of Drosophila heat shock genes and RNA polymerase II phosphorylation in vivo. *Mol. Cell. Biol.* **23**, 6876–6886 (2003). doi: 10.1128/MCB.23.19.6876-6886.2003; pmid: 12972606
6. E. Y. Shim, A. K. Walker, Y. Shi, T. K. Blackwell, CDK-9/cyclin T (P-TEFb) is required in two postinitiation pathways for transcription in the C. elegans embryo. *Genes Dev.* **16**, 2135–2146 (2002). doi: 10.1101/gad.999002; pmid: 12183367
7. S. J. Dubbury, P. L. Boutz, P. A. Sharp, CDK12 regulates DNA repair genes by suppressing intronic polyadenylation. *Nature* **564**, 141–145 (2018). doi: 10.1038/s41586-018-0758-y; pmid: 30487607
8. M. Krajewska *et al*., CDK12 loss in cancer cells affects DNA damage response genes through premature cleavage and polyadenylation. *Nat. Commun.* **10**, 1757 (2019). doi: 10.1038/s41467-019-09703-y; pmid: 30988284
9. J. K. Rimel *et al*., Selective inhibition of CDK7 reveals high-confidence targets and new models for TFIIH function in transcription. *Genes Dev.* **34**, 1452–1473 (2020). doi: 10.1101/gad.341545.120; pmid: 33060135
10. E. Chipumuro *et al*., CDK7 inhibition suppresses super-enhancer-linked oncogenic transcription in MYCN-driven cancer. *Cell* **159**, 1126–1139 (2014). doi: 10.1016/j.cell.2014.10.024; pmid: 25416950
11. C. L. Christensen *et al*., Targeting transcriptional addictions in small cell lung cancer with a covalent CDK7 inhibitor. *Cancer Cell* **26**, 909–922 (2014). doi: 10.1016/j.ccell.2014.10.019; pmid: 25490451
12. N. Kwiatkowski *et al*., Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. *Nature* **511**, 616–620 (2014). doi: 10.1038/nature13393; pmid: 25043025
13. Y. Wang *et al*., CDK7-dependent transcriptional addiction in triple-negative breast cancer. *Cell* **163**, 174–186 (2015). doi: 10.1016/j.cell.2015.08.063; pmid: 26406377
14. H. E. Sheppard *et al*., Targeted brachyury degradation disrupts a highly specific autoregulatory program controlling chordoma cell identity. *Cell Rep. Med.* **2**, 100188 (2021). doi: 10.1016/j.xcrm.2020.100188; pmid: 33521702
15. Deciphering Developmental Disorders Study, Prevalence and architecture of de novo mutations in developmental disorders. *Nature* **542**, 433–438 (2017). doi: 10.1038/nature21062; pmid: 28135719
16. A. Sifrim *et al*., Distinct genetic architectures for syndromic and nonsyndromic congenital heart defects identified by exome sequencing. *Nat. Genet.* **48**, 1060–1065 (2016). doi: 10.1038/ng.3627; pmid: 27479907
17. W. M. R. van den Akker *et al*., De novo variants in CDK13 associated with syndromic ID/DD: Molecular and clinical delineation of 15 individuals and a further review. *Clin. Genet.* **93**, 1000–1007 (2018). doi: 10.1111/cge.13225; pmid: 29393965
18. B. L. Bostwick *et al*., Phenotypic and molecular characterisation of CDK13-related congenital heart defects, dysmorphic facial

features and intellectual developmental disorders. *Genome Med.* **9**, 73 (2017). doi: 10.1186/s13073-017-0463-8; pmid: 28807008
19. M. J. Hamilton *et al*., Heterozygous mutations affecting the protein kinase domain of *CDK13* cause a syndromic form of developmental delay and intellectual disability. *J. Med. Genet.* **55**, 28–38 (2018). doi: 10.1136/jmedgenet-2017-104620; pmid: 29021403
20. B. Gu, D. Eick, O. Bensaude, CTD serine-2 plays a critical role in splicing and termination factor recruitment to RNA polymerase II in vivo. *Nucleic Acids Res.* **41**, 1591–1603 (2013). doi: 10.1093/nar/gks1327; pmid: 23275552
21. L. Spraggon, L. Cartegni, U1 snRNP-dependent suppression of polyadenylation: Physiological role and therapeutic opportunities in cancer. *Int. J. Cell Biol.* **2013**, 846510 (2013). doi: 10.1155/2013/846510; pmid: 24285958
22. N. Meola *et al*., Identification of a nuclear exosome decay pathway for processed transcripts. *Mol. Cell* **64**, 520–533 (2016). doi: 10.1016/j.molcel.2016.09.025; pmid: 27871484
23. K. Ogami *et al*., An Mtr4/ZFC3H1 complex facilitates turnover of unstable nuclear RNAs to prevent their cytoplasmic transport and global translational repression. *Genes Dev.* **31**, 1257–1271 (2017). doi: 10.1101/gad.302604.117; pmid: 28733371
24. K. A. Hoadley *et al*., Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell* **173**, 291–304.e6 (2018). doi: 10.1016/j.cell.2018.03.022; pmid: 29625048
25. Cancer Genome Atlas Network, Genomic classification of cutaneous melanoma. *Cell* **161**, 1681–1696 (2015). doi: 10.1016/j.cell.2015.05.044; pmid: 26091043
26. D. Liu *et al*., Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nat. Med.* **25**, 1916–1927 (2019). doi: 10.1038/s41591-019-0654-5; pmid: 31792460
27. W. Hugo *et al*., Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* **165**, 35–44 (2016). doi: 10.1016/j.cell.2016.02.065; pmid: 26997480
28. A. Snyder *et al*., Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* **371**, 2189–2199 (2014). doi: 10.1056/NEJMoa1406498; pmid: 25409260
29. E. M. Van Allen *et al*., Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207–211 (2015). doi: 10.1126/science.aad0095; pmid: 26359337
30. E. Hodis *et al*., A landscape of driver mutations in melanoma. *Cell* **150**, 251–263 (2012). doi: 10.1016/j.cell.2012.06.024; pmid: 22817889
31. M. Krauthammer *et al*., Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. *Nat. Genet.* **44**, 1006–1014 (2012). doi: 10.1038/ng.2359; pmid: 22842228
32. E. M. Van Allen *et al*., The genetic landscape of clinical resistance to RAF inhibition in metastatic melanoma. *Cancer Discov.* **4**, 94–109 (2014). doi: 10.1158/2159-8290.CD-13-0617; pmid: 24265153
33. N. K. Hayward *et al*., Whole-genome landscapes of major melanoma subtypes. *Nature* **545**, 175–180 (2017). doi: 10.1038/nature22071; pmid: 28467829
34. A. Gonzalez-Perez, N. Lopez-Bigas, Functional impact bias reveals cancer drivers. *Nucleic Acids Res.* **40**, e169 (2012). doi: 10.1093/nar/gks743; pmid: 22904074
35. A. K. Greifenberg *et al*., Structural and functional analysis of the Cdk13/Cyclin K complex. *Cell Rep.* **14**, 320–331 (2016). doi: 10.1016/j.celrep.2015.12.025; pmid: 26748711
36. A. C. Carrera, K. Alexandrov, T. M. Roberts, The conserved lysine of the catalytic domain of protein kinases is actively involved in the phosphotransfer reaction and not required for anchoring ATP. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 442–446 (1993). doi: 10.1073/pnas.90.2.442; pmid: 8421674
37. J. Ablain *et al*., Human tumor genomics and zebrafish modeling identify *SPRED1* loss as a driver of mucosal melanoma. *Science* **362**, 1055–1060 (2018). doi: 10.1126/science.aau6509; pmid: 30385465
38. C. J. Ceol *et al*., The histone methyltransferase SETDB1 is recurrently amplified in melanoma and accelerates its onset. *Nature* **471**, 513–517 (2011). doi: 10.1038/nature09806; pmid: 21430779
39. E. E. Patton *et al*., BRAF mutations are sufficient to promote nevi formation and cooperate with p53 in the genesis of melanoma. *Curr. Biol.* **15**, 249–254 (2005). doi: 10.1016/j.cub.2005.01.031; pmid: 15694309
40. D. Blazek *et al*., The Cyclin K/Cdk12 complex maintains genomic stability via regulation of expression of DNA damage response genes. *Genes Dev.* **25**, 2158–2172 (2011). doi: 10.1101/gad.16962311; pmid: 22012619
41. K. Liang *et al*., Characterization of human cyclin-dependent kinase 12 (CDK12) and CDK13 complexes in C-terminal domain phosphorylation, gene transcription, and RNA processing. *Mol. Cell. Biol.* **35**, 928–938 (2015). doi: 10.1128/MCB.01426-14; pmid: 25561469
42. Q. Dai *et al*., Cyclin K-containing kinase complexes maintain self-renewal in murine embryonic stem cells. *J. Biol. Chem.* **287**, 25344–25352 (2012). doi: 10.1074/jbc.M111.321760; pmid: 22547058
43. S. Anders, A. Reyes, W. Huber, Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**, 2008–2017 (2012). doi: 10.1101/gr.133744.111; pmid: 22722343
44. B. Schwalb *et al*., TT-seq maps the human transient transcriptome. *Science* **352**, 1225–1228 (2016). doi: 10.1126/science.aad9841; pmid: 27257258
45. T. Silla, E. Karadoulama, D. Mąkosa, M. Lubas, T. H. Jensen, The RNA exosome adaptor ZFC3H1 functionally competes with nuclear export activity to retain target transcripts. *Cell Rep.* **23**, 2199–2210 (2018). doi: 10.1016/j.celrep.2018.04.061; pmid: 29768216
46. W. A. Khyrul, D. P. LaLonde, M. C. Brown, H. Levinson, C. E. Turner, The integrin-linked kinase regulates cell morphology and motility in a rho-associated kinase-dependent manner. *J. Biol. Chem.* **279**, 54131–54139 (2004). doi: 10.1074/jbc.M410051200; pmid: 15485819
47. M. J. May, R. B. Marienfeld, S. Ghosh, Characterization of the Ikappa B-kinase NEMO binding domain. *J. Biol. Chem.* **277**, 45992–46000 (2002). doi: 10.1074/jbc.M206494200; pmid: 12244103
48. C. G. Lian *et al*., Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell* **150**, 1135–1146 (2012). doi: 10.1016/j.cell.2012.07.033; pmid: 22980977
49. M. Ashburner *et al*., Gene ontology: Tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000). doi: 10.1038/75556; pmid: 10802651
50. G. Wu *et al*., A two-layered targeting mechanism underlies nuclear RNA sorting by the human exosome. *Cell Rep.* **30**, 2387–2401.e5 (2020). doi: 10.1016/j.celrep.2020.01.068; pmid: 32075771
51. C. Pak *et al*., Mutation of the conserved polyadenosine RNA binding protein, ZC3H14/dNab2, impairs neural function in *Drosophila* and humans. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 12390–12395 (2011). doi: 10.1073/pnas.1107103108; pmid: 21734151
52. J. Rha *et al*., The RNA-binding protein, ZC3H14, is required for proper poly(A) tail length control, expression of synaptic proteins, and brain function in mice. *Hum. Mol. Genet.* **26**, 3663–3681 (2017). doi: 10.1093/hmg/ddx248; pmid: 28666327
53. J. J. Allen *et al*., A semisynthetic epitope for kinase substrates. *Nat. Methods* **4**, 511–516 (2007). doi: 10.1038/nmeth1048; pmid: 17486086
54. K. J. Morris, A. H. Corbett, The polyadenosine RNA-binding protein ZC3H14 interacts with the THO complex and coordinately regulates the processing of neuronal transcripts. *Nucleic Acids Res.* **46**, 6561–6575 (2018). doi: 10.1093/nar/gky446; pmid: 29912477
55. M. Nováková *et al*., Mouse model of congenital heart defects, dysmorphic facial features and intellectual developmental disorders as a result of non-functional CDK13. *Front. Cell Dev. Biol.* **7**, 155 (2019). doi: 10.3389/fcell.2019.00155; pmid: 31440507
56. J. R. MacDonald, R. Ziman, R. K. Yuen, L. Feuk, S. W. Scherer, The Database of Genomic Variants: A curated collection of structural variation in the human genome. *Nucleic Acids Res.* **42** (D1), D986–D992 (2014). doi: 10.1093/nar/gkt958; pmid: 24174537
57. H. V. Firth *et al*., DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am. J. Hum. Genet.* **84**, 524–533 (2009). doi: 10.1016/j.ajhg.2009.03.010; pmid: 19344873
58. D. Cui, S. Wang, A. Zhang, A. Liu, Q. Hu, Case report: Hemophagocytic lymphohistiocytosis prior to the onset of leukemia in a boy with CDK13-related disorder. *Front. Genet.* **13**, 858668 (2022). doi: 10.3389/fgene.2022.858668; pmid: 35651941
59. Z. Fan *et al*., CDK13 cooperates with CDK12 to control global RNA polymerase II processivity. *Sci. Adv.* **6**, eaaz5041 (2020). doi: 10.1126/sciadv.aaz5041; pmid: 32917631
60. S. H. Lee *et al*., Widespread intronic polyadenylation inactivates tumour suppressor genes in leukaemia. *Nature* **561**, 127–131 (2018). doi: 10.1038/s41586-018-0465-8; pmid: 30150773
61. B. R. So *et al*., A complex of U1 snRNP with cleavage and polyadenylation factors controls telescripting, regulating mRNA transcription in human cells. *Mol. Cell* **76**, 590–599.e4 (2019). doi: 10.1016/j.molcel.2019.08.007; pmid: 31522989
62. S. Shuai *et al*., The U1 spliceosomal RNA is recurrently mutated in multiple cancers. *Nature* **574**, 712–716 (2019). doi: 10.1038/s41586-019-1651-z; pmid: 31597163
63. R. Kumar *et al*., THOC2 mutations implicate mRNA-export pathway in X-linked intellectual disability. *Am. J. Hum. Genet.* **97**, 302–310 (2015). doi: 10.1016/j.ajhg.2015.05.021; pmid: 26166480
64. S. R. Viswanathan *et al*., Structural alterations driving castration-resistant prostate cancer revealed by linked-read genome sequencing. *Cell* **174**, 433–447.e19 (2018). doi: 10.1016/j.cell.2018.05.036; pmid: 29909985
65. Cancer Genome Atlas Research Network, Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609–615 (2011). doi: 10.1038/nature10166; pmid: 21720365
66. K. M. Ekumi *et al*., Ovarian carcinoma CDK12 mutations misregulate expression of DNA repair genes via deficient formation and function of the Cdk12/CycK complex. *Nucleic Acids Res.* **43**, 2575–2589 (2015). doi: 10.1093/nar/gkv101; pmid: 25712099
67. Data for: M. L. Insco *et al*., Oncogenic CDK13 mutations impede nuclear RNA surveillance, Zenodo (2023); https://zenodo.org/badge/latestdoi/577358557.

ACKNOWLEDGMENTS

The CDK13 antibody used for mouse experiments and to detect truncated CDK13 was a gift from A. L. Greenleaf. The pAC4 and PBNeoTetO-Dest vectors were gifts from A. Cheng. We acknowledge the Harvard Nascent Transcriptomics Core for assistance with TT-seq. The human cancer cell template used in fig. S4 was adapted from BioRender. Analyses presented here are in part based upon data generated by the TCGA Research Network (https://www.cancer.gov/tcga). This work used the Biowulf cluster computing system at the National Institutes of Health (NIH). **Funding:** This work was supported by the Damon Runyon Cancer Foundation (Fellowship Award to M.I.); the American Society of Clinical Oncology (Young Investigator Award to M.I.); the NIH [National Heart, Lung and Blood Institute grant T32HL116324 to M.I.; grant T32 GM135134 to D.E.E.; grant R35 GM144283 to R.A.Y.; grants P01CA042063, R01GM034277, and R01CA133404 to P.A.S., S.J.D, and P.B.; grant R01CA103846 to L.I.Z.; grant R01GM141544 to P.L.B.; National Cancer Institute (NCI) grant K08CA248727 to M.I.; and Intramural Research Program (IRP) of the Division of Cancer Epidemiology and Genetics, NCI, NIH, grant to T.Z., K.M.B., and S.J.C.)]; the Charles S. Memorial Sloan Kettering Starr Foundation Cancer Consortium; Hope Funds for Cancer Research Grillo-Marxuach Family Fellowship (B.J.A.); the American Lebanese Syrian Associated Charities (ALSAC) (B.J.A.); the University of Rochester Health Sciences Center (computational innovation high performance computational resource grant to P.L.B.); the David H. Koch Foundation (fellowship to S.J.D.); the Ludwig Center at Harvard (K.A. and B.E.J.M.); and the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy (EXC2151–390873048 and grant GE 976/9-2 to M.G.). **Author contributions:** M.L.I. planned and performed the experiments with assistance from K.Y.C., T.F., A.M.C., R.M., and C.W. with the following exceptions. S.J.D. planned and performed the mouse ES cell experiments and I.H.K. and S.D. planned and performed the kinase assays. M.L.I., B.J.A., P.L.B., S.B., B.E.J.M., C.G.L., T.Z., and D.E.E. analyzed the data. D.L. coded the survival analysis. M.G. provided the plots of the mutations on the CDK13 crystal structure and insight into the effects the mutations could have on CDK13 function. T.H. advised on TT-seq experiments. R.A.Y., P.A.S., P.L.B., K.A., M.G., K.M.B., S.J.C., and L.I.Z. supervised the project. M.L.I., B.J.A., P.L.B., and L.I.Z. wrote the manuscript. **Competing interests:** R.A.Y. is a founder and shareholder of Syros Pharmaceuticals, CAMP4 Therapeutics, Omega Therapeutics, and Dewpoint Therapeutics. P.A.S. is a shareholder of Syros Pharmaceuticals. L.I.Z. is a founder and stockholder of Fate Therapeutics, CAMP4 Therapeutics, and Scholar Rock and is a consultant for Celularity. B.J.A. is a shareholder in Syros Pharmaceuticals. K.A. consults for Syros Pharmaceuticals and is on the scientific advisory board of CAMP4 Therapeutics. P.L.B. is a paid scientific consultant for Skyhawk Therapeutics, Inc. L.I.Z. and M.L.I. are inventors on patent application no. PCT/US2022/037839 filed on 07-21-2022 by the Children's Medical Center Corporation titled "Selective nuclear RNA degradation as a therapeutic mechanism." The remaining authors declare no competing interests. **Data and materials availability:** The datasets generated and/or analyzed during the present study have been uploaded to the Gene Expression Omnibus (GEO Superseries GSE131334). Newly

generated code was deposited at Zenodo (*67*). The TCGA patient RNA-seq bams are controlled-access data with unique identifiers so they cannot be publicly shared. All zebrafish strains, cell lines, and DNA vectors are readily available through the corresponding author. Where possible, we can share remaining zebrafish tumor material; however, this material is limited in abundance. All antibodies are commercially available. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn7625
Materials and Methods
Figs. S1 to S5
Tables S1 to S8
References (*68*–*87*)
MDAR Reproducibility Checklist

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

Submitted 17 December 2021; resubmitted 30 August 2022
Accepted 27 February 2023
10.1126/science.abn7625

RESEARCH ARTICLE SUMMARY

GENOME EDITING

Base editing rescue of spinal muscular atrophy in cells and in mice

Mandana Arbab†, Zaneta Matuszek†, Kaitlyn M. Kray, Ailing Du, Gregory A. Newby, Anton J. Blatnik, Aditya Raguram, Michelle F. Richter, Kevin T. Zhao, Jonathan M. Levy, Max W. Shen, W. David Arnold, Dan Wang, Jun Xie, Guangping Gao, Arthur H. M. Burghes, David R. Liu*

INTRODUCTION: Spinal muscular atrophy (SMA) is the leading genetic cause of infant mortality. SMA results from survival motor neuron (SMN) protein insufficiency after homozygous loss of the *SMN1* gene. A closely related gene, *SMN2*, differs from *SMN1* by a C6T substitution (i.e., a C-to-T transition at position 6) in exon 7 that results in a truncated SMNΔ7 protein that fails to fully compensate for *SMN1* loss. Two recently approved SMA drugs partially restore SMN protein levels through splice isoform switching. A third drug uses viral gene complementation to restore SMN levels. Although up-regulation of SMN levels by these approved drugs effectively treats SMA, current therapies circumvent endogenous regulation of SMN, do not fully restore SMN levels, and either require repeated dosing or may fade over time. A one-time, permanent treatment that restores endogenous gene expression and preserves native SMN regulation may address these limitations of existing SMA therapies.

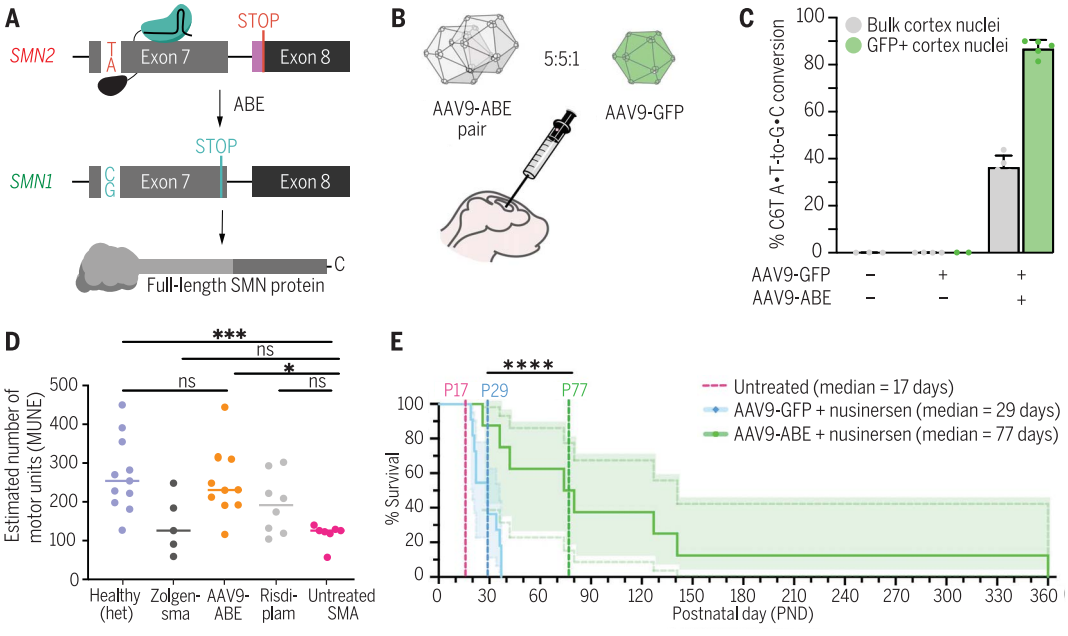
RATIONALE: Genome editing of *SMN2*, which is present in all SMA patients, could enable a one-time treatment for SMA that restores normal SMN transcript and protein levels while preserving their endogenous regulatory mechanisms. We developed one-time genome editing approaches targeting endogenous *SMN2* that restore SMN protein abundance to normal levels and rescue disease phenotypes in cell and mouse models of SMA. We tested 79 base editing and nuclease strategies that modify five posttranscriptional and posttranslational regulatory regions in *SMN2* to increase SMN protein levels.

RESULTS: Each of the *SMN2* nuclease and base editing strategies tested durably increased SMN protein levels between 9- and 50-fold. Base editing efficiently converted *SMN2* to *SMN1* genes and, unlike nuclease editing strategies or current SMA drugs, fully restored SMN transcript and protein levels to those of wild-type cells (~40-fold increase) with minimal off-target editing across the genome and transcriptome. Intracerebroventricular injection of adeno-associated virus serotype 9 encoding an adenine base editor (AAV9-ABE) resulted in 87% average conversion of *SMN2* C6T among transduced cells in the central nervous system of Δ7SMA mice, improved motor function, and extended life span, despite Δ7SMA mice having a much shorter window for treatment t[o] human patients (≤6 days for mice versus mo[nths] to years for humans) that ends earlier than typical in vivo base editing time scales (weeks). One-time in vivo coadministration of AAV9-ABE with the antisense oligonucleotide drug nusinersen expanded the therapeutic window for gene correction, further improving the life span of AAV9-ABE–treated animals to an average of 111 days, compared with an average of 17 days for untreated animals.

CONCLUSION: Despite the incongruent timeline of base editing–mediated rescue for ideal rescue of Δ7SMA mice, AAV9-ABE treatment yielded substantial improvements in life span and motor function. Combination treatment with nusinersen enables Δ7SMA mouse rescue that resembles presymptomatic up-regulation of SMN levels. In humans, the therapeutic window is much longer. Therefore, we anticipate that AAV9-ABE may achieve presymptomatic rescue as a standalone therapeutic in SMA patients. Our study also demonstrates the compatibility of base editing with nusinersen, which may inform future clinical applications. Together, these findings support the potential of base editing as a future one-time treatment for SMA that restores native SMN production while preserving endogenous regulatory mechanisms of SMN expression. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: drliu@fas.harvard.edu
†These authors contributed equally to this work.
Cite this article as M. Arbab *et al.*, *Science* **380**, eadg6518 (2023). DOI: 10.1126/science.adg6518

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adg6518



Base editing of *SMN2* rescues SMA in mice. (**A**) A customized ABE converts insufficient *SMN2* genes into healthy *SMN1* genes to produce full-length SMN protein. (**B**) Dual-AAV9–mediated delivery of ABE and green fluorescent protein (GFP) into SMA neonates. (**C**) In vivo conversion of *SMN2* C6T in the central nervous system of treated animals. (**D**) Motor unit number estimation (MUNE) in SMA mouse muscle after base editing treatment. het, heterozygous. (**E**) Survival of SMA mice after base editing treatment. ns, not significant. *$P \leq 0.02$, ***$P \leq 0.005$, ****$P \leq 0.001$.

RESEARCH ARTICLE

GENOME EDITING

Base editing rescue of spinal muscular atrophy in cells and in mice

Mandana Arbab[1,2,3,4]†, Zaneta Matuszek[3,4,5]†, Kaitlyn M. Kray[6], Ailing Du[7], Gregory A. Newby[3,4], Anton J. Blatnik[6], Aditya Raguram[3,4], Michelle F. Richter[3,4], Kevin T. Zhao[3,4], Jonathan M. Levy[3,4], Max W. Shen[3,4,8], W. David Arnold[9,10], Dan Wang[7,11], Jun Xie[7], Guangping Gao[7,12], Arthur H. M. Burghes[6], David R. Liu[3,4,13]*

Spinal muscular atrophy (SMA), the leading genetic cause of infant mortality, arises from survival motor neuron (SMN) protein insufficiency resulting from *SMN1* loss. Approved therapies circumvent endogenous SMN regulation and require repeated dosing or may wane. We describe genome editing of *SMN2*, an insufficient copy of *SMN1* harboring a C6>T mutation, to permanently restore SMN protein levels and rescue SMA phenotypes. We used nucleases or base editors to modify five *SMN2* regulatory regions. Base editing converted *SMN2* T6>C, restoring SMN protein levels to wild type. Adeno-associated virus serotype 9–mediated base editor delivery in Δ7SMA mice yielded 87% average T6>C conversion, improved motor function, and extended average life span, which was enhanced by one-time base editor and nusinersen coadministration (111 versus 17 days untreated). These findings demonstrate the potential of a one-time base editing treatment for SMA.

Spinal muscular atrophy (SMA) is a progressive motor neuron disease and the leading genetic cause of infant mortality (*1*–*3*). SMA is caused by homozygous loss or mutation of the essential survival motor neuron 1 (*SMN1*) gene. One or more copies of the nearly identical (>99.9% sequence identity) *SMN2* gene partially compensate for the loss of *SMN1* (*1*, *4*, *5*). However, *SMN1* and *SMN2* differ by a silent C•G-to-T•A substitution at nucleotide position 6 of exon 7 (C6T) that results in exon 7 skipping in mRNA transcripts (Fig. 1A) (*6*, *7*). The resulting truncated SMNΔ7 protein is rapidly degraded, causing SMN protein insufficiency that results in loss of motor neurons, paralysis, and death (*8*–*10*). Untreated patients with the most common form of SMA (type I) live a median of 6 months (*11*, *12*).

Up-regulation of SMN protein can rescue motor function and substantially improve the prognosis of SMA patients (*13*–*15*). However, endogenous SMN protein is subject to multiple levels of regulation that differs across tissues (*16*–*18*). Whereas *SMN* underexpression can fail to rescue SMN phenotypes, *SMN* overexpression can cause aggregation, toxicity, and tissue pathology (*19*–*21*). Three breakthrough therapeutics effectively rescue many SMA phenotypes and improve life span by up-regulating SMN protein (*22*). The antisense oligonucleotide (ASO) nusinersen (Spinraza) and the small-molecule risdiplam (Evrysdi) both promote splicing inclusion of exon 7, resulting in ~2-fold up-regulation of SMN levels, and have proven highly effective in the clinic (*23*, *24*). However, SMN protein is reduced by ~ 85% in the spinal cord of untreated SMA patients (*25*–*27*). The partial recovery of SMN protein promoted by these therapeutics may be insufficient at early time points and in damaged tissues, potentially underlying the limited rescue observed in some patients (*28*, *29*). Moreover, the transient nature of these therapeutics necessitates repeated administration of costly drugs throughout patients' lifetimes (*30*, *31*).

Adeno-associated virus (AAV)–mediated gene complementation of full-length SMN cDNA by onasemnogene abeparvovec-xioi (Zolgensma) leads to constitutive production of SMN in transduced cells that is not under endogenous control (*32*–*34*). In the spinal cord, Zolgensma up-regulates *SMN* transcript levels by ~25% (*35*), while in other tissues such as the liver and dorsal root ganglia, gene complementation may cause *SMN* overexpression that under some circumstances can cause long-term toxicity (*21*). We do not yet know whether *SMN* overexpression induces toxicity in patients treated with Zolgensma or how long AAV-mediated expression will persist in motor neurons in patients (*36*, *37*). As such, a therapeutic modality that restores endogenous gene expression and preserves native SMN regulation by a one-time permanent treatment may address remaining limitations of existing SMA therapies. Genome editing of *SMN2*, which is present in all SMA patients regardless of the nature of their *SMN1* mutation, could enable a one-time treatment for SMA that restores native SMN transcript and protein levels while preserving their endogenous regulatory mechanisms.

Results

Predictable and precise nuclease editing of SMN2 *ISS-N1 increases SMN protein levels*

SMN protein production from *SMN1* and *SMN2* genes is constrained by transcriptional, transcriptomic, and posttranslational regulatory sequences. We explored using Cas nucleases to create gain-of-function alleles in *SMN2* regulatory sequences that up-regulate SMN levels. The inclusion of exon 7, which underlies SMN protein stability, is strongly influenced by the downstream intronic splicing silencer ISS-N1 that harbors two heterogeneous nuclear ribonucleoprotein (hnRNP) A1/A2 binding sites (Fig. 1A) (*38*). Deletions within and downstream of the 3′ hnRNP A1/A2 binding domain improve exon 7 inclusion (*38*–*41*). We speculated that Cas9 nuclease–mediated disruption of the ISS-N1 genomic locus might increase exon 7 inclusion in *SMN2* splicing and thereby increase SMN protein levels (strategy A, Fig. 1B).

We used inDelphi, a machine learning model of SpCas9 nuclease editing outcomes, to predict insertion and deletion (indel) outcomes at the ISS-N1 locus that disrupt hnRNP A1/A2 binding and improve full-length SMN splicing of *SMN2* (Fig. 1B) (*42*). InDelphi identified 10 spacer sequences predicted to induce ≥4-nucleotide (nt) deletions at ISS-N1 and loss of ≥1 nt of the 3′ hnRNP A1/A2 domain ("predicted % precision"). We estimated editing efficiencies of these strategies on the basis of the reported protospacer adjacent motif (PAM) compatibility of these spacer sequences with SpCas9-variant nucleases ("predicted % PAM efficiency") (*43*–*46*). From 19 possible nuclease editing strategies (A1 to A19, defined as different combinations of genome editing agents and guide RNAs), we selected nine (A2, A3, A5, A6, A13, A14, A16, A17, and A19) for experimental testing.

We cotransfected Δ7SMA mouse embryonic stem cells (mESCs)—which lack endogenous

[1]Department of Neurology, Rosamund Stone Zander Translational Neuroscience Center, Boston Children's Hospital, Boston, MA 02115, USA. [2]Department of Neurobiology, Harvard Medical School, Boston, MA 02115, USA. [3]Merkin Institute of Transformative Technologies in Healthcare, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA. [4]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA. [5]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA. [6]Department of Biological Chemistry and Pharmacology, The Ohio State University Wexner Medical Center, Columbus, OH 43210, USA. [7]Horae Gene Therapy Center, UMass Chan Medical School, University of Massachusetts, Worcester, MA 01605, USA. [8]Computational and Systems Biology Program, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [9]Department of Neurology, The Ohio State University Wexner Medical Center, Columbus, OH 43210, USA. [10]NextGen Precision Health, University of Missouri, Columbia, MO 65212, USA. [11]RNA Therapeutics Institute, UMass Chan Medical School, University of Massachusetts, Worcester, MA 01605, USA. [12]Microbiology and Physiological Systems, UMass Chan Medical School, University of Massachusetts, Worcester, MA 01605, USA. [13]Howard Hughes Medical Institute, Harvard University, Cambridge, MA 02138, USA.
*Corresponding author. Email: drliu@fas.harvard.edu
†These authors contributed equally to this work.

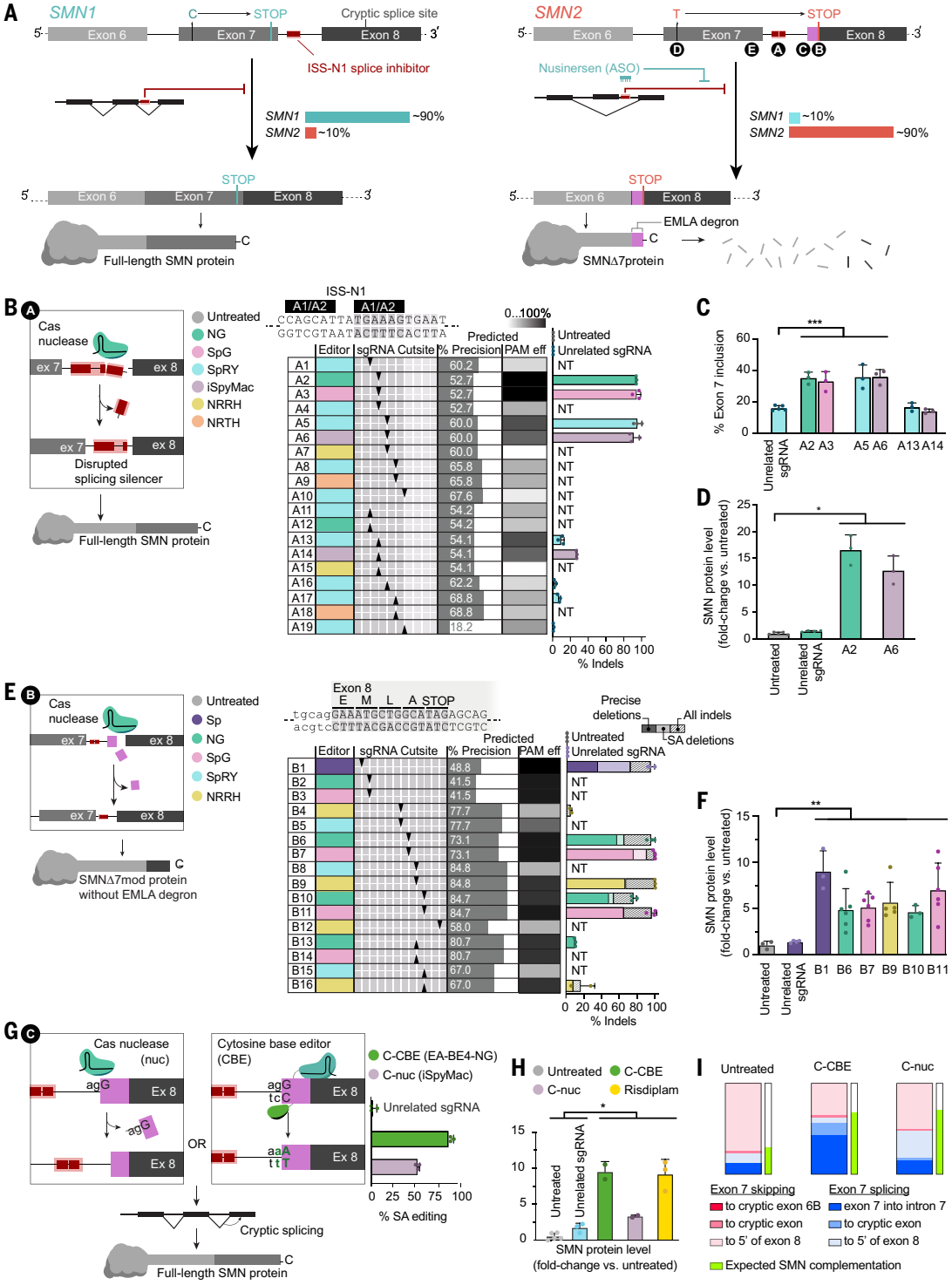


Fig. 1. Editing *SMN2* regulatory regions. (**A**) Schematic diagrams of genomic *SMN* exons 6 to 8 and *SMN* mRNA and protein products. (**B**) Nuclease editing strategy and genome editing outcomes of ISS-N1 targeting (strategy A). The table shows combinations of six nucleases paired with 10 sgRNAs complementary to the top strand (A1 to A10) or bottom strand (A11 to A19) identified by arrows that show the DSB site of the sgRNAs relative to the sequence above. PAM eff, protospacer adjacent motif efficiency. (**C**) Exon 7 inclusion in *SMN* mRNA after editing, as indicated, measured by automated electrophoresis. (**D**) SMN protein levels after editing, as indicated, normalized to histone H3. (**E**) Nuclease editing strategy targeting and genome editing outcomes of targeting the first five codons of exon 8 (strategy B). The table shows combinations of five nucleases paired with nine sgRNAs complementary to the top strand (B1 to B12) or bottom strand (B13 to B16) identified by arrows that indicate their DSB site, as in (B). (**F**) Total SMN protein levels after editing. (**G**) Nuclease and cytosine base editing strategies and genome editing outcomes of 3′ splice acceptor disruption at exon 8 (strategy C). (**H**) SMN protein levels after C-nuc and C-CBE editing or treatment with risdiplam, normalized to histone H3. (**I**) Distribution of *SMN2* transcript variants after C-nuc and C-CBE editing. Experiments were performed in Δ7SMA mESCs. NT, no treatment; *$P \le 0.05$, **$P \le 0.01$, ***$P \le 0.005$. Error bars indicate standard deviations.

Smn1, are homozygous for the full-length human *SMN2* gene, carry human *SMNΔ7*-cDNA transgenes, and harbor a *Mnx1:GFP* reporter of motor neurons (*SMN2*$^{+/+}$; *SMNΔ7*; *Smn*$^{-/-}$; *Mnx1:GFP*) (*47*)—with nuclease expression plasmids that carry a blasticidin-resistance cassette and single-guide RNA (sgRNA) plasmids that carry a hygromycin-resistance cassette. Both plasmids also contain Tol2 transposase sequences to enable stable transposon-mediated genomic integration and antibiotic selection. We achieved 92 ± 5.6% average indel frequencies for the top four strategies targeting the ISS-N1 locus (A2, A3, A5, and A6) (Fig. 1B).

To assess whether nuclease-mediated editing of ISS-N1 improved exon 7 inclusion, we performed reverse transcription PCR (RT-PCR) of *SMN2* from exons 6 to 8 and quantified *SMNΔ7* and full-length *SMN* products (Fig. 1C). We found that all strategies that edited ISS-N1 with high efficiency (≥85%) resulted in a significant increase in exon 7 inclusion averaging 2.2-fold relative to an unrelated sgRNA control (Welch's two-tailed *t* test, $P = 0.01$). The increase in exon 7 inclusion caused a substantial increase in SMN protein of 17-fold by A2 and 13-fold by A6 relative to untreated controls

(values normalized to histone H3, Welch's two-tailed t test, $P = 0.02$; Fig. 1D and fig. S1A). Collectively, these results demonstrate that disruption of the ISS-N1 genomic locus can stably increase full-length SMN splicing and protein phenotypes of SMA.

Predictable and precise genome editing of SMN2 *exon 8 increases SMN protein levels*

In an alternative nuclease-mediated approach to up-regulate SMN levels, we disrupted post-translational regulatory sequences in *SMN2* to increase SMNΔ7 protein stability. The critical difference between full-length SMN and the unstable SMNΔ7 protein is the substitution of 16 amino acids encoded by exon 7 with EMLA, a four-residue degron encoded by exon 8 (Fig. 1A) (*8*). Extending the coding sequence of exon 8 with five or more heterologous amino acids obscures SMNΔ7 C-terminal degradation signals. These modified SMNΔ7 (SMNΔ7mod) protein variants have increased stability and rescue survival and motor phenotypes of severe SMA mice (*48*). We designed strategies for Cas nuclease–mediated disruption of exon 8 to generate similar stabilized SMNΔ7mod proteins with therapeutic potential (strategies B1 to B16; supplementary text and Fig. 1E) and observed up to a 7.0-fold increase in SMN protein levels by B11 (Welch's two-tailed t test, $P = 0.007$; Fig. 1F and fig. S1B).

Some exon 8 editing strategies improved SMN protein stability more than expected given the observed edited genotypes (Fig. 1, E and F). For example, precision-edited genotypes were 1.9-fold higher in frequency after B9 editing than B1, yet SMNΔ7mod protein levels were greater in cells edited with B1 (9.1-fold) than B9 (5.7-fold). These data suggest that additional edited genotypes may improve SMN protein stability. Inspection of the non-precisely edited fraction of edited alleles revealed that B1 editing frequently induces indels at the exon 8 splice acceptor. Thus, we hypothesized that disrupting splicing of exon 8 improves SMN protein stability (*49*).

To test this hypothesis, we disrupted the canonical AG splice acceptor (SA) motif of exon 8 using either a nuclease or cytosine base editor (C-nuc or C-CBE, respectively) in Δ7SMA mESCs (Fig. 1G) (*45*, *50*) and observed 54 ± 2.3% indels from C-nuc and 89 ± 2.3% cytosine base editing from C-CBE. Notably, C-nuc editing resulted in a complex mixture of indel genotypes at the intron-exon junction that resulted in deletion of additional nucleotides beyond the AG motif. Both strategies significantly increased SMN levels in Δ7SMA mESCs, similar to treatment with risdiplam (3.3-fold for C-nuc, 9.5-fold for C-CBE, and 9.1-fold for risdiplam relative to untreated; Welch's two-tailed t test, $P < 0.05$; Fig. 1H and fig. S1, D to G), indicating that alternative splicing at exon 8 improves the stability of *SMN2* gene products.

We investigated how exon 8 SA disruption affects *SMN2* transcripts (supplementary text). C-CBE editing induced a minor increase in *SMN2* mRNA that only partially explains the 9.5-fold increase in SMN levels (fig. S1H). We also observed a profound shift in *SMN2* splice products (Fig. 1I). We investigated whether these alternative splice isoforms improve stability of SMN proteins and found that transcripts including exon 7 were increased twofold by C-CBE (63 ± 2.0%) and 1.6-fold by C-nuc (50 ± 1.1%) relative to untreated cells (24 ± 1.4%). These transcripts often retain intron 7 as in some functional transcript variants of *SMN2* (ENST00000511812.5, fig. S1I). Notably, all transcripts that include exon 7 encode full-length SMN protein and can therefore complement loss of *SMN1*. Thus, the substantial increase in SMN protein levels after exon 8 SA editing predominantly arises from an increase in normal full-length SMN.

Collectively, the tested *SMN2* editing strategies permanently increase SMN protein levels up to 17-fold (strategy A2), 9.1-fold (strategy B1), and 9.5-fold (strategy C-CBE). As a 1.5- to 2-fold increase in SMN protein is therapeutic for SMA patients (*23*, *24*), these strategies represent promising approaches for further studies.

Efficient and precise base editing of SMN2 *splice regulatory elements*

Several single-nucleotide substitutions in exon 7 strongly regulate splicing of *SMN2*, including the C-to-T transition at position 6 (C6T) that differentiates *SMN1* (C) from *SMN2* (T) genes (Fig. 1A), and T44C, G52A, and A54G at the 3′ end of exon 7 (*51*). Using existing and newly developed BE-Hive predictive models of base-editing outcomes (supplementary text and fig. S2, A to E), we identified 42 strategies (combinations of base editors and guide RNAs) to modify exon 7 splicing regulatory elements (SREs) (Fig. 2, A to C, and fig. S2, F and G). We designed 13 spacers targeting C6T using ABE8e (strategies D1 to D19) or targeting C6T, T44C, G52A, and A54G using ABE8e, ABE7.10, and EA-BE4 deaminases (strategies E1 to E23). We paired these spacers with 12 compatible SpCas9 variants on the basis of reported PAM preferences ("predicted % PAM efficiency") (*43*, *46*, *50*). We validated these strategies in Δ7SMA mESCs and found that the BE-Hive models of SpCas9 base editors predicted edited outcomes of Cas-variant base editors with high accuracy [Cas9-NG (*46*), NRTH, NRRH, and NRCH (*44*), Pearson's correlation coefficient (r) = 0.810; chimeric SpyMac and iSpMac (*45*), Pearson's $r = 0.910$; supplementary text and Fig. 2D].

Base editing of exon 7 SREs was highly efficient. At 3′ SREs, we achieved 69 ± 5.0% T44C editing by E14, 92 ± 4.0% G52A editing by E20, and 95 ± 5.1% A54G editing by E23 (fig. S2, F and G). We achieved nearly complete (94 to 99.5%) C6T A•T-to-G•C conversion by strategies targeting C6T at positions *P5* (D1 and D2), *P8* (D10 and D11), and *P10* (D18 and D19) within the protospacer (Fig. 2, A to C). The deaminase in ABE7.10 enabled up to 64 ± 2.5% conversion of T6>C (E7, fig. S2G) (*52*, *53*).

The frequency of edited alleles with single-nucleotide T6>C conversion alone (i.e., without any bystander edits or indels) varied substantially between the most efficient C6T editing strategies, ranging from 82 ± 1.9% for D10 editing to 40 ± 13% for D19 editing (Fig. 2E). Prior studies suggest that the coding sequence at the SMN C terminus beyond exon 6 does not strongly affect SMN protein function, and it is therefore unlikely that single-nucleotide editing precision of C6T is imperative for rescue of SMA (*8*, *48*, *54*). Maximizing the sequence similarity of modified *SMN2* genes to native *SMN1*, however, may preserve additional regulatory interactions, including those not yet known. D10, the strategy with the highest precision and efficiency (99 ± 0.7%), did not induce measurable indels, and its induced bystander missense nucleotide substitutions (18 ± 2.4%) have previously been shown to benefit inclusion of exon 7 by improved protein binding at the exonic splicing enhancer (fig. S2H) (*55*, *56*). Together, these results establish efficient base editing strategies to convert *SMN2* T6>C with high fidelity and few undesirable by-products.

Base editing of SMN2 *splice regulatory elements rescues SMN protein levels*

Next, we sought to determine whether base editing of exon 7 SREs results in functional rescue of cellular SMA phenotypes. The top six ABE8e editing strategies that converted C6T in >97% of alleles increased exon 7 inclusion to 78 ± 10.2% on average, up to 9.7-fold higher than untreated cells (87 ± 1.5% by D10 compared with 9.0 ± 6.6% in untreated; Welch's two-tailed t test, $P < 0.002$; Fig. 2F). These results are on par with, or exceed, maximum exon 7 inclusion by risdiplam or nusinersen treatment of Δ7SMA mESCs (89 ± 4.3% and 80 ± 0.3%, respectively; Fig. 2F and fig. S1E) and resemble splicing ratios of *SMN1* genes (82 ± 7.3% in U2OS cells) (*38*, *39*). Base editing of 3′ SREs in exon 7 also improved inclusion, averaging 60 ± 3.2% after T44C editing by E14, 76 ± 12% after G52A editing by E20, and 50 ± 8.6% after A54G editing by E23 (fig. S2I). These data demonstrate that base editing of various exon 7 SREs can increase full-length *SMN* splice products.

Base editing of 3′ SREs increased SMN protein levels in ways that did not closely mirror observed improvements in exon 7 inclusion. We detected a 3.4-fold increase in SMN protein by E14 base editing of T44C, a 23-fold increase by E20 editing of G52A, and a 1.6-fold

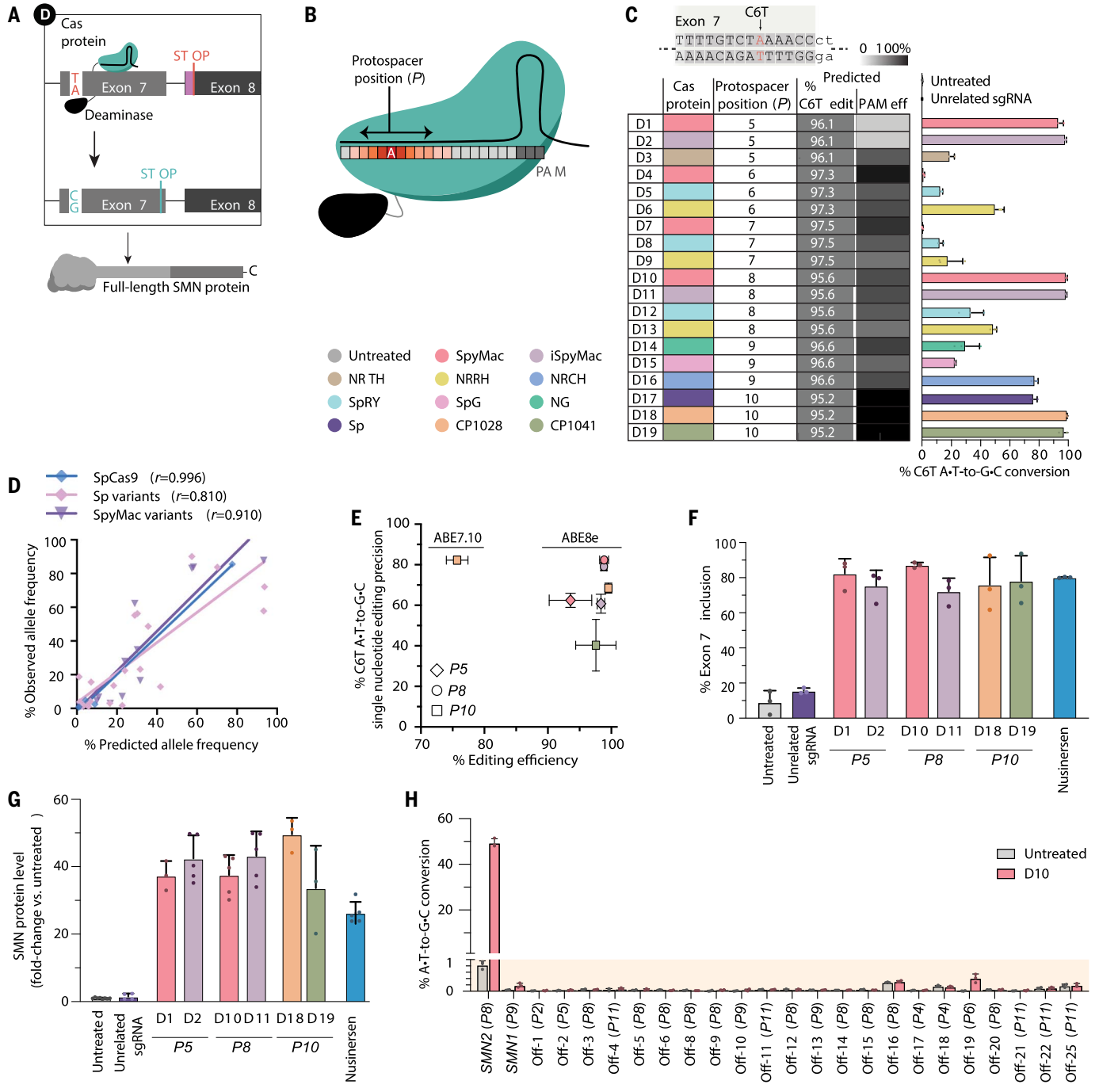


Fig. 2. Adenine base editing of *SMN2* C6T. (**A**) Adenine base editing of *SMN2* C6T (strategy D). (**B**) Target nucleotide position within the protospacer (*P#*) for base editing. A typical base editor activity window is illustrated as a heatmap. (**C**) The table shows ABE8e editing strategies with color-coded Cas-variant domains and their corresponding spacers. The protospacer position of the C6T target nucleotide (*P#*) is indicated. Graph shows genome editing outcomes in Δ7SMA mESCs. (**D**) Correlation of BE-Hive predicted editing outcomes with observed allele frequencies after base editing with ABE7.10 or ABE8e deaminases fused to different Cas variants. Pearson's *r* is shown; 95% confidence interval (CI) ranges are 0.9408 to 0.9998 for SpCas9, 0.5823 to 0.9201 for SpCas9 engineered and evolved variants, and 0.7557 to 0.9689 for SpyMac Cas variants. (**E**) Plot of base editing efficiency and single-nucleotide editing precision of C6T by the indicated ABE and spacer combinations. (**F**) Exon 7 inclusion in *SMN* mRNA after editing by the indicated strategies, measured by automated electrophoresis. (**G**) SMN protein levels after editing by the indicated strategies, normalized to histone H3. (**H**) On-target and off-target base editing of strategy D10 in HEK293T cells. Bars show editing of the most frequently edited nucleotide at each locus, with the *P#* position shown in parentheses. Error bars indicate standard deviations.

increase by E23 editing of A54G (Welch's two-tailed *t* test, $P = 0.02$), despite all three edits inducing comparable improvements in exon 7 inclusion (figs. S2I and S3, A and B). We hypothesized that unintended bystander edits may underlie this persistent protein instability and found that the T44C and A54G editing strategies frequently ablate the nearby TAA stop codon in exon 7 (fig. S2, F and G). A failure to terminate translation in exon 7 leads to the extension of full-length SMN proteins with the EMLA degron encoded by exon 8

(Fig. 1A). Thus, imprecise editing of T44C or A54G by E14 or E23 results in the translation of unstable full-length SMN-EMLA fusions that prevent up-regulation of SMN protein levels. Editing of G52A by E20 uses the EA-BE4 cytosine deaminase that does not recognize TAA as a substrate and therefore does not induce nonsilent bystander changes in 99 ± 0.1% of edited alleles, resulting in a 23-fold increase in SMN protein levels.

Base editing of exon 7 C6T resulted in the greatest up-regulation of SMN protein. The top six ABE8e editing strategies that correct C6T in >97% of alleles induced a 41-fold average increase in SMN protein levels compared with untreated controls (normalized to H3, Welch's two-tailed t test, $P < 0.0002$; Fig. 2G and fig. S3C), indicating complete rescue of normal SMN protein levels in Δ7SMA mESCs, which are reduced >95% relative to wild-type mESCs (*47*). Despite inducing a comparable increase in exon 7 inclusion, base editing of C6T enabled a 4.5-fold and 1.5-fold greater increase in SMN protein levels than risdiplam and nusinersen treatment of Δ7SMA mESCs (9-fold and 17-fold, respectively, compared with 41-fold on average across the top six strategy D approaches; Figs. 1H and 2G and figs. S1, D, F, and G, and S3, C, E, and F). Normal levels of SMN protein are essential to the function, survival, and long-term health of all species in the animal kingdom (*57*–*60*). Restoring wild-type levels of SMN protein as achieved through a base editing strategy may thus best maximize the long-term health of SMA patients.

Among all genome editing strategies tested, base editing of C6T by D10 induces the greatest increase in exon 7 inclusion (87 ± 1.5%) and best recapitulates native SMN protein levels (95% of wild-type levels, a 38-fold increase versus untreated Δ7SMA mESCs). D10 base editing is highly efficient (99 ± 0.7%) with high on-target precision (82 ± 0.0%). The *SMN2* gene arose from a duplication of the chromosomal region containing *SMN1* and shares an identical promoter and >99.9% sequence identity with *SMN1*, including 100% DNA conservation of its protein-coding sequence other than exon 7 C6T (*1*, *4*, *5*). We performed reverse transcription–quantitative PCR and quantified *SMN2* mRNA levels in edited cells, confirming that *SMN2* mRNA abundance is not affected by D10 base editing compared with untreated Δ7SMA mESCs or after ABE8e transfection with an unrelated sgRNA (fig. S3G). Together, these data indicate that D10 editing of *SMN2* faithfully reproduces the genomic sequence and function of native *SMN1* alleles. Therefore, we selected strategy D10 for further study.

Off-target editing analysis of ABE8e targeting SMN2 *C6T in the human genome*

Some base editors can induce off-target deamination in cells, including Cas-dependent off-target DNA editing and Cas-independent off-target DNA or RNA editing (*61*–*65*). Genomic and transcriptomic off-target deamination by adenine base editors without involvement of the Cas protein component is rare, and deaminase variants that minimize these events have been reported (*61*, *66*). We assessed the Cas-dependent genome specificity of the D10 strategy (ABE8e-SpyMac and *P8* sgRNA) characterizing SpyMac Cas9 nuclease with *P8* sgRNA using CIRCLE-seq (circularization for in vitro reporting of cleavage effects by sequencing) (*67*), an unbiased and sensitive empirical in vitro off-target detection method. Potential off-target sites nominated by CIRCLE-seq can then be sequenced in-depth in base-edited human cells to provide a sensitive genome-wide analysis of off-target genome editing events induced by the D10 strategy (*67*, *68*).

We generated purified D10 strategy ribonucleoprotein (RNP) complexes containing SpyMac nuclease and *P8* sgRNA to treat human genomic DNA from human embryonic kidney (HEK) 293T cells in vitro and analyzed rare off-target genomic cleavage events (fig. S3H). We identified 55 candidate SpyMac-dependent DNA off-target loci nominated by the CIRCLE-seq method. Next, we measured on-target and genomic off-target editing at the top 23 CIRCLE-seq–nominated loci in human cells (supplementary text, Fig. 2H, and fig. S3I). We achieved 49 ± 1.8% C6T on-target base editing at *SMN2* in HEK293T cells and observed minimal base editing at *SMN1* (0.15 ± 0.07%), which is generally absent in SMA patients. We detected minor levels of D10 base editing at the off-target site ranked 19 (0.41 ± 0.14%), which is in an intergenic region of chromosome 15, and no evident base editing (≤0.03% over untreated cells) at the other 21 assayed potential off-target loci. These data indicate high genomic target specificity of the D10 base editing strategy for the on-target locus.

Together, these experiments did not detect any coding mutations or sequence changes of anticipated physiological significance in the human genome, and they support continued preclinical evaluation of the D10 strategy, including assessment of base editor off-target editing measured in various tissues that may accumulate over an extended period of time. We refer to the D10 editing strategy as the "ABE strategy" hereafter.

Viral delivery of ABE enables efficient in vivo conversion of SMN2 C6T

To enable in vivo *SMN2* C6T conversion in an animal model of SMA, we designed an AAV strategy to package ABE8e-SpyMac and the *P8* sgRNA for delivery (v6 AAV-ABE8e; supplementary text, Fig. 3A, and fig. S3J). The AAV serotype 9 (AAV9) has a well-established tropism for neurons in the central nervous system (CNS) of a wide range of organisms, including Δ7SMA mice and human patients (*69*–*71*). In the cortex, AAV9 has been shown to almost exclusively target neurons (*71*), and intracerebroventricular (ICV) or systemic injection in neonates results in efficient transduction of spinal motor neurons to enable rescue of SMA disease phenotypes and lethality in both mice and humans (*13*, *32*, *72*). Thus, we selected AAV9 for delivery of our D10 ABE strategy ("AAV9-ABE") to Δ7SMA neonates by ICV injection to correct the *SMN2* C6T target in vivo (Fig. 3B).

We intracerebroventricularly injected SMA neonates with 2.7×10^{13} vector genomes per kilogram of body weight (vg/kg) of the dual AAV9-ABE vectors, along with 2.7×10^{12} vg/kg AAV9-Cbh-eGFP-KASH (Klarsicht/ANC-1/Syne-1 homology domain, hereafter AAV9-GFP) (*73*) to serve as a viral transduction control. This dose is comparable to doses used for PND0 ICV AAV administration of Zolgensma for rescue of Δ7SMA mice, and of other base editor AAVs that enable efficient genome editing in mice (*32*, *73*). We observed typical transduction patterns of AAV9 in the spinal cord (Fig. 3, C to E; supplementary text; and fig. S4A) (*32*, *33*, *74*). We quantified green fluorescent protein (GFP) and choline acetyltransferase (ChAT) double-positive cells in the ventral horn of spinal cords from injected mice and observed a mean transduction efficiency of 43% in spinal motor neurons (Fig. 3F), consistent with transduction efficiencies >20% previously shown to enable significant phenotypic rescue of Δ7SMA mice after ICV injection of self-complementary AAV9-SMN (Zolgensma) (*32*). Transduction of spinal motor neurons using 2.97×10^{13} vg/kg AAV9-GFP alone was similar (median: 46%) to transduction efficiencies using the 10-fold lower concentration of 2.7×10^{12} vg/kg, suggesting that the low-dose cotransduction of AAV9-GFP accurately represents the subset of cells transduced by AAV9-ABE.

Next, we assessed base editing in transduced cells (supplementary text and fig. S4B). We isolated cortical nuclei of treated animals and enriched for AAV9-transduction by sorting GFP-positive cells as previously described (*73*, *75*). We observed 87 ± 3.5% conversion of *SMN2* C6T among transduced cells (Fig. 3G), a 2.4-fold enrichment over unsorted tissue (37% ± 4.7%), with high single-nucleotide precision for C6T alone (73 ± 2.7%) and few indels (<0.4 ± 0.1%) or bystander edits, similar to D10 editing in Δ7SMA mESCs (Fig. 2E and figs. S2H and S4C). Collectively, these data confirm that ICV injection of AAV9-ABE in Δ7SMA neonates enables efficient and precise conversion of *SMN2* C6T in the CNS of treated animals with minimal undesirable by-products (*55*, *56*, *76*).

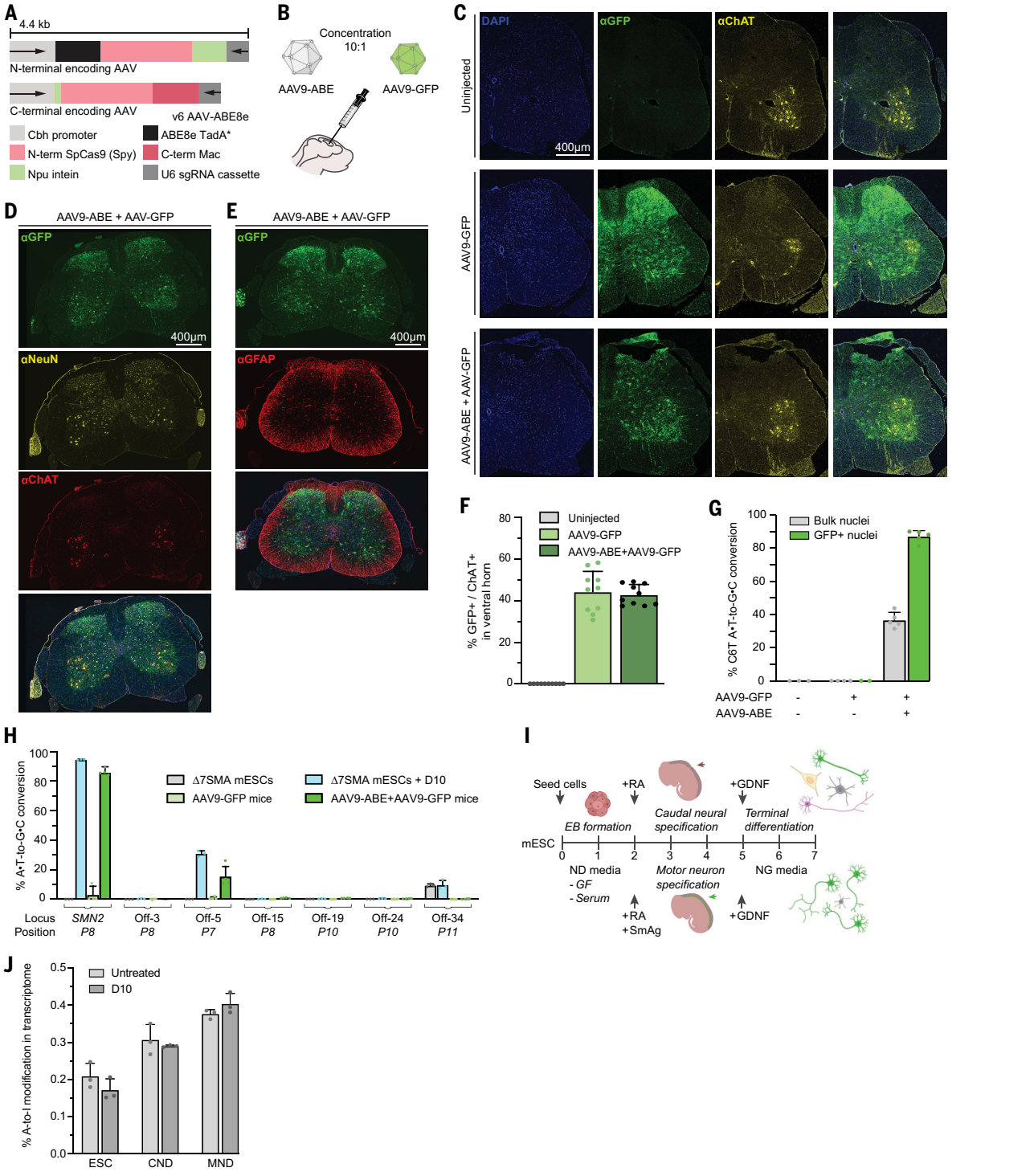


Fig. 3. Adenine base editing in Δ7SMA mice. (**A**) Dual-AAV vectors encoding split-intein ABE8e-SpyMac and *P8* sgRNA cassettes (v6 AAV9-ABE8e). (**B**) Neonatal ICV injections in Δ7SMA mice with AAV9-ABE, and AAV9-GFP as a transduction control. (**C** to **E**) Immunofluorescence images of lumbar spinal cord sections from wild-type Δ7SMA mice at 25 weeks old, intracerebroventricularly injected on PND0 and PND1 with AAV9-ABE, AAV9-GFP, or uninjected, as indicated. GFP indicates transduction, ChAT labels spinal motor neurons in the ventral horn, NeuN labels postmitotic neurons, GFAP labels astrocytes, and 4′,6-diamidino-2-phenylindole (DAPI) stains all nuclei. (**F**) Quantification of GFP and ChAT double-positive cells within the ventral horn (n = 3 animals). (**G**) Base editing in bulk and GFP+ flow-sorted nuclei of Δ7SMA mice treated with AAV9-ABE+AAV9-GFP (n = 5), AAV9-GFP (n = 4), or uninjected (n = 3). (**H**) On-target and off-target editing after VIVO analysis of strategy D10 in Δ7SMA mESCs compared with AAV9-ABE+AAV9-GFP treatment in Δ7SMA mice. Bars show editing of the most frequently edited nucleotide at each locus, with the *P#* position shown along the bottom of the graph. (**I**) Schematic of motor neuron differentiation (MND) and caudal neural differentiation (CND) of Δ7SMA mESCs. EB, embryoid body; RA, retinoic acid; GDNF, glial cell line–derived neurotrophic factor; ND, neural differentiation; GF, growth factor; SmAG, smoothened agonist; NG, neural growth. (**J**) Whole-transcriptome A-to-I RNA off-target editing analysis in Δ7SMAmESCs (n = 3) and CND (n = 3) and MND (n = 3) differentiated cells stably expressing the D10 strategy. Error bars indicate standard deviations.

Base editing conversion of C6T effectively converts the native *SMN2* gene to *SMN1*, thereby restoring SMN protein levels to that of wild-type cells. Current SMA drugs induce non-native SMN levels (*23*, *24*, *32*–*35*) and require repeated dosing or may fade over time. The permanent and precise editing of endogenous *SMN2* genes that preserves native transcript levels and native regulatory mechanisms governing *SMN* expression thus may address shortcomings of existing SMA therapies (*1*, *4*, *5*, *21*, *28*, *77*).

In vitro and in vivo DNA and RNA off-target analysis of ABE8e targeting SMN2 *C6T*

In addition to the off-target analysis in human cells described above, we also assessed the DNA and RNA specificity of the ABE strategy in mouse cells in vitro and in vivo. We performed CIRCLE-seq and validated the top 35 nominated sites in Δ7SMA mESCs (supplementary text and fig. S4D). We achieved 95 ± 0.0% on-target editing at the *SMN2* transgene and only observed substantial editing at off-target site 5 in an intron of the mucin 16 gene (Muc16, 31 ± 1.9%) that is not expressed in the CNS (fig. S4E) (*78*). Next, we compared this analysis to off-target editing in vivo after AAV9-ABE ICV injection in Δ7SMA neonates by performing verification of in vivo off targets (VIVO) (*79*). We observed 10–27% (average 15 ± 7%) editing at off-target site 5 in intron 54 of Muc16 and 0.1–0.9% (average 0.5 ± 0.3%) editing at the noncoding off-target site rank 15, compared with 87 ± 3.5% average on-target editing of *SMN2* among GFP-positive cells in the CNS across five animals (Fig. 3, G and H). These animals ranged from 4 to 18 weeks of age at the time of off-target analysis (26, 36, 42, 80, and 127 days old), and we observed no increase in off-target editing events over time. Thus, off-target editing outcomes observed in cell culture experiments were consistent with those observed in vivo over 18 weeks (*79*). The ABE strategy did not induce any detected coding mutations in either human or mouse genomes, and off-target editing in vivo was lower than in cell culture (reduced by 50% at Muc16 intron 54), likely because of lower copy number and expression levels in transduced cells in vivo or in vivo gene silencing over time (*33*, *36*, *37*).

Cas-independent RNA off-target adenine base editing in vivo is typically indistinguishable from background A-to-I conversion owing to the low copy number of ABE-expressing transgenes (*33*, *80*). We investigated RNA off-target editing in Δ7SMA mESCs and differentiated neural lineages, including motor neurons, that stably produce ABE8e from low gene copy numbers similar to those resulting from AAV9 transduction (Fig. 3I; supplementary text; and fig. S4, F to H). Consistent with previous reports (*80*, *81*), whole-transcriptome sequencing did not reveal detected accumulation of RNA A-to-I edits over background levels of endogenous A-to-I and A-to-G changes (Fig. 3J and fig. S4G).

Collectively, these in vitro and in vivo analyses did not reveal off-target edits of anticipated clinical or physiological significance in human or mouse cells, suggesting high target specificity of the D10 base editing approach. Continued preclinical assessment and minimization of off-target editing is important to ensure the safety of a potential base editing therapeutic for the treatment of SMA in patients.

ABE-mediated rescue of SMA pathophysiology in mice

The physiology of AAV9-ABE–treated Δ7SMA mice was improved compared with that of untreated animals (movies S1 and S2). We assessed the rescue of motor phenotypes by electrophysiological measurements in AAV9-ABE–treated Δ7SMA mice. We measured compound muscle action potential (CMAP) amplitude and performed motor unit number estimation (MUNE) in the gastrocnemius muscle to assess loss of motor neuron functional integrity, a key feature of SMA and preclinical SMA models (*82*). We compared outcomes with US Food and Drug Administration (FDA)–approved therapeutics for SMA including ICV injection of Zolgensma and daily intraperitoneal (IP) injection of risdiplam (Evrysdi) at doses that were previously demonstrated to confer a survival benefit to these mice (3.3×10^{13} vg/kg Zolgensma and 0.1 mg/kg risdiplam; Fig. 4A) (*30*, *32*). MUNE was reduced by 50% in untreated Δ7SMA animals compared with heterozygous mice at postnatal day (PND) 12, and Zolgensma or 0.1 mg/kg risdiplam showed little to no improvement (50 and 75% relative to heterozygotes, respectively; Kruskal-Wallis test, $P > 0.6$). In contrast, MUNE in SMA mice treated with 3.3×10^{13} vg/kg of AAV9-ABE was significantly improved compared with untreated animals (Kruskal-Wallis test, $P < 0.02$) and did not significantly differ from heterozygous animals, with values averaging 91% that of heterozygotes. CMAP amplitudes were also higher for AAV9-ABE–treated mice compared with risdiplam-treated or untreated Δ7SMA mice, whereas CMAP amplitudes did not significantly differ between heterozygotes, Zolgensma-treated mice, and AAV9-ABE–treated animals (Kruskal-Wallis one-way analysis of variance, $P > 0.2$). Thus, neonatal ICV injection of AAV9-ABE measurably rescues SMA pathophysiology of spinal motor neurons.

Next, we assessed survival of intracerebroventricularly AAV9-ABE–injected Δ7SMA mice. In SMA type I patients, therapeutic intervention can meaningfully improve disease outcomes if administered in the first several months of life (*83*–*86*); however, in Δ7SMA mice, survival drops precipitously when animals receive treatment past PND6 (Fig. 4B) (*87*). This large difference is due in part to the highly accelerated (~150-fold greater) rate of maturation of mice compared with humans in the first month, early perinatal reduction in SMN expression that occurs in mice (*88*) and humans (*28*), and the rapid early-onset loss of motor units, which consist of spinal motor neurons and the muscle fibers that they innervate (*82*, *89*). Restoration of SMN protein levels using inducible transgenes demonstrates that high levels of SMN are required by PND4 to PND6 to rescue Δ7SMA mice, and delays of a small number of days are strongly anti-correlated with survival (*32*, *87*, *88*, *90*–*92*). In cells, complete mRNA rescue is not achieved until 7 days after D10 transfection (fig. S5A), and the time to restore SMN protein levels in vivo surpasses the extremely short therapeutic window in Δ7SMA mice.

The accumulation of SMN protein after transduction with the dual single-stranded AAV9-ABE8e vectors used in this study requires completion of (i) second-strand synthesis of each AAV9-ABE genome (*93*, *94*), (ii) transcription and translation of the split-intein ABE protein segments, (iii) assembly and trans-splicing of the split ABE protein, (iv) RNP assembly and base editing of *SMN2*, (v) transcription of full-length C6T-modified endogenous *SMN2* pre-mRNA driven by its native promoter, and (vi) splicing and translation of corrected *SMN2* transcripts. Thus, the timing for SMN protein rescue after AAV9-ABE administration is slower than fast-acting splice-switching drugs or constitutive gene complementation from SMN cDNA encoded by a self-complementary AAV9-SMN vector such as Zolgensma (*93*–*95*). We recently demonstrated that in vivo base editing affects protein levels by ~1 to 3 weeks after administration (*80*).

Despite the incongruent timeline of base editing–mediated rescue for ideal rescue of Δ7SMA mice, AAV9-ABE increased the life span of treated animals by ~33% in two colonies in different institutions (supplementary text; Materials and methods; Fig. 4C; and fig. S5, B to D). Life span of treated animals improved from an average of 17 days (median: 17 days; maximum: 20 days) to 23 days (median: 22 days; maximum: 33 days; Mantel-Cox test, $P < 0.02$). As anticipated, the life-span extension resulting from AAV9-ABE treatment is similar to that achieved by scAAV9-SMN gene therapy in postsymptomatic (>PND7) Δ7SMA mice (Fig. 4B) (*32*, *72*, *87*, *92*). Collectively, these data demonstrate that postnatal conversion of *SMN2* C6T by AAV9-ABE rescues SMA motor phenotypes in mice, including the number (MUNE) and output (CMAP) of functional motor units innervating muscle, and that the prolonged process of AAV9-ABE–mediated SMN restoration results in mostly

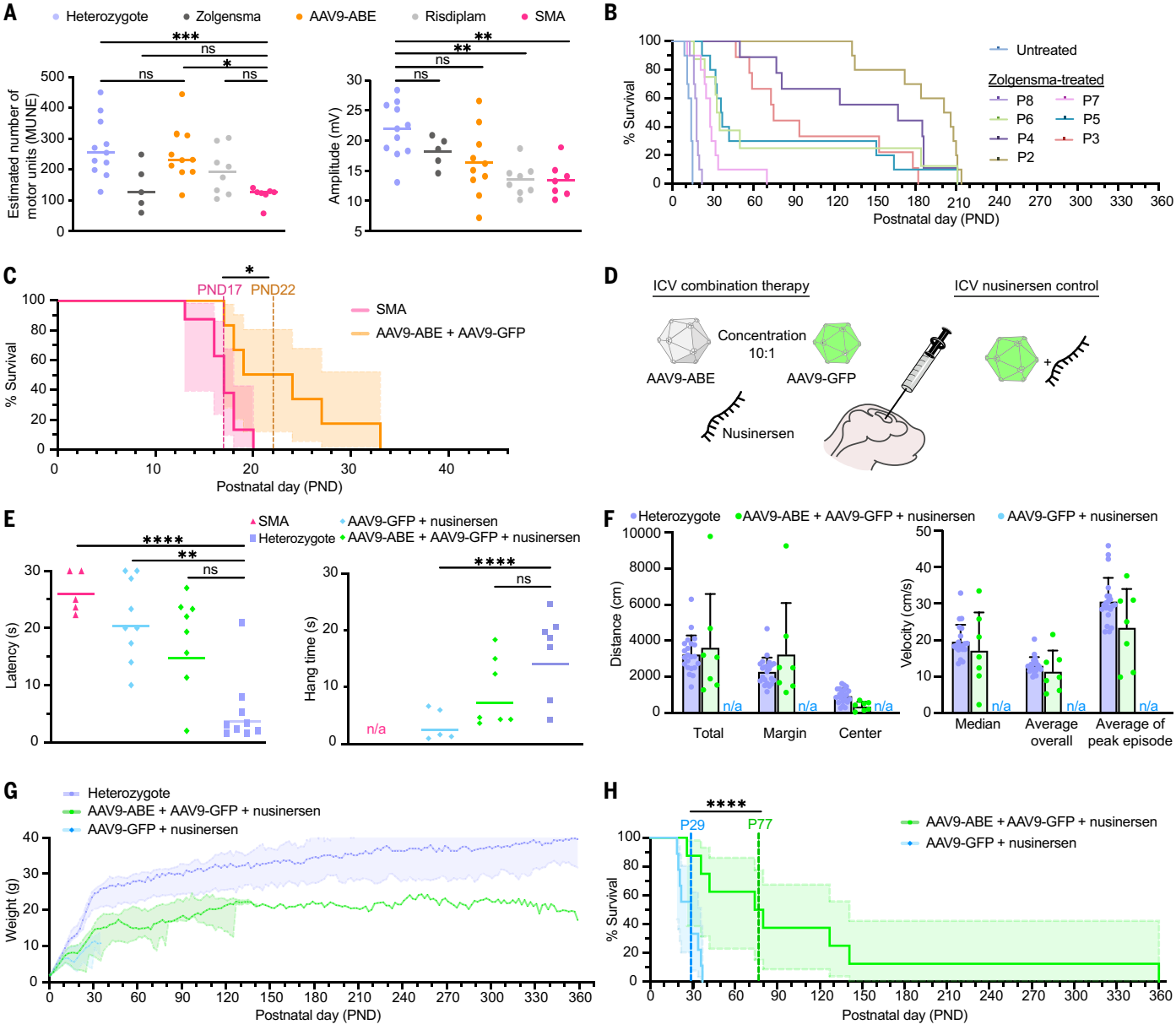


Fig. 4. AAV9-ABE–mediated rescue of Δ7SMA mice. (**A**) (Left) Motor unit number estimation (MUNE) and (right) compound muscle action potential (CMAP) amplitude at PND12 in heterozygotes (n = 11), and Δ7SMA mice treated with Zolgensma (n = 5), AAV9-ABE (n = 10), risdiplam (n = 8), or uninjected (n = 7). (**B**) Kaplan-Meier curve of Δ7SMA neonates intracerebroventricularly injected with Zolgensma from Robbins *et al.* (*87*) (data extracted using PlotDigitizer). Average (av), median (md), and longest (lng) survival in days: untreated (avg: 13; med: 14; lng: 15), PND2 (avg: 187; med: 204; lng: 214), PND3 (avg: 102; med: 75; lng: 182), PND4 (avg: 141; med: 167; lng: 211), PND5 (avg: 76; med: 37; lng: 211), PND6 (avg: 73; med: 34; lng: 211), PND7 (avg: 30; med: 28; lng: 70), and PND8 (avg: 18; med: 18; lng: 22). (**C**) Kaplan-Meier curve in AAV9-ABE treated (n = 6) and uninjected (n = 8) Δ7SMA mice. (**D**) Neonatal ICV co-injections with AAV9-ABE, AAV9-GFP, and nusinersen. (**E**) (Left) The time required for Δ7SMA mice to right themselves in the righting reflex assay at PND7. (Right) The hang time of Δ7SMA mice in the inverted screen test at PND25. (**F**) Analysis of voluntary movement by open field tracking at PND40. (Left) Traveled distance in centimeters. (Right) Velocity in centimeters per second. (**G** and **H**) Body weight in grams and Kaplan-Meier curve of Δ7SMA mice. Graph line shading represents (G) standard deviation or (H) 95% CI. Animals were treated as indicated. Dots represent individual animals. *$P \leq 0.02$, **$P \leq 0.01$, ***$P \leq 0.005$, ****$P \leq 0.001$. Error bars indicate standard deviations.

postsymptomatic rescue in Δ7SMA mice that results in a statistically significant, but limited, improvement in animal life span.

Up-regulation of SMN protein levels improves motor function and life expectancy of SMA patients and animal models if achieved before the onset of neuromuscular pathology and symptoms (*13*, *32*, *85*–*87*, *92*), yet even high levels of SMN protein cannot correct neuromuscular junction defects once SMA has progressed to an advanced stage, and loss of motor neurons upon cell death is irreversible. We therefore sought to extend the effective therapeutic window for gene editing by transient early administration of an existing approved SMA drug to attenuate disease progression, as has previously been applied to study milder forms of SMA in mice (*72*, *96*, *97*). Given that SMA patients in a gene editing clinical trial would likely be receiving an SMA drug, repeating the base editing treatment in mice receiving an existing SMA drug would also inform a potential future clinical application of this approach.

Combination therapy improves the life span of ABE-treated SMA mice

Transient SMA drug administration can ameliorate SMA pathology and extend survival of Δ7SMA mice. We hypothesized that attenuating disease progression using nusinersen could extend the unusually short therapeutic window of Δ7SMA mice and allow AAV9-ABE–mediated rescue to begin before extensive irreversible SMA damage occurs. The mechanism of nusinersen (binding to *SMN2* pre-mRNA) is orthogonal to base editing of *SMN2* genes, and cotransfection of 20 nM nusinersen did not affect base editing outcomes or inclusion of exon 7 in spliced *SMN* transcripts after D10 in Δ7SMA mESCs (fig. S5, A and E). We assessed whether coadministration of nusinersen can improve phenotypic rescue from AAV9-ABE treatment. A single ICV injection of nusinersen at PND0 has been shown to extend survival of Δ7SMA mice by several weeks (*98*), thus we co-injected a single low dose (1 μg) of nusinersen together with AAV9-ABE and AAV9-GFP in Δ7SMA neonates (supplementary text). As a control, we also treated Δ7SMA neonates with 1 μg nusinersen and AAV9-GFP but no base editor (Fig. 4D). We assessed motor coordination and overall muscle strength at PND7 using the righting reflex test, which measures the time needed for a mouse placed on its back to right itself (Fig. 4E). We observed a significant difference between heterozygotes and nusinersen-treated or untreated Δ7SMA mice (Kruskal-Wallis test, $P \leq 0.01$) but no significant difference between mice treated with a combination of AAV9-ABE and nusinersen (hereafter AAV9-ABE+nusinersen) compared with heterozygous littermates (Kruskal-Wallis test, $P > 0.1$).

Next, we assessed motor strength and coordination of treated and heterozygous mice using an inverted screen test, which measures how long a mouse can hang inverted from a screen mesh surface. At PND25, Δ7SMA animals treated with nusinersen alone performed significantly worse at inverted screen testing than did healthy heterozygous mice (Kruskal-Wallis test, $P = 0.007$; Fig. 4E). In contrast, the AAV9-ABE+nusinersen–treated animals showed no significant difference in the inverted screen assay from healthy heterozygous mice. Notably, half of nusinersen-only–treated animals were deceased by this time point, and age-matched untreated Δ7SMA mice do not survive long enough for this PND25 assay.

For a more complete behavioral assessment of treated and heterozygous animals, we performed extensive multiparametric analysis of voluntary movement by open field tracking at PND40 (Fig. 4F and fig. S5, F to J). Across 33 parameters, including traveled distances, velocity, duration, and counts of various activities, the measured behaviors of AAV9-ABE+nusinersen–treated animals showed no significant difference with those of heterozygous mice (Mann-Whitney test, $P > 0.5$). Neither nusinersen-only treated or untreated age-matched Δ7SMA mice were available as a reference for this PND40 assay owing to their short life span.

We also assessed the effect of AAV9-ABE+nusinersen treatment on weight and life span of Δ7SMA mice. The weight of nusinersen-only– and AAV9-ABE+nusinersen–treated Δ7SMA mice steadily increased and was indistinguishable for the first week of life, after which weight gain slowed in the nusinersen-only cohort (Fig. 4G). Combination-treated animals maintained, on average, 61 ± 4.0% of the weight of heterozygous animals throughout their life spans. The nusinersen-only injection improved life span of Δ7SMA mice from an average of 17 days (median: 17; maximum: 20 days; Fig. 4C) to an average of 28 days (median: 29; maximum: 37 days; Mantel-Cox test, $P = 0.0001$; Fig. 4H). Notably, AAV9-ABE+nusinersen treatment improved survival of Δ7SMA mice to an average of 111 days (median: 77; Mantel-Cox test, $P = 0.002$), with >60% of animals surviving beyond nusinersen-only controls, and a 10-fold increase in maximum life span (37 days maximum with nusinersen only compared with 360 days maximum with AAV9-ABE). AAV9-ABE+nusinersen–treated SMA mice also exhibited normal behavior and vitality well beyond the life span of nusinersen-only controls (P40, P96, and P200 in movies S3 to S5). Collectively, these data indicate that transient extension of the very narrow therapeutic window in Δ7SMA mice can greatly improve phenotypic rescue of SMA from base editing of *SMN2*.

Whereas neonatal AAV9-ABE ICV injection alone enables life extension in Δ7SMA mice that resembles >PND7 ICV injection with Zolgensma (Fig. 4, B and C) (*87*), coadministration of 1 μg nusinersen temporarily slows disease progression and broadens the narrow therapeutic window, allowing base editing the opportunity to enable life-span rescue that more closely resembles that of presymptomatic Zolgensma administration at ≤PND3 (Fig. 4H). Moreover, these data demonstrate compatibility of AAV9-ABE with nusinersen as a one-time treatment without evident adverse effects and with apparent synergy to improve therapeutic outcomes. Such a combination therapy approach may play an important role in future clinical trial designs for one-time SMA treatments that permanently correct a genetic cause of the disease and for clinical application in patients already receiving treatment.

Discussion

Current treatment options for SMA have revolutionized care for thousands of patients, effectively extending life span, preventing the loss of motor function in presymptomatic patients, and delaying progression in symptomatic patients by increasing full-length SMN protein levels (*13*, *24*, *85*, *86*, *90*, *99*). However, current therapies do not restore endogenous protein levels and native regulation of SMN, which could result in pathogenic SMN insufficiency in motor neurons or potential long-term toxicity in other tissues (*21*, *23*–*28*, *35*). Furthermore, the transient therapies nusinersen and risdiplam require repeated dosing throughout a patient's lifetime, and it is unclear how long Zolgensma gene complementation will persist in motor neurons (*36*, *37*). Thus, achieving permanent and endogenously regulated rescue of SMN protein levels is an important goal of a future therapeutic for SMA patients. The optimized D10 ABE strategy developed in this work is a one-time treatment that enables permanent and precise editing of endogenous *SMN2* genes while preserving native transcript levels and regulatory mechanisms that govern *SMN* expression (*1*, *4*, *5*, *28*, *77*, *100*). As such, a future base editing therapeutic approach could offer substantial benefits over existing SMA therapies.

We compared a total of 79 nuclease and base editing strategies targeting five regions of *SMN2* to induce either posttranscriptional or posttranslational regulatory changes in *SMN2* that up-regulate SMN protein production. BE-Hive and inDelphi machine learning models enabled the design of precise editing strategies that, in some cases, were not obvious and also helped preselect sgRNAs for genotypic and phenotypic validation of editing outcomes. All SMA patients, regardless of their *SMN1* mutations, must carry the *SMN2* gene to complete gestation (*7*), and thus the genome editing strategies identified in this study have the potential to benefit all SMA patients.

While on-target Cas nuclease editing at *SMN2* can be precise, double-strand breaks (DSBs) can result in large deletions and chromosomal rearrangements, especially when induced simultaneously at multiple genomic loci (*101*). Given that SMA patients usually have multiple copies of *SMN2*, nuclease editing may result in unintended restructuring of the chromosome region (5q13) that harbors *SMN* genes (*102*, *103*). In contrast, base editors precisely convert nucleotides without inducing DSBs (*50*, *104*, *105*) and result in greater SMN protein up-regulation than the nuclease strategies in this study (up to 50-fold by base editors compared with up to 17-fold by nucleases). We therefore recommend that future gene editing therapeutic strategies for SMA use base editing rather than nucleases.

ABE strategy D10 demonstrated high on-target efficiency and specificity, with minimal Cas-dependent or Cas-independent off-target DNA or RNA editing. It is possible that extended base editor expression in cells, as can

result from AAV-delivery, could result in a greater accumulation of genomic and transcriptomic off-target events. Therefore, a deeper assessment of genomic and transcriptomic off-targets and efforts to minimize off-target editing risk will be important in the preclinical development of a potential base editing therapeutic for SMA. If needed, Cas-independent editing events can be further minimized by alternative delivery strategies that shorten exposure to base editors (*61*) and by the use of tailored deaminases such as the V106W variant of TadA*-8e (*61*, *63*) or TadA-8.17-m (*106*).

SMA has variable presentation in humans that largely correlates with the copy number of *SMN2* (*107*–*111*). Type I SMA patients have two *SMN2* copies and present with symptoms within the first 6 months of life, type II patients have three copies and present with symptoms by 18 months, whereas type III patients have 3 or 4 *SMN2* copies and later onset of symptoms. Early intervention is paramount to achieving the best outcomes for SMA patients. The window to effectively treat type II and III patients is broader than for type I patients, who should ideally receive treatment within the first few months of life and up to 18 months (*13*, *24*, *83*–*86*, *99*). Indeed, we directly observed the critical role of differences in timing on the order of days in determining the efficacy of an AAV9-ABE treatment in Δ7SMA mice, which have an unusually short (≤6 days) therapeutic window compared with the time scale of base editing (weeks) (*87*). We show that the FDA-approved ASO drug nusinersen can extend the very short therapeutic window for rescue in Δ7SMA mice, allowing base editing–mediated rescue of SMN protein levels to occur to a greater extent (*80*). We anticipate that the broader therapeutic window in human SMA patients would provide ample opportunity for AAV9-ABE–mediated restoration of SMN protein levels to take place without the need for co-administration of a transient therapeutic. Nevertheless, our study demonstrates the compatibility of base editing with nusinersen as a combination therapy approach to treat SMA in animals, which may be valuable for future clinical applications.

The intracerebroventricularly injected AAV9-ABE animals in our study exhibited mouse-specific peripheral disease phenotypes that are common in SMA mouse models, including necrosis of the extremities (*112*), while exhibiting otherwise normal behavior and vitality without displays of progressive muscle weakness. However, SMA treatment that is restricted to the CNS also reveals a later-onset (≥2 months) lethal cardiac abnormality specific to Δ7SMA mice (*32*, *113*–*116*), which likely underlies the sudden late-stage fatality observed in intracerebroventricularly AAV9-ABE–treated animals in this study. Treating both CNS and peripheral tissues may ameliorate this murine cardiac phenotype to improve life span of treated Δ7SMA mice compared with intracerebroventricularly injected animals (*113*, *117*). Nevertheless, given that patients have been successfully treated intrathecally with Spinraza, peripheral restoration of SMN protein does not appear to be required to rescue SMA lethality in humans (*23*, *31*, *83*, *90*, *118*).

As demonstrated in this work, dual-AAV delivery of base editors supports therapeutic levels of editing in mouse models of human disease (*119*, *120*). After these in vivo experiments were completed, our lab developed efficient in vivo base editing using single–AAV9-ABE systems that use size-minimized AAV vector components and one of a suite of small Cas protein domains that are highly active as ABEs (*80*). Such single-AAV base editing systems may simplify the development of future base editor therapeutics and potentially minimize the required dose and potential side effects of AAV in clinical settings (*121*).

Materials and methods

Cell culture

Culture of mESCs, HEK293T, and U2OS cells was performed according to previously published protocols (*122*). mESCs were maintained on 0.2% gelatin-coated plates feeder-free in mESC media composed of Knockout Dulbecco's Modified Eagle medium (DMEM; Life Technologies) supplemented with 15% defined fetal bovine serum (FBS, HyClone), 0.1 mM nonessential amino acids (NEAA, Life Technologies), Glutamax (GM, Life Technologies), 0.55 mM 2-mercaptoethanol (b-ME, Sigma-Aldrich), 1X ESGRO LIF (Millipore), with the addition of 2i: 5 nM GSK-3 inhibitor XV (Sigma-Aldrich), and 500 nM UO126 (Sigma-Aldrich). Δ7SMA mESCs were a kind gift from L. L. Rubin. HEK293T cells were purchased from ATCC (CRL-3216) and were maintained in DMEM (Life Technologies) supplemented with 10% FBS (Thermo Fisher Scientific). U2OS cells were purchased from ATCC (HTB-96) and were maintained in McCoy's 5a medium (Life Technologies) supplemented with 10% FBS (Thermo Fisher Scientific). All cells were regularly tested for mycoplasma.

For genome editing experiments, cells were seeded 1 day prior to be ~70 to 80% confluent on the day of transfection and transfected with sgRNA and genome editing plasmids at a 1:1 molar ratio using Lipofectamine 3000 (Thermo Fisher Scientific) in accordance with the manufacturer's protocols. For stable integration of plasmids, cells were cotransfected with Tol2 transposase at an equimolar ratio. Cells that did not undergo antibiotic selection were cultured for 3 to 5 days before harvesting. For antibiotic selection, Δ7SMA mESCs were treated with 50 μg/ml hygromycin B (Life Technologies) and/or 6.67 μg/ml blasticidin as indicated, starting 24 hours after transfection. For transient selection, antibiotics were removed from the media after 48 hours. Selected cells were allowed to recover and expand before harvesting. All sgRNA sequences designed for this study are listed in the supplementary materials.

For Δ7SMA mESC nusinersen experiments, cells were transfected with 20 nM of fully 2′-*O*-methoxyethyl (MOE)–modified ASO (5′-TCACTTTCATAATGCTGG-3′) on a phosphorothioate backbone (TriLink), using Lipofectamine 3000 (Thermo Fisher Scientific). After 24 hours, media was replaced every other day with fresh mESC+2i media. For splicing rescue by risdiplam, mESC media was supplemented with 0.1 to 1 μM of risdiplam (RG7916, Selleck Chemicals LLC) in dimethyl sulfoxide, as indicated. Cells were harvested at the indicated time points.

High-throughput sequencing of genomic DNA

Sequencing library preparation was performed according to previously published protocols (*50*). Primers are listed in the supplementary materials. Briefly, we isolated genomic DNA (gDNA) with the QIAamp DNA mini kit (Qiagen) and used 250 to 1000 ng of gDNA for individual locus editing experiments and 20 μg of gDNA for comprehensive context library samples. Sequencing libraries were amplified in two steps, first to amplify the locus of interest and second to add full-length Illumina sequencing adapters using the NEBNext Index Primer Sets 1 and 2 [New England Biolabs (NEB)] or internally ordered primers with equivalent sequences. All PCRs were performed using NEBNext Ultra II Q5 Master Mix. Samples were pooled using Tape Station (Agilent) and quantified using a KAPA Library Quantification Kit (KAPA Biosystems). The pooled samples were sequenced using Illumina NextSeq or MiSeq. Alignment of fastq files and quantification of editing frequency for individual loci were performed using CRISPResso2 in batch mode (*66*). The editing frequency for each site was calculated as the ratio between the number of modified reads (i.e., containing nucleotide conversions or indels) and the total number of reads. Base editing characterization library analysis was performed as previously described (*50*).

Quantification of SMN splice products

We isolated mRNA from Δ7 mESCs with the RNeasy mini kit (Qiagen) and performed reverse transcription using SuperScript IV (Thermo Fisher) according to the manufacturer's protocols. For targeted *SMN2* splice product quantitation by qPCR, high-throughput sequencing (HTS), or automated electrophoresis, we performed reverse transcription with random hexamers. Inclusion of *SMN2* exon 7 was quantified by automated electrophoresis using Tape Station (Agilent). For unbiased

SMN2 splice product analysis by high-throughput sequencing, we performed reverse transcription using a custom oligo-dT primer with a Read 2 Illumina sequencing stub. The pooled samples were sequenced using Illumina MiSeq. All PCRs were performed using NEBNext Ultra II Q5 Master Mix, with the addition of Sybr Green for qPCR. Primers are listed in table S3.

Western blot

Cells harvested for Western blot were washed with ice-cold phosphate-buffered saline (PBS) and incubated at 4°C for 30 min while rocking in RIPA lysis buffer (Thermo Fisher) supplemented with 1 mM phenylmethylsulfonyl fluoride (Thermo Fisher) and cOmplete EDTA-free protease inhibitor cocktail (Roche). Lysates were clarified by centrifugation at 12,000 rpm at 4°C for 20 min. Lysates were normalized using bicinchoninic acid (BCA; Pierce BCA Protein Assay Kit) and combined with 4× Laemelli buffer (BioRad) and dithiothreitol (Thermo Fisher) at a final concentration of 1 mM. We loaded 10 μg of reduced protein per gel lane and performed transfer with an iBlot 2 dry blotting system (Thermo Fisher) using the following program: 20 V for 1 min, then 23 V for 4 min, then 25 V for 2 min, for a total transfer time of 7 min. Blocking was performed at room temperature for 60 min with block buffer: 1% bovine serum albumin (BSA) in TBST (150 mM NaCl, 0.5% Tween-20, 50 mM Tris-Cl, pH 7.5). Membranes were then incubated in primary antibody diluted in block buffer for 2 hours at room temperature. After a washing, secondary antibodies diluted in TBST were added and incubated for 1 hour at room temperature. Membranes were washed again and imaged using a LI-COR Odyssey. Wash steps were 3× 5-min washes in TBST. Primary antibodies used were mouse anti-human SMN (Proteintech 2C6D9), mouse anti-mouse and human SMN (Proteintech 3A8G1), and rabbit anti-histone H3 (Cell Signaling D1H2); secondary antibodies used were LI-COR IRDye 680RD goat anti-rabbit (#926–68071) and goat anti-mouse (#926–68070).

Base editor characterization library assay

For characterization of the ABE8e-SpCas9 base editor, we used mouse ESCs carrying the comprehensive context library according to previously published protocols (*42*, *50*). Briefly, 15-cm plates with >10^7 initial cells were transfected with a total of 50 μg of p2T-ABE8e-SpCas9 and 30 μg of Tol2 plasmid to allow for stable genomic integration with Lipofectamine 3000 according to manufacturer protocols and selected with 10 μg/ml blasticidin starting the day after transfection for 4 days before harvesting. We maintained an average coverage of ~300× per library cassette throughout. We collected gDNA from cells 5 days after transfection, after 4 days of antibiotic selection.

Cloning

Base editor plasmids were constructed by replacing deaminase and Cas-protein domains of the p2T-CMV-ABE7.10-BlastR (Addgene 152989) plasmid by USER cloning (NEB) (*50*). Individual sgRNAs were cloned into the SpCas9-hairpin U6 sgRNA expression plasmid (Addgene 71485) using BbsI plasmid digest and Gibson assembly (NEB). Protospacer sequences and gene-specific primers used for amplification followed by HTS are listed in table S1. Constructs were transformed into Mach1 chemically competent *Escherichia coli* (Thermo Fisher) grown on LB agar plates, and liquid cultures were grown in LB broth overnight at 37°C with 100 μg/ml ampicillin. Individual colonies were validated by TempliPhi rolling circle amplification (Thermo Fisher) followed by Sanger sequencing. Verified plasmids were prepared by mini, midi, or maxiprep (Qiagen).

AAV vectors were cloned by Gibson assembly (NEB) using NEB Stable Competent *E. coli* (High Efficiency) to insert the sgRNA sequence and C-terminal base editor half of ABE8e-SpyMac into v5 Cbh-AAV-ABE-NpuC+U6-sgRNA (Addgene 137177), and the N-terminal base editor half and a second U6-sgRNA cassette into v5 Cbh-AAV-ABE-NpuN (Addgene 137178) (*73*).

Neural differentiation

Differentiation of Δ7SMA mESCs was performed according to established protocols (*123*, *124*). Briefly, Δ7SMA mESCs maintained on 0.2% gelatin-coated plates feeder-free in mESC media + 2i were plated onto irradiated mouse embryonic fibroblast (iMEF) feeders on 0.2% gelatin-coated plates in mESC media for 7 days to wean cells from 2i factors. Cells were then seeded at 10^6 in 10-cm tissue culture–treated plates for 48 hours for priming and depletion of feeders. Media was replaced with neural differentiation (ND) media composed of 1:1 DMEM:F12 and Neurobasal media (Life Technologies) supplemented with 10% knockout serum-replacement (KOSR, Life Technologies), Glutamax (GM, Life Technologies), and 0.55 mM 2-mercaptoethanol (b-ME, Sigma-Aldrich) for one hour before trypsinization and seeding of 2 × 10^6 cells in 10-cm non–tissue culture–treated dishes for 24 hours. Single cells and small early embryoid bodies (EBs) in suspension were collected and transferred to 10-cm tissue culture–treated plates in fresh ND media for 24 hours. Small EBs that remained in suspension were collected and transferred to 10-cm tissue culture–treated plates in fresh ND media with the addition of 1 μM retinoic acid (RA; Sigma-Aldrich R2625) for caudal neural differentiation (CND), or with 1 μM RA and 0.5 μM smoothened agonist (SmAg; Calbiochem 566660) for motor neuron differentiation (MND) for 72 hours. Large EBs were collected and split into two 10-cm tissue culture–treated plates in neural growth (NG) media composed of 1:1 DMEM:F12 and Neurobasal media supplemented with GM, B27 (Life Technologies), and 10 ng/ml human recombinant glial cell line–derived neurotrophic factor (GDNF; R&D Systems 212-GD-010) for 48 hours. EBs were monitored for *Mnx1:GFP* expression to assess motor neuron differentiation efficiency and imaged using a Zeiss inverted fluorescence microscope or collected for downstream whole-transcriptome analysis.

Whole-transcriptome RNA sequencing

Library preparation, sequencing and analysis were performed by SMART-seq2 as previously described (*125*). Briefly, total RNA was harvested from cells using the RNeasy Mini kit (Qiagen). First, we incubated 20 ng purified total RNA with RNase inhibitor (Clontech Takara 2313B), deoxynucleoside triphosphate (dNTP) mix (Thermo Fisher R0192), and the 3′-RT primer [5′-AAGCAGTGGTATCAACGCAGAGTAC(T30)VN-3′] at 72°C for 3 min to anneal the RT primer. Next, we performed first-strand synthesis using the template switching oligo (TSO): (5′-AGCAGTGGTATCAACGCAGAGTACrGrG+G-3′ Exiqon, Qiagen) together with RNase inhibitor, betaine (Sigma Aldrich B0300-1VL), $MgCl_2$ (Sigma Aldrich 1028), and Maxima RNase H-minus RT (Thermo Fisher EP0751), according to the manufacturer's protocols. We performed preamplification of first-strand libraries with the ISPCR primer: 5′-AAGCAGTGGTATCAACGCAGAGT-3′ using KAPA HiFi HotStart (KAPA KK2601) and SYBR green (Thermo Fisher). Whole-transcriptome amplification (WTA) product was washed using DNA SPRI beads (Beckman Coulter A63881) and quantified by Agilent Tapestation. We performed Tagmentation and library preparation of 0.25 ng WTA using the Nextera XT kit (Illumina) and Nextera i7 and Nextera i5 barcoding primers. Samples were pooled and washed using DNA SPRI beads and quantified by Agilent Tapestation and the KAPA Universal Library Quantification kit (Roche KK4824). Libraries were run on Illumina NextSeq 550.

FASTQs were generated using bcl2fastq v2.20 and processed with Trim Galore v0.6.7 in paired-end mode with default parameters to remove low-quality bases, adapter sequences, and unpaired sequences. Trimmed reads were aligned to the GENCODE mouse reference genome M31 (GRCm39) using STAR (v2.7.10a), quantified using kallisto (*126*), and refined to canonical coding sequences using CCDS release 21 (*127*). For RNA A-to-I off-target analysis, REDItools v1.3 was used to quantify the average frequency of A-to-I editing among all sequenced adenosines in each sample (*128*), excluding adenosines with read depth <10 or a read quality score <30. The transcriptome-wide A-to-I editing frequency was calculated independently for each biological replicate as:

(number of reads in which an adenosine was called as a guanosine)/(total number of reads covering all analyzed adenosines).

Purification of SpyMac Cas nuclease protein

SpyMac Cas nuclease protein was cloned into the expression plasmid pD881-SR (Atum, Cat. No. FPB-27E-269). The resulting plasmid was transformed into BL21 Star DE3 competent cells (Thermo Fisher, Cat. No. C601003). Colonies were picked for overnight growth in terrific broth (TB) and 25 μg/ml kanamycin at 37°C. The next day, 2 liters of prewarmed TB were inoculated with overnight culture at a starting OD_{600} of 0.05. Cells were shaken at 37°C for ~2.5 hours until the OD_{600} was ~1.5. Cultures were cold-shocked in an ice-water slurry for 1 hour, after which L-rhamnose was added to a final concentration of 0.8% to induce protein production. Cultures were then incubated at 18°C with shaking for 24 hours to produce protein. After induction, cells were pelleted and flash-frozen in liquid nitrogen and stored at −80°C. The next day, cells were resuspended in 30 ml cold lysis buffer (1 M NaCl, 100 mM Tris-HCl pH 7.0, 5 mM tris(2-carboxyethyl)phosphine (TCEP), 20% glycerol, with five tablets of cOmplete, EDTA-free protease inhibitor cocktail (Millipore Sigma, Cat. No. 4693132001). Cells were passed three times through a homogenizer (Avestin Emulsiflex-C3) at ~18,000 psi to lyse. Cell debris was pelleted for 20 min using a 20,000*g* centrifugation at 4°C. Supernatant was collected and spiked with 40 mM imidazole, followed by a 1-hour incubation at 4°C with 1 ml of Ni-NTA resin slurry (G Bioscience Cat. No. 786-940, prewashed once with lysis buffer). Protein-bound resin was washed twice with 12 ml of lysis buffer in a gravity column at 4°C. Protein was eluted in 3 ml of elution buffer (300 mM imidazole, 500 mM NaCl, 100 mM Tris-HCl pH 7.0, 5 mM TCEP, 10% glycerol). Eluted protein was diluted in 40 ml of low-salt buffer (100 mM Tris-HCl, pH 7.0, 1 mM TCEP, 20% glycerol) just before loading into a 50 ml Akta Superloop for ion exchange purification on the Akta Pure25 FPLC. Ion exchange chromatography was conducted on a 5 ml GE Healthcare HiTrap SP HP pre-packed column (Cat. No. 17115201). After washing the column with low-salt buffer, the diluted protein was flowed through the column to bind. The column was then washed in 15 ml of low-salt buffer before being subjected to an increasing gradient to a maximum of 80% high-salt buffer (1 M NaCl, 100 mM Tris-HCl, pH 7.0, 5 mM TCEP, 20% glycerol) over the course of 50 ml, at a flow rate of 5 ml per minute. 1-ml fractions were collected during this ramp to high-salt buffer. Peaks were assessed by SDS–polyacrylamide gel electrophoresis to identify fractions containing the desired protein, which were concentrated first using an Amicon Ultra 15-ml centrifugal filter (100-kDa cutoff, Cat. No. UFC910024), followed by a 0.5-ml 100-kDa cutoff Pierce concentrator (Cat. No. 88503). Concentrated protein was quantified using a BCA assay and determined to be 12.6 mg/ml (Thermo Fisher, Cat. No. 23227).

CIRCLE-seq off-target editing analysis

Off-target analysis using CIRCLE-seq was performed as previously described (*67*, *129*). Briefly, genomic DNA from HEK293T cells or NIH3T3 cells was isolated using Gentra Puregene Kit (Qiagen) according to manufacturer's instructions. Purified genomic DNA was sheared with a Covaris S2 instrument to an average length of 300 base pairs (bp). The fragmented DNA was end-repaired, poly-A tailed, and ligated to an uracil-containing stem-loop adaptor using the KAPA HTP Library Preparation Kit, PCR Free (KAPA Biosystems). Adaptor-ligated DNA was treated with Lambda Exonuclease (NEB) and *E. coli* Exonuclease I (NEB), then with USER enzyme (NEB) and T4 polynucleotide kinase (NEB). Intramolecular circularization of the DNA was performed with T4 DNA ligase (NEB) and residual linear DNA was degraded by Plasmid-Safe ATP-dependent DNase (Lucigen). In vitro cleavage reactions were performed with 250 ng of Plasmid-Safe ATP-dependent DNase-treated circularized DNA, 90 nM of SpyMac Cas9 nuclease protein, Cas9 nuclease buffer (NEB), and 90 nM of synthetic chemically modified sgRNA (Synthego), in a total volume of 100 μl. Cleaved products were poly-A tailed, ligated with a hairpin adaptor (NEB), treated with USER enzyme (NEB), and amplified by PCR with barcoded universal primers NEBNext Multiplex Oligos for Illumina (NEB), using Kapa HiFi Polymerase (KAPA Biosystems). Libraries were sequenced with 150-bp paired-end reads on an Illumina MiSeq instrument. CIRCLE-seq data analyses were performed using open-source CIRCLE-seq analysis software and default recommended parameters (https://github.com/tsailabSJ/circleseq).

Husbandry of Δ7SMA mice

All experiments in animals were approved by the Institutional and Animal Care and Use Committee of the Broad Institute of MIT and Harvard and Ohio State University (OSU). Δ7SMA heterozygous mice (*Smn*$^{+/-}$; *SMN2*$^{+/+}$; *SMN*Δ*7*$^{+/+}$) were purchased from the Jackson Laboratory (005025) (*54*) and maintained in the Broad Institute and OSU vivaria according to recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. Pairs of Δ7SMA heterozygotes were crossed to generate Δ7 SMA mice (*Smn*$^{-/-}$; *SMN2*$^{+/+}$; *SMN*Δ*7*$^{+/+}$). On date of birth (PND0), pups were microtattooed on the foot pads (Aramis) with animal-grade permanent ink (Ketchum) using a sterile hypodermic needle (BD) to enable identification of individual pups. Subsequently, biopsies of ~1 mm tissue were taken from the tail using a sterile blade, lysed for genomic DNA extraction, and used for genotyping by PCR. Litter size was controlled to five pups, including 1 to 3 homozygous mutants, by culling and cross-fostering among same-age mice. Mice of both sexes were included in the study, although sex has been reported to not have a substantial impact on the phenotype of SMA mice (TreatNMD SOP Code: SMA_M.2.2.003).

Electrophysiology experiments were performed at OSU. All other animal studies were performed at the Broad Institute unless indicated otherwise in the text. At the Broad Institute, the mean birthweight of heterozygous animals was 1.7 ± 0.1 g, and 1.5 ± 0.1 g for SMA pups, and any animal weighing <1.5 g at time of birth was excluded from the study. The average weight of SMA neonates at injection on PND0 at the Broad Institute was 1.6 ± 0.2 g. At OSU, the mean birthweight of heterozygous animals on the day of birth was 1.3 ± 0.1 g and 1.2 ± 0.1 g for SMA pups, and any SMA, heterozygous, or wild-type pup weighing ≤1.0 g at time of birth was excluded from the study. The average weight of SMA neonates at injection on PND0 at OSU was 1.3 ± 0.13 g, and animals were injected with 3.3×10^{13} vg/kg of the dual AAV-ABE vectors. By facility, each litter was subjected to the same exclusion criterion (TreatNMD SOP Code: SMA_M.2.2.003). Cohort sizes were chosen on the basis of prior experience with these animals, known to allow for determination of statistical significance. Animals were monitored daily for morbidity and mortality and weighed every other day from day of birth.

Intracerebroventricular injections

Neonatal ICV injections were performed as previously described (*73*, *130*). Briefly, glass capillaries (Drummond 5–000-1001-X10) were pulled to a tip diameter of ~100 μm. High-titer qualified AAV was obtained through the Viral Vector Core at UMass Medical School and concentrated using Amicon Ultra-15 centrifugal filter units (Millipore), quantified by qPCR (AAVpro Titration Kit v.2, Clontech), and stored at 4°C until use. For injection, a small amount of Fast Green was added to the AAV injection solution to assess ventricle targeting. The injection solution was loaded through front-filling using the included Drummond plungers. Δ7SMA pups were anesthetized by placement on ice for 2 to 3 min, until they were immobile and unresponsive to a toe pinch. Up to 4.5 μl of injection mix was injected freehand into each ventricle on PND0 and PND1.

Immunofluorescence imaging of spinal cord sections

For immunofluorescence staining of transduced spinal motor neurons, Δ7SMA mice

were perfused at 25 weeks with ice-cold PBS and ice-cold 4% paraformaldehyde (PFA), the CNS was exposed, and the whole carcass was fixed overnight in 4% PFA. Whole spinal cord was isolated and fixed in 4% PFA overnight, then consecutively transferred to 10%, 20%, and 30% sucrose in three overnight incubations before embedding in OCT for long-term storage at −80°C. Embedded tissue was cryosectioned and stained with goat anti-ChAT (Millipore AB144P), mouse anti-NeuN (EMD Millipore MAB377), mouse anti-GFAP (Sigma-Aldrich MAB3402), rabbit anti-GFP (Thermo scientific A-11122), and Alexa-Fluor secondary antibodies (Life Technologies) and imaged on an SP8 confocal microscope (Leica).

Nuclear isolation and sorting of tissues

Tissue harvest and nuclear isolation was performed as previously described (*73*). Briefly, deceased Δ7SMA mice were stored at −80°C until dissection of the brain and spinal cord tissue. The cortex and cerebella were separated from the brain postmortem using surgical scissors. Hemispheres were separated using a scalpel, and the cortex was separated from underlying midbrain tissue with a curved spatula. For nuclear isolation, dissected tissue was homogenized using a glass dounce homogenizer (Sigma D8938) (20 strokes with pestle A followed by 20 strokes with pestle B) in 2 ml ice-cold EZ-PREP buffer (Sigma NUC-101). Samples were incubated for 5 min with an additional 2 ml EZ-PREP buffer. Nuclei were centrifuged at 500*g* for 5 min, and the supernatant removed. For spinal cord tissue, wash steps were repeated 10 times. Samples were resuspended with gentle pipetting in 4 ml ice-cold nuclei suspension buffer (NSB) consisting of 100 μg/ml BSA and 3.33 μM Vybrant DyeCycle Ruby (Thermo Fisher) in PBS and centrifuged at 500*g* for 5 min. The supernatant was removed, and nuclei were resuspended in 1 to 2 ml NSB, passed through a 35-μm strainer, and sorted into 200 μl Agencourt DNAdvance lysis buffer using a MoFlo Astrios (Beckman Coulter) at the Broad Institute flow cytometry core. All steps were performed on ice or at 4°C. Genomic DNA was purified according to the DNAdvance (Agencourt) instructions for 200 μl volume.

Behavioral assays

Righting reflex was recorded on PND7 by placing neonates on their backs and recording with a stopwatch, up to a maximum of 30 s, the duration of time that it took for the mice to right themselves. For inverted screen testing, we subjected juvenile mice to the horizontal grid test for mice (Maze Engineers) on PND25 by placing the animals on a wire-mesh screen, which the mice are capable of gripping, then inverting the screen over the course of 2 s, animal head first, over a padded surface made of 4- to 5-cm-high bedding. The time it took for the animal to fall onto the bedding was recorded with a stopwatch. Each mouse was assessed with three measurements. The procedure concluded when the animal fell onto the bedding, or if the animal exceeded 120 s for the measurement, in which case the screen was reverted so that the mouse was upright, and the mouse was manually removed from the screen.

Voluntary movement of adult mice was recorded on PND40 by open field testing (Omnitech Electronics). Mice were brought into the testing room under normal lighting conditions and allowed 30 to 60 min of acclimation. The animals were placed into the locomotor activity chamber with infrared beams crossing the x, y, and z axes that plotted their ambulatory and fine motor movements and rearing behavior. Recordings are analyzed using Fusion 5.1 SuperFlex software.

Electrophysiological measurements

Compound muscle action potential (CMAP) and motor unit number estimate (MUNE) measurements were performed as previously described (*131*). Briefly, at PND12, the right sciatic nerve was stimulated with a pair of insulated 28-gauge monopolar needles (Teca, Oxford Instruments Medical, NY) placed in proximity to the sciatic nerve in the proximal hind limb. Recording electrodes consisted of a pair of fine ring wire electrodes (Alpine Biomed, Skovlunde, Denmark). The active recording electrode (E1) was placed distal to the knee joint over the proximal portion of the triceps surae muscle, and the reference electrode (E2) was placed over the metatarsal region of the foot. A disposable strip electrode (Carefusion, Middleton, WI) was placed on the tail to serve as the ground electrode. For CMAP, supramaximal responses were generated maintaining stimulus currents <10 mA, and baseline-to-peak amplitude measurements were made.

For MUNE, an incremental stimulus technique similar to a previously described procedure was used (*131*). Submaximal stimulation was used to obtain 10 incremental responses to calculate the average single motor unit potential (SMUP) amplitude. The first increment was obtained by delivering square wave stimulations at 1 Hz at an intensity between 0.21 and 0.70 mA to obtain the minimal all-or-none SMUP response. If the initial response did not occur with stimulus intensity between 0.21 and 0.70 mA, the stimulating cathode position was adjusted either closer to or farther away from the position of the sciatic nerve in the proximal thigh to decrease or increase the required stimulus intensity, respectively. This first incremental response was accepted if three duplicate responses were observed. To obtain the subsequent incremental responses, the stimulation intensity was adjusted in 0.03-mA steps, and incremental responses were distinguished visually in real-time to obtain nine additional increments. To be accepted, each increment was required to be: (i) observed for a total of three duplicate responses, (ii) visually distinct from the prior increment, and (iii) at least 25 μV larger than the prior increment. The peak-to-peak amplitude of each individual incremental response was calculated by subtracting the amplitude of the prior response. The 10 incremental values were averaged to estimate average peak-to-peak SMUP amplitude. The maximum CMAP amplitude (peak-to-peak) was divided by the average SMUP amplitude to yield the MUNE.

Statistical analysis

Welch's two-tailed t tests were used to compare sequencing, splicing, mRNA levels, and immunostaining data. Error bars represent standard deviations of ≥3 independent biological replicates. Root mean squared error (RMSE) and Pearson's r-correlation were used for correlation analysis of predicted and observed genome editing outcomes, where appropriate. Kruskal-Wallis tests were used to compare physiology measurements and behaviors of mouse cohorts under experimental conditions. Mann-Whitney tests were used to compare multiparametric measurements of voluntary behaviors of mouse cohorts. The logrank Mantel-Cox test was used to compare body weight and life span of mouse cohorts. All statistical tests were calculated by GraphPad Prism 9.4.1 and Microsoft Excel v16.64.

REFERENCES AND NOTES

1. S. Lefebvre *et al*., Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* **80**, 155–165 (1995). doi: 10.1016/0092-8674(95)90460-3; pmid: 7813012
2. E. A. Sugarman *et al*., Pan-ethnic carrier screening and prenatal diagnosis for spinal muscular atrophy: Clinical laboratory analysis of >72,400 specimens. *Eur. J. Hum. Genet.* **20**, 27–32 (2012). doi: 10.1038/ejhg.2011.134; pmid: 21811307
3. D. F. Roberts, J. Chavez, S. D. M. Court, The genetic component in child mortality. *Arch. Dis. Child.* **45**, 33–38 (1970). doi: 10.1136/adc.45.239.33; pmid: 4245389
4. B. Boda *et al*., Survival motor neuron *SMN1* and *SMN2* gene promoters: Identical sequences and differential expression in neurons and non-neuronal cells. *Eur. J. Hum. Genet.* **12**, 729–737 (2004). doi: 10.1038/sj.ejhg.5201217; pmid: 15162126
5. C. F. Rochette, N. Gilbert, L. R. Simard, *SMN* gene duplication and the emergence of the *SMN2* gene occurred in distinct hominids: *SMN2* is unique to *Homo sapiens*. *Hum. Genet.* **108**, 255–266 (2001). doi: 10.1007/s004390100473; pmid: 11354640
6. C. L. Lorson, E. Hahnen, E. J. Androphy, B. Wirth, A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6307–6311 (1999). doi: 10.1073/pnas.96.11.6307; pmid: 10339583
7. U. R. Monani *et al*., A single nucleotide difference that alters splicing patterns distinguishes the *SMA* gene *SMN1* from the copy gene *SMN2*. *Hum. Mol. Genet.* **8**, 1177–1183 (1999). doi: 10.1093/hmg/8.7.1177; pmid: 10369862
8. S. Cho, G. Dreyfuss, A degron created by *SMN2* exon 7 skipping is a principal contributor to spinal muscular atrophy severity. *Genes Dev.* **24**, 438–442 (2010). doi: 10.1101/gad.1884910; pmid: 20194437

9. J. Vitte *et al*., Refined characterization of the expression and stability of the SMN gene products. *Am. J. Pathol.* **171**, 1269–1280 (2007). doi: 10.2353/ajpath.2007.070399; pmid: 17717146
10. B. G. Burnett *et al*., Regulation of SMN protein stability. *Mol. Cell. Biol.* **29**, 1107–1115 (2009). doi: 10.1128/MCB.01262-08; pmid: 19103745
11. J. M. Cobben *et al*., Survival in SMA type I: A prospective analysis of 34 consecutive cases. *Neuromuscul. Disord.* **18**, 541–544 (2008). doi: 10.1016/j.nmd.2008.05.008; pmid: 18579378
12. S. J. Kolb *et al*., Natural history of infantile-onset spinal muscular atrophy. *Ann. Neurol.* **82**, 883–891 (2017). doi: 10.1002/ana.25101; pmid: 29149772
13. J. R. Mendell *et al*., Single-dose gene-replacement therapy for spinal muscular atrophy. *N. Engl. J. Med.* **377**, 1713–1722 (2017). doi: 10.1056/NEJMoa1706198; pmid: 29091557
14. E. W. Ottesen, ISS-N1 makes the first FDA-approved drug for spinal muscular atrophy. *Transl. Neurosci.* **8**, 1–6 (2017). doi: 10.1515/tnsci-2017-0001; pmid: 28400976
15. S. M. Hoy, Onasemnogene Abeparvovec: First global approval. *Drugs* **79**, 1255–1262 (2019). doi: 10.1007/s40265-019-01162-5; pmid: 31270752
16. L. E. Kernochan *et al*., The role of histone acetylation in SMN gene expression. *Hum. Mol. Genet.* **14**, 1171–1182 (2005). doi: 10.1093/hmg/ddi130; pmid: 15772088
17. C. d'Ydewalle *et al*., The antisense transcript *SMN-AS1* regulates SMN expression and is a novel therapeutic target for spinal muscular atrophy. *Neuron* **93**, 66–79 (2017). doi: 10.1016/j.neuron.2016.11.033; pmid: 28017471
18. C. J. Woo *et al*., Gene activation of SMN by selective disruption of lncRNA-mediated recruitment of PRC2 for the treatment of spinal muscular atrophy. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E1509–E1518 (2017). doi: 10.1073/pnas.1616521114; pmid: 28193854
19. M. Cucchiarini, H. Madry, E. F. Terwilliger, Enhanced expression of the central survival of motor neuron (SMN) protein during the pathogenesis of osteoarthritis. *J. Cell. Mol. Med.* **18**, 115–124 (2014). doi: 10.1111/jcmm.12170; pmid: 24237934
20. H. M. Blauw *et al*., *SMN1* gene duplications are associated with sporadic ALS. *Neurology* **78**, 776–780 (2012). doi: 10.1212/WNL.0b013e318249f697; pmid: 22323753
21. M. Van Alstyne *et al*., Gain of toxic function by long-term AAV9-mediated SMN overexpression in the sensorimotor circuit. *Nat. Neurosci.* **24**, 930–940 (2021). doi: 10.1038/s41593-021-00827-3; pmid: 33795885
22. H. Chaytow, K. M. E. Faller, Y. T. Huang, T. H. Gillingwater, Spinal muscular atrophy: From approved therapies to future therapeutic targets for personalized medicine. *Cell Rep. Med.* **2**, 100346 (2021). doi: 10.1016/j.xcrm.2021.100346; pmid: 34337562
23. C. A. Chiriboga *et al*., Results from a phase 1 study of nusinersen (ISIS-SMN_{Rx}) in children with spinal muscular atrophy. *Neurology* **86**, 890–897 (2016). doi: 10.1212/WNL.0000000000002445; pmid: 26865511
24. G. Baranello *et al*., Risdiplam in type 1 spinal muscular atrophy. *N. Engl. J. Med.* **384**, 915–923 (2021). doi: 10.1056/NEJMoa2009965; pmid: 33626251
25. S. Lefebvre *et al*., Correlation between severity and SMN protein level in spinal muscular atrophy. *Nat. Genet.* **16**, 265–269 (1997). doi: 10.1038/ng0797-265; pmid: 9207792
26. D. D. Coovert *et al*., The survival motor neuron protein in spinal muscular atrophy. *Hum. Mol. Genet.* **6**, 1205–1214 (1997). doi: 10.1093/hmg/6.8.1205; pmid: 9259265
27. A. L. Patrizi *et al*., SMN protein analysis in fibroblast, amniocyte and CVS cultures from spinal muscular atrophy patients and its relevance for diagnosis. *Eur. J. Hum. Genet.* **7**, 301–309 (1999). doi: 10.1038/sj.ejhg.5200286; pmid: 10234506
28. D. M. Ramos *et al*., Age-dependent SMN expression in disease-relevant tissue and implications for SMA treatment. *J. Clin. Invest.* **129**, 4817–4831 (2019). doi: 10.1172/JCI124120; pmid: 31589162
29. S. Kariya *et al*., Requirement of enhanced Survival Motoneuron protein imposed during neuromuscular junction maturation. *J. Clin. Invest.* **124**, 785–800 (2014). doi: 10.1172/JCI72017; pmid: 24463453
30. H. Ratni *et al*., Discovery of risdiplam, a selective survival of motor neuron-2 (*SMN2*) gene splicing modifier for the treatment of spinal muscular atrophy (SMA). *J. Med. Chem.* **61**, 6501–6517 (2018). doi: 10.1021/acs.jmedchem.8b00741; pmid: 30044619
31. C. D. Wurster, A. C. Ludolph, Nusinersen for spinal muscular atrophy. *Ther. Adv. Neurol. Disord.* **11**, 1756285618754459 (2018). doi: 10.1177/1756285618754459; pmid: 29568328
32. K. Meyer *et al*., Improving single injection CSF delivery of AAV9-mediated gene therapy for SMA: A dose-response study in mice and nonhuman primates. *Mol. Ther.* **23**, 477–487 (2015). doi: 10.1038/mt.2014.210; pmid: 25358252
33. N. Armbruster *et al*., Efficacy and biodistribution analysis of intracerebroventricular administration of an optimized scAAV9-SMN1 vector in a mouse model of spinal muscular atrophy. *Mol. Ther. Methods Clin. Dev.* **3**, 16060 (2016). doi: 10.1038/mtm.2016.60; pmid: 27652289
34. A. K. Bevan *et al*., Systemic gene delivery in large species for targeting spinal cord, brain, and peripheral tissues for pediatric disorders. *Mol. Ther.* **19**, 1971–1980 (2011). doi: 10.1038/mt.2011.157; pmid: 21811247
35. G. Thomsen *et al*., Biodistribution of onasemnogene abeparvovec DNA, mRNA and SMN protein in human tissue. *Nat. Med.* **27**, 1701–1711 (2021). doi: 10.1038/s41591-021-01483-7; pmid: 34608334
36. A. Das *et al*., Epigenetic silencing of recombinant adeno-associated virus genomes by NP220 and the HUSH complex. *J. Virol.* **96**, e0203921 (2022). doi: 10.1128/jvi.02039-21; pmid: 34878926
37. J. A. Greig *et al*., Loss of transgene expression limits liver gene therapy in primates. bioRxiv 2022.03.24.485675 [Preprint] (2022). https://doi.org/10.1101/2022.03.24.485675.
38. N. K. Singh, N. N. Singh, E. J. Androphy, R. N. Singh, Splicing of a critical exon of human *Survival Motor Neuron* is regulated by a unique silencer element located in the last intron. *Mol. Cell. Biol.* **26**, 1333–1346 (2006). doi: 10.1128/MCB.26.4.1333-1346.2006; pmid: 16449646
39. Y. Hua, T. A. Vickers, H. L. Okunola, C. F. Bennett, A. R. Krainer, Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects *SMN2* splicing in transgenic mice. *Am. J. Hum. Genet.* **82**, 834–848 (2008). doi: 10.1016/j.ajhg.2008.01.014; pmid: 18371932
40. J. J. Li *et al*., Disruption of splicing-regulatory elements using CRISPR/Cas9 to rescue spinal muscular atrophy in human iPSCs and mice. *Natl. Sci. Rev.* **7**, 92–101 (2019). doi: 10.1093/nsr/nwz131; pmid: 34691481
41. M. Zhou *et al*., Targeted-deletion of a tiny sequence via prime editing to restore SMN expression. *Int. J. Mol. Sci.* **23**, 7941 (2022). doi: 10.3390/ijms23147941; pmid: 35887289
42. M. W. Shen *et al*., Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* **563**, 646–651 (2018). doi: 10.1038/s41586-018-0686-x; pmid: 30405244
43. R. T. Walton, K. A. Christie, M. N. Whittaker, B. P. Kleinstiver, Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* **368**, 290–296 (2020). doi: 10.1126/science.aba8853; pmid: 32217751
44. S. M. Miller *et al*., Continuous evolution of SpCas9 variants compatible with non-G PAMs. *Nat. Biotechnol.* **38**, 471–481 (2020). doi: 10.1038/s41587-020-0412-8; pmid: 32042170
45. P. Chatterjee *et al*., A Cas9 with PAM recognition for adenine dinucleotides. *Nat. Commun.* **11**, 2474 (2020). doi: 10.1038/s41467-020-16117-8; pmid: 32424114
46. H. Nishimasu *et al*., Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* **361**, 1259–1262 (2018). doi: 10.1126/science.aas9129; pmid: 30166441
47. N. Rodriguez-Muela *et al*., Single-cell analysis of SMN reveals its broader role in neuromuscular disease. *Cell Rep.* **18**, 1484–1498 (2017). doi: 10.1016/j.celrep.2017.01.035; pmid: 28178525
48. E. C. Wolstencroft, V. Mattis, A. A. Bajer, P. J. Young, C. L. Lorson, A non-sequence-specific requirement for SMN protein activity: The role of aminoglycosides in inducing elevated SMN protein levels. *Hum. Mol. Genet.* **14**, 1199–1210 (2005). doi: 10.1093/hmg/ddi131; pmid: 15790598
49. C. Madocsai, S. R. Lim, T. Geib, B. J. Lam, K. J. Hertel, Correction of SMN2 pre-mRNA splicing by antisense U7 small nuclear RNAs. *Mol. Ther.* **12**, 1013–1022 (2005). doi: 10.1016/j.ymthe.2005.08.022; pmid: 16226920
50. M. Arbab *et al*., Determinants of base editing outcomes from target library analysis and machine learning. *Cell* **182**, 463–480.e30 (2020). doi: 10.1016/j.cell.2020.05.037; pmid: 32533916
51. N. N. Singh, R. N. Singh, E. J. Androphy, Modulating role of RNA structure in alternative splicing of a critical exon in the spinal muscular atrophy genes. *Nucleic Acids Res.* **35**, 371–389 (2007). doi: 10.1093/nar/gkl1050; pmid: 17170000
52. B. L. Oakes *et al*., CRISPR-Cas9 circular permutants as programmable scaffolds for genome modification. *Cell* **176**, 254–267.e16 (2019). doi: 10.1016/j.cell.2018.11.052; pmid: 30633905
53. T. P. Huang *et al*., Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. *Nat. Biotechnol.* **37**, 626–631 (2019). doi: 10.1038/s41587-019-0134-y; pmid: 31110355
54. T. T. Le *et al*., SMNΔ7, the major product of the centromeric survival motor neuron (*SMN2*) gene, extends survival in mice with spinal muscular atrophy and associates with full-length SMN. *Hum. Mol. Genet.* **14**, 845–857 (2005). doi: 10.1093/hmg/ddi078; pmid: 15703193
55. N. N. Singh, E. J. Androphy, R. N. Singh, An extended inhibitory context causes skipping of exon 7 of *SMN2* in spinal muscular atrophy. *Biochem. Biophys. Res. Commun.* **315**, 381–388 (2004). doi: 10.1016/j.bbrc.2004.01.067; pmid: 14766219
56. L. Cartegni, M. L. Hastings, J. A. Calarco, E. de Stanchina, A. R. Krainer, Determinants of exon 7 splicing in the spinal muscular atrophy genes, *SMN1* and *SMN2*. *Am. J. Hum. Genet.* **78**, 63–77 (2006). doi: 10.1086/498853; pmid: 16385450
57. B. Schrank *et al*., Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9920–9925 (1997). doi: 10.1073/pnas.94.18.9920; pmid: 9275227
58. U. R. Monani, Spinal muscular atrophy: A deficiency in a ubiquitous protein; a motor neuron-specific disease. *Neuron* **48**, 885–895 (2005). doi: 10.1016/j.neuron.2005.12.001; pmid: 16364894
59. R. N. Singh, M. D. Howell, E. W. Ottesen, N. N. Singh, Diverse role of survival motor neuron protein. *Biochim. Biophys. Acta. Gene Regul. Mech.* **1860**, 299–315 (2017). doi: 10.1016/j.bbagrm.2016.12.008; pmid: 28095296
60. M. Briese *et al*., Deletion of *smn-1*, the *Caenorhabditis elegans* ortholog of the spinal muscular atrophy gene, results in locomotor dysfunction and reduced lifespan. *Hum. Mol. Genet.* **18**, 97–104 (2009). doi: 10.1093/hmg/ddn320; pmid: 18829666
61. M. F. Richter *et al*., Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.* **38**, 883–891 (2020). doi: 10.1038/s41587-020-0453-z; pmid: 32433547
62. J. L. Doman, A. Raguram, G. A. Newby, D. R. Liu, Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. *Nat. Biotechnol.* **38**, 620–628 (2020). doi: 10.1038/s41587-020-0414-6; pmid: 32042165
63. H. A. Rees, C. Wilson, J. L. Doman, D. R. Liu, Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci. Adv.* **5**, eaax5717 (2019). doi: 10.1126/sciadv.aax5717; pmid: 31086823
64. Y. Yu *et al*., Cytosine base editors with minimized unguided DNA and RNA off-target events and high on-target activity. *Nat. Commun.* **11**, 2052 (2020). doi: 10.1038/s41467-020-15887-5; pmid: 32345976
65. C. Zhou *et al*., Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature* **571**, 275–278 (2019). doi: 10.1038/s41586-019-1314-0; pmid: 31181567
66. K. Clement *et al*., CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **37**, 224–226 (2019). doi: 10.1038/s41587-019-0032-3; pmid: 30809026
67. S. Q. Tsai *et al*., CIRCLE-seq: A highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat. Methods* **14**, 607–614 (2017). doi: 10.1038/nmeth.4278; pmid: 28459458
68. Y. Zhao *et al*., Sequence-specific inhibition of microRNA via CRISPR/CRISPRi system. *Sci. Rep.* **4**, 3943 (2014). doi: 10.1038/srep03943; pmid: 24487629
69. K. D. Foust *et al*., Intravascular AAV9 preferentially targets neonatal neurons and adult astrocytes. *Nat. Biotechnol.* **27**, 59–65 (2009). doi: 10.1038/nbt.1515; pmid: 19098898
70. S. L. Hammond, A. N. Leek, E. H. Richman, R. B. Tjalkens, Cellular selectivity of AAV serotypes for gene delivery in neurons and astrocytes by neonatal intracerebroventricular injection. *PLOS ONE* **12**, e0188830 (2017). doi: 10.1371/journal.pone.0188830; pmid: 29244806
71. S. N. Mathiesen, J. L. Lock, L. Schoderboeck, W. C. Abraham, S. M. Hughes, CNS transduction benefits of AAV-PHP.eB over AAV9 are dependent on administration route and mouse

strain. *Mol. Ther. Methods Clin. Dev.* **19**, 447–458 (2020). doi: 10.1016/j.omtm.2020.10.011; pmid: 33294493
72. W. Arnold *et al*., The neuromuscular impact of symptomatic SMN restoration in a mouse model of spinal muscular atrophy. *Neurobiol. Dis.* **87**, 116–123 (2016). doi: 10.1016/j.nbd.2015.12.014; pmid: 26733414
73. J. M. Levy *et al*., Cytosine and adenine base editing of the brain, liver, retina, heart and skeletal muscle of mice via adeno-associated viruses. *Nat. Biomed. Eng.* **4**, 97–110 (2020). doi: 10.1038/s41551-019-0501-5; pmid: 31937940
74. D. J. Schuster *et al*., Biodistribution of adeno-associated virus serotype 9 (AAV9) vector after intrathecal and intravenous delivery in mouse. *Front. Neuroanat.* **8**, 42 (2014). doi: 10.3389/fnana.2014.00042; pmid: 24959122
75. L. Swiech *et al*., In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat. Biotechnol.* **33**, 102–106 (2015). doi: 10.1038/nbt.3055; pmid: 25326897
76. L. Cartegni, A. R. Krainer, Disruption of an SF2/ASF-dependent exonic splicing enhancer in *SMN2* causes spinal muscular atrophy in the absence of *SMN1*. *Nat. Genet.* **30**, 377–384 (2002). doi: 10.1038/ng854; pmid: 11925564
77. E. J. N. Groen *et al*., Temporal and tissue-specific variability of SMN protein levels in mouse models of spinal muscular atrophy. *Hum. Mol. Genet.* **27**, 2851–2862 (2018). doi: 10.1093/hmg/ddy195; pmid: 29790918
78. M. Karlsson *et al*., A single–cell type transcriptomics map of human tissues. *Sci. Adv.* **7**, eabh2169 (2021). doi: 10.1126/sciadv.abh2169; pmid: 34321199
79. P. Akcakaya *et al*., In vivo CRISPR editing with no detectable genome-wide off-target mutations. *Nature* **561**, 416–419 (2018). doi: 10.1038/s41586-018-0500-9; pmid: 30209390
80. J. R. Davis *et al*., Efficient in vivo base editing via single adeno-associated viruses with size-optimized genomes encoding compact adenine base editors. *Nat. Biomed. Eng.* **6**, 1272–1283 (2022). doi: 10.1038/s41551-022-00911-4; pmid: 35902773
81. T. Rothgangl *et al*., In vivo adenine base editing of *PCSK9* in macaques reduces LDL cholesterol levels. *Nat. Biotechnol.* **39**, 949–957 (2021). doi: 10.1038/s41587-021-00933-4; pmid: 34012094
82. W. D. Arnold *et al*., Electrophysiological biomarkers in spinal muscular atrophy: Proof of concept. *Ann. Clin. Transl. Neurol.* **1**, 34–44 (2014). doi: 10.1002/acn3.23; pmid: 24511555
83. A. Tscherter *et al*., Evaluation of real-life outcome data of patients with spinal muscular atrophy treated with nusinersen in Switzerland. *Neuromuscul. Disord.* **32**, 399–409 (2022). doi: 10.1016/j.nmd.2022.02.001; pmid: 35337708
84. T. Dangouloff, L. Servais, Clinical evidence supporting early treatment of patients with spinal muscular atrophy: current perspectives. *Ther. Clin. Risk Manag.* **15**, 1153–1161 (2019). doi: 10.2147/TCRM.S172291; pmid: 31632042
85. K. A. Strauss *et al*., Onasemnogene abeparvovec for presymptomatic infants with two copies of *SMN2* at risk for spinal muscular atrophy type 1: The Phase III SPR1NT trial. *Nat. Med.* **28**, 1381–1389 (2022). doi: 10.1038/s41591-022-01866-4; pmid: 35715566
86. D. C. De Vivo *et al*., Nusinersen initiated in infants during the presymptomatic stage of spinal muscular atrophy: Interim efficacy and safety results from the Phase 2 NURTURE study. *Neuromuscul. Disord.* **29**, 842–856 (2019). doi: 10.1016/j.nmd.2019.09.007; pmid: 31704158
87. K. L. Robbins, J. J. Glascock, E. Y. Osman, M. R. Miller, C. L. Lorson, Defining the therapeutic window in a severe animal model of spinal muscular atrophy. *Hum. Mol. Genet.* **23**, 4559–4568 (2014). doi: 10.1093/hmg/ddu169; pmid: 24722206
88. T. T. Le *et al*., Temporal requirement for high SMN expression in SMA mice. *Hum. Mol. Genet.* **20**, 3578–3591 (2011). doi: 10.1093/hmg/ddr275; pmid: 21672919
89. K. Flurkey, J. M. Currer, D. E. Harrison, Mouse models in aging research. *Mouse Biomed. Res.* **3**, 637–672 (2007).
90. E. Mercuri *et al*., Nusinersen versus sham control in later-onset spinal muscular atrophy. *N. Engl. J. Med.* **378**, 625–635 (2018). doi: 10.1056/NEJMoa1710504; pmid: 29443664
91. R. S. Finkel *et al*., One-time intrathecal (IT) administration of AVXS-101 IT gene-replacement therapy for spinal muscular atrophy: Phase 1 Study (STRONG). *Neurology* **94** (suppl. 15), 2493 (2020).
92. C. M. Lutz *et al*., Postsymptomatic restoration of SMN rescues the disease phenotype in a mouse model of severe spinal muscular atrophy. *J. Clin. Invest.* **121**, 3029–3041 (2011). doi: 10.1172/JCI57291; pmid: 21785219
93. Z. Wang *et al*., Rapid and highly efficient transduction by double-stranded adeno-associated virus vectors *in vitro* and *in vivo*. *Gene Ther.* **10**, 2105–2111 (2003). doi: 10.1038/sj.gt.3302133; pmid: 14625564
94. B. Hauck, W. Zhao, K. High, W. Xiao, Intracellular viral processing, not single-stranded DNA accumulation, is crucial for recombinant adeno-associated virus transduction. *J. Virol.* **78**, 13678–13686 (2004). doi: 10.1128/JVI.78.24.13678-13686.2004; pmid: 15564477
95. D. M. McCarty, P. E. Monahan, R. J. Samulski, Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Ther.* **8**, 1248–1254 (2001). doi: 10.1038/sj.gt.3301514; pmid: 11509958
96. Z. Feng *et al*., Pharmacologically induced mouse model of adult spinal muscular atrophy to evaluate effectiveness of therapeutics after disease onset. *Hum. Mol. Genet.* **25**, 964–975 (2016). doi: 10.1093/hmg/ddv629; pmid: 26758873
97. N. A. Naryshkin *et al*., *SMN2* splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science* **345**, 688–693 (2014). doi: 10.1126/science.1250127; pmid: 25104390
98. M. A. Passini *et al*., Antisense oligonucleotides delivered to the mouse CNS ameliorate symptoms of severe spinal muscular atrophy. *Sci. Transl. Med.* **3**, 72ra18 (2011). doi: 10.1126/scitranslmed.3001777; pmid: 21368223
99. R. S. Finkel *et al*., Nusinersen versus sham control in infantile-onset spinal muscular atrophy. *N. Engl. J. Med.* **377**, 1723–1732 (2017). doi: 10.1056/NEJMoa1702752; pmid: 29091570
100. R. I. Wadman *et al*., A comparative study of SMN protein and mRNA in blood and fibroblasts in patients with spinal muscular atrophy and healthy controls. *PLOS ONE* **11**, e0167087–e0167087 (2016). doi: 10.1371/journal.pone.0167087; pmid: 27893852
101. S. Shmakov *et al*., Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. *Mol. Cell* **60**, 385–397 (2015). doi: 10.1016/j.molcel.2015.10.008; pmid: 26593719
102. M. E. R. Butchbach, Copy number variations in the *survival motor neuron* genes: Implications for spinal muscular atrophy and other neurodegenerative diseases. *Front. Mol. Biosci.* **3**, 7 (2016). doi: 10.3389/fmolb.2016.00007; pmid: 27014701
103. T. Kato *et al*., Creation of mutant mice with megabase-sized deletions containing custom-designed breakpoints by means of the CRISPR/Cas9 system. *Sci. Rep.* **7**, 59 (2017). doi: 10.1038/s41598-017-00140-9; pmid: 28246396
104. N. M. Gaudelli *et al*., Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017). doi: 10.1038/nature24644; pmid: 29160308
105. A. C. Komor, Y. B. Kim, M. S. Packer, J. A. Zuris, D. R. Liu, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016). doi: 10.1038/nature17946; pmid: 27096365
106. N. M. Gaudelli *et al*., Directed evolution of adenine base editors with increased activity and therapeutic application. *Nat. Biotechnol.* **38**, 892–900 (2020). doi: 10.1038/s41587-020-0491-6; pmid: 32284586
107. P. E. McAndrew *et al*., Identification of proximal spinal muscular atrophy carriers and patients by analysis of SMN^T and SMN^C gene copy number. *Am. J. Hum. Genet.* **60**, 1411–1422 (1997). doi: 10.1086/515465; pmid: 9199562
108. M. D. Mailman *et al*., Molecular analysis of spinal muscular atrophy and modification of the phenotype by SMN2. *Genet. Med.* **4**, 20–26 (2002). doi: 10.1097/00125817-200201000-00004; pmid: 11839954
109. C. Ruhno *et al*., Complete sequencing of the *SMN2* gene in SMA patients detects *SMN* gene deletion junctions and variants in *SMN2* that modify the SMA phenotype. *Hum. Genet.* **138**, 241–256 (2019). doi: 10.1007/s00439-019-01983-0; pmid: 30788592
110. M. Calucho *et al*., Correlation between SMA type and *SMN2* copy number revisited: An analysis of 625 unrelated Spanish patients and a compilation of 2834 reported cases. *Neuromuscul. Disord.* **28**, 208–215 (2018). doi: 10.1016/j.nmd.2018.01.003; pmid: 29433793
111. T. W. Prior *et al*., A positive modifier of spinal muscular atrophy in the *SMN2* gene. *Am. J. Hum. Genet.* **85**, 408–413 (2009). doi: 10.1016/j.ajhg.2009.08.002; pmid: 19716110
112. A. Besse *et al*., AAV9-mediated expression of SMN restricted to neurons does not rescue the spinal muscular atrophy phenotype in mice. *Mol. Ther.* **28**, 1887–1901 (2020). doi: 10.1016/j.ymthe.2020.05.011; pmid: 32470325
113. A. K. Bevan *et al*., Early heart failure in the SMNΔ7 model of spinal muscular atrophy and correction by postnatal scAAV9-SMN delivery. *Hum. Mol. Genet.* **19**, 3895–3905 (2010). doi: 10.1093/hmg/ddq300; pmid: 20639395
114. M. Shababi *et al*., Cardiac defects contribute to the pathology of spinal muscular atrophy models. *Hum. Mol. Genet.* **19**, 4059–4071 (2010). doi: 10.1093/hmg/ddq329; pmid: 20696672
115. V. L. McGovern *et al*., SMN expression is required in motor neurons to rescue electrophysiological deficits in the SMNΔ7 mouse model of SMA. *Hum. Mol. Genet.* **24**, 5524–5541 (2015). doi: 10.1093/hmg/ddv283; pmid: 26206889
116. M. A. Passini *et al*., CNS-targeted gene therapy improves survival and motor function in a mouse model of spinal muscular atrophy. *J. Clin. Invest.* **120**, 1253–1264 (2010). doi: 10.1172/JCI41615; pmid: 20234094
117. Y. Hua *et al*., Peripheral SMN restoration is essential for long-term rescue of a severe spinal muscular atrophy mouse model. *Nature* **478**, 123–126 (2011). doi: 10.1038/nature10485; pmid: 21979052
118. S. L. Lipnick *et al*., Systemic nature of spinal muscular atrophy revealed by studying insurance claims. *PLOS ONE* **14**, e0213680 (2019). doi: 10.1371/journal.pone.0213680; pmid: 30870495
119. L. Villiger *et al*., Treatment of a metabolic liver disease by in vivo genome base editing in adult mice. *Nat. Med.* **24**, 1519–1525 (2018). doi: 10.1038/s41591-018-0209-1; pmid: 30297904
120. W. H. Yeh, H. Chiang, H. A. Rees, A. S. B. Edge, D. R. Liu, In vivo base editing of post-mitotic sensory cells. *Nat. Commun.* **9**, 2184 (2018). doi: 10.1038/s41467-018-04580-3; pmid: 29872041
121. D. A. Kuzmin *et al*., The clinical landscape for AAV gene therapies. *Nat. Rev. Drug Discov.* **20**, 173–174 (2021). doi: 10.1038/d41573-021-00017-7; pmid: 33495615
122. M. Arbab, S. Srinivasan, T. Hashimoto, N. Geijsen, R. I. I. Sherwood, Cloning-free CRISPR. *Stem Cell Reports* **5**, 908–917 (2015). doi: 10.1016/j.stemcr.2015.09.022; pmid: 26527385
123. H. Wichterle, I. Lieberam, J. A. Porter, T. M. Jessell, Directed differentiation of embryonic stem cells into motor neurons. *Cell* **110**, 385–397 (2002). doi: 10.1016/S0092-8674(02)00835-8; pmid: 12176325
124. H. Wichterle, M. Peljto, Differentiation of mouse embryonic stem cells to spinal motor neurons. *Curr. Protoc. Stem Cell Biol.* **5**, 1H.1.1–1H.1.9 (2008). doi: 10.1002/9780470151808.sc01h01s5; pmid: 18770634
125. S. Picelli *et al*., Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014). doi: 10.1038/nprot.2014.006; pmid: 24385147
126. N. L. Bray, H. Pimentel, P. Melsted, L. Pachter, Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016). doi: 10.1038/nbt.3519; pmid: 27043002
127. S. Pujar *et al*., Consensus coding sequence (CCDS) database: A standardized set of human and mouse protein-coding regions supported by expert curation. *Nucleic Acids Res.* **46**, D221–D228 (2018). doi: 10.1093/nar/gkx1031; pmid: 29126148
128. M. A. Diroma, L. Ciaccia, G. Pesole, E. Picardi, Elucidating the editome: Bioinformatics approaches for RNA editing detection. *Brief. Bioinform.* **20**, 436–447 (2019). doi: 10.1093/bib/bbx129; pmid: 29040360
129. C. R. Lazzarotto *et al*., Defining CRISPR-Cas9 genome-wide nuclease activities with CIRCLE-seq. *Nat. Protoc.* **13**, 2615–2642 (2018). doi: 10.1038/s41596-018-0055-0; pmid: 30341435
130. P. N. Porensky *et al*., A single administration of morpholino antisense oligomer rescues spinal muscular atrophy in mouse. *Hum. Mol. Genet.* **21**, 1625–1638 (2012). doi: 10.1093/hmg/ddr600; pmid: 22186025
131. W. D. Arnold *et al*., Electrophysiological motor unit number estimation (MUNE) measuring compound muscle action potential (CMAP) in mouse hindlimb muscles. *J. Vis. Exp.* **2015**, e52899 (2015). pmid: 26436455

ACKNOWLEDGMENTS

We thank A. Vieira for assistance with editing, M. O'Reilly for assistance with figures, and A. Hsu for assistance with computational analysis. Δ7SMA mESCs were a kind gift from L. L. Rubin. **Funding:** This work was supported by US National Institutes of Health grants U01 AI142756, RM1 HG009490, R01 EB022376, R35 GM118062, and P01 HL053749 (D.R.L.); the Bill and Melinda Gates Foundation (D.R.L.); the Howard Hughes Medical Institute (D.R.L.); a Friedreich's Ataxia Accelerator grant

(D.R.L.); a Netherlands Organization for Scientific Research Rubicon Fellowship (M.A.); US National Institutes of Health K99 Pathway to Independence Award NS119743-01A1 (M.A.); a Helen Hay Whitney Fellowship (G.A.N.); US National Institutes of Health K99 Pathway to Independence Award HL163805 (G.A.N.); National Science Foundation Graduate Research Fellowship (A.R. and M.W.S.); a Howard Hughes Medical Institute Hanna Gray Fellowship (M.F.R.); Harvard Chemical Biology Training Grant T32 GM095450 (K.T.Z.); UMass Chan Internal Funding (G.G.); and US National Institutes of Health grant R01HD060586 (A.H.M.B.).
Author contributions: Conceptualization: M.A. and D.R.L. Methodology: M.A., W.D.A., D.W., and A.H.M.B. Software: A.R. and M.W.S. Validation: M.A. and K.M.K. Formal analysis: M.A., Z.M., G.A.N., A.R., and M.W.S. Investigation: M.A., Z.M., K.M.K., A.D., G.A.N., and A.J.B. Resources: M.F.R., K.T.Z., J.M.L., J.X., and G.G. Writing – original draft: M.A. and D.R.L. Visualization: M.A. and M.W.S. Supervision: M.A. and D.R.L. Project administration: M.A. and D.R.L. Funding acquisition: M.A. and D.R.L. **Competing interests:** M.A. and D.R.L. have filed patent applications on this work (WO2021158999A1). D.R.L. is a consultant and equity owner of Beam Therapeutics, Prime Medicine, Pairwise Plants, Chroma Medicine, and Nvelop Therapeutics, companies that use or deliver genome editing or genome engineering agents. A.H.M.B is a consultant for Novartis. W.D.A. has served as a consultant to NMD Pharma, Genentech, Catalyst Pharmaceuticals, Dyne, Avidity Biosciences, Design Therapeutics, Argenx, and Novartis and has received research funding from Novartis, Avidity Biosciences, NMD Pharma, and Biogen. **Data and materials availability:** The plasmids used in this study are available through AddGene (depositor: David R. Liu). DNA sequencing files can be accessed using the NCBI SRA (SUB# PRJNA871232). All data are available in the main text or the supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse. This article is subject to HHMI's Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the author-accepted manuscript of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adg6518
Supplementary Text
Figs. S1 to S5
Tables S1 to S3
References (*132*–*160*)
MDAR Reproducibility Checklist
Movies S1 to S5

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

Submitted 11 January 2023; accepted 21 March 2023
Published online 30 March 2023
10.1126/science.adg6518

RESEARCH ARTICLE

SLEEP

Brain activity of diving seals reveals short sleep cycles at depth

Jessica M. Kendall-Bar[1,2]*, Terrie M. Williams[2], Ritika Mukherji[3], Daniel A. Lozano[2], Julie K. Pitman[4], Rachel R. Holser[5], Theresa Keates[6], Roxanne S. Beltran[2], Patrick W. Robinson[2], Daniel E. Crocker[7], Taiki Adachi[2], Oleg I. Lyamin[8,9], Alexei L. Vyssotski[10], Daniel P. Costa[2,5]

Sleep is a crucial part of the daily activity patterns of mammals. However, in marine species that spend months or entire lifetimes at sea, the location, timing, and duration of sleep may be constrained. To understand how marine mammals satisfy their daily sleep requirements while at sea, we monitored electroencephalographic activity in wild northern elephant seals (*Mirounga angustirostris*) diving in Monterey Bay, California. Brain-wave patterns showed that seals took short (less than 20 minutes) naps while diving (maximum depth 377 meters; 104 sleeping dives). Linking these patterns to accelerometry and the time-depth profiles of 334 free-ranging seals (514,406 sleeping dives) revealed a North Pacific sleepscape in which seals averaged only 2 hours of sleep per day for 7 months, rivaling the record for the least sleep among all mammals, which is currently held by the African elephant (about 2 hours per day).

Across the animal kingdom, sleep is critical for energy conservation, immune function, memory, and learning (*1*). Disruptions to sleep, including obstructive sleep apnea and shift work, negatively affect human health (*2*, *3*). By comparison, diverse sleeping habits among wild animals reflect adaptations to resolve conflicts between sleeping or feeding while avoiding predation and exhaustion (*4*–*6*). In response to these trade-offs, cows sleep-chew, horses sleep-stand, ostriches sleep-stare, and frigate birds sleep-fly (*7*–*10*).

Marine mammals face unique challenges in obtaining adequate daily sleep (*11*). Most of them feed underwater and breathe at the ocean surface, where predators typically attack (*12*). Activity budgets of mammals at sea reflect the balance between these survival needs, which often push the animals toward physiological extremes such as large body size, prolonged activity, and enhanced oxygen stores (*13*). For example, northern elephant seals (*Mirounga angustirostris*) travel >10,000 km during 7-month-long foraging trips. Seals minimize time at the surface (~2 min between 10- to 30-min dives) to reduce predation risk by killer whales and white sharks while maximizing foraging time (*12*, *14*–*16*). They also feed around the clock on small prey to satisfy the energy requirements associated with their large body size (*17*). Given these ecophysiological demands, a long-standing question has been when, where, and how do seals sleep at sea?

[1]Scripps Institution of Oceanography, University of California San Diego, San Diego, CA, USA. [2]Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA, USA. [3]Department of Neuroscience, University of Oxford, Oxford, UK. [4]Sleep Health MD, Santa Cruz, CA, USA. [5]Institute of Marine Sciences, University of California Santa Cruz, Santa Cruz, CA, USA. [6]Ocean Sciences Department, University of California Santa Cruz, Santa Cruz, CA, USA. [7]Department of Biology, Sonoma State University, Rohnert Park, CA, USA. [8]Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, CA, USA. [9]A.N. Severtsov Institute of Ecology and Evolution, Moscow, Russia. [10]Institute of Neuroinformatics, University of Zurich and Swiss Federal Institute of Technology (ETH), Zurich, Switzerland.
*Corresponding author. Email: jkb@ucsc.edu

A new tool to detect sleep at sea

We developed a new submersible system to record brain activity (electroencephalogram, EEG) and heart rate (electrocardiogram, ECG) concurrently with dive depth and motion of elephant seals at sea [Fig. 1E; (*18*)]. These sensors identified sleep states [rapid eye movement (REM) sleep and slow-wave sleep (SWS); figs. S1 to S3 and table S2], swimming effort (stroke rate), and three-dimensional (3D) diving behavior in freely moving female juvenile seals (n = 13 seals) (*18*). We recorded sleep in a controlled laboratory environment (n = 5 seals) and in the wild (n = 8 seals) at four locations, including on the beach, in shallow water, offshore along the continental shelf (depth <250 m), and in the open ocean (depth >250 m; table S1). EEG recordings allowed us to pair sleep states with diving behavior recorded in time-depth profiles for juveniles over multiple days at sea (104 sleeping dives). We used these sleep signatures to estimate sleep patterns across >3 million dives by 334 free-ranging adult females over prolonged trips at sea (514,406 sleeping dives across 53,581 recording days).

On a typical sleeping dive, seals transitioned from an awake glide into SWS. Although asleep, they could maintain their upright posture for several minutes (Fig. 1 and movie S1). These results underscore the importance of EEG in assessing sleep state (*18*). As seals shifted from SWS to REM sleep, sleep pa[ra]lysis resulted in a loss of postural control. S[eals] turned upside down and drifted downwards in a "sleep spiral." Sleep spirals tightened from a median diameter of 7.5 ± 7.9 m [median ± interquartile range (IQR)] at 71 ± 97 s (median ± IQR) in SWS to 3.3 ± 3.5 m loops at 40 ± 29 s in REM (Fig. 1). Sleep spirals consisted of two to 13 consecutive 360-degree loops at 82 to 377 m depth. On the continental shelf, seals slept motionless on the ocean floor at 64 to 249 m depth.

The predation risk of sleep at sea

Among marine mammals, unihemispheric sleep (SWS in only one hemisphere) allows captive cetaceans and otariids (fur seals and sea lions) to swim and keep one eye open during sleep (*11*, *19*). This suggests that cetaceans and otariids can sleep while monitoring predators (*20*, *21*). Unihemispheric sleep has not been detected in captive true seals (family *Phocidae*) such as elephant seals (*22*). Similarly, our study did not reveal sleep asymmetry between hemispheres (<2-fold difference). This suggests that true seals use an alternative solution to mitigate predation risk. This study experimentally confirms the hypothesis (*22*, *23*) that in the absence of unihemispheric sleep, elephant seals' extreme diving abilities allow sleep deep below the ocean surface, out of view of visual predators.

The sleep paralysis that co-occurs with REM sleep would make seals especially vulnerable to predation (*1*). REM is often minimized for aquatic mammals because the accompanying paralysis can also prevent access to air (*22*). In captive fur seals confined to water, REM is virtually eliminated (*24*). Elephant seals at sea reduce REM sleep, as is seen in captive fur seals and true seals in water (*24*–*27*), but unexpectedly exhibit a large proportion of REM [26.5 ± 5.0% (mean ± SD) in total sleep time overall and 29.1 ± 4.3% of at-sea total sleep time; table S3]. This compares to 11%, 6%, 5%, and 1% in aquatic sleep for captive Caspian seals, harp seals, walruses, and fur seals, respectively [(*22*, *24*–*27*); see the materials and methods).

Our at-sea deployments occurred during late spring, when juvenile seal aggregations attract predators (*14*). While transiting over the continental shelf, juvenile seals alter their swimming behavior to avoid predation (*28*). Unexpectedly, we found that seals slept proportionally more on the continental shelf than in the open ocean (Figs. 2 and 3A). One seal performed up to 36 consecutive sleeping dives on the continental shelf but fewer than five at sea. This suggests that seals can safely sleep at depth despite elevated coastal predation risk.

Finding time to sleep at sea

Without unihemispheric sleep allowing continuous vigilance, seals are vulnerable and

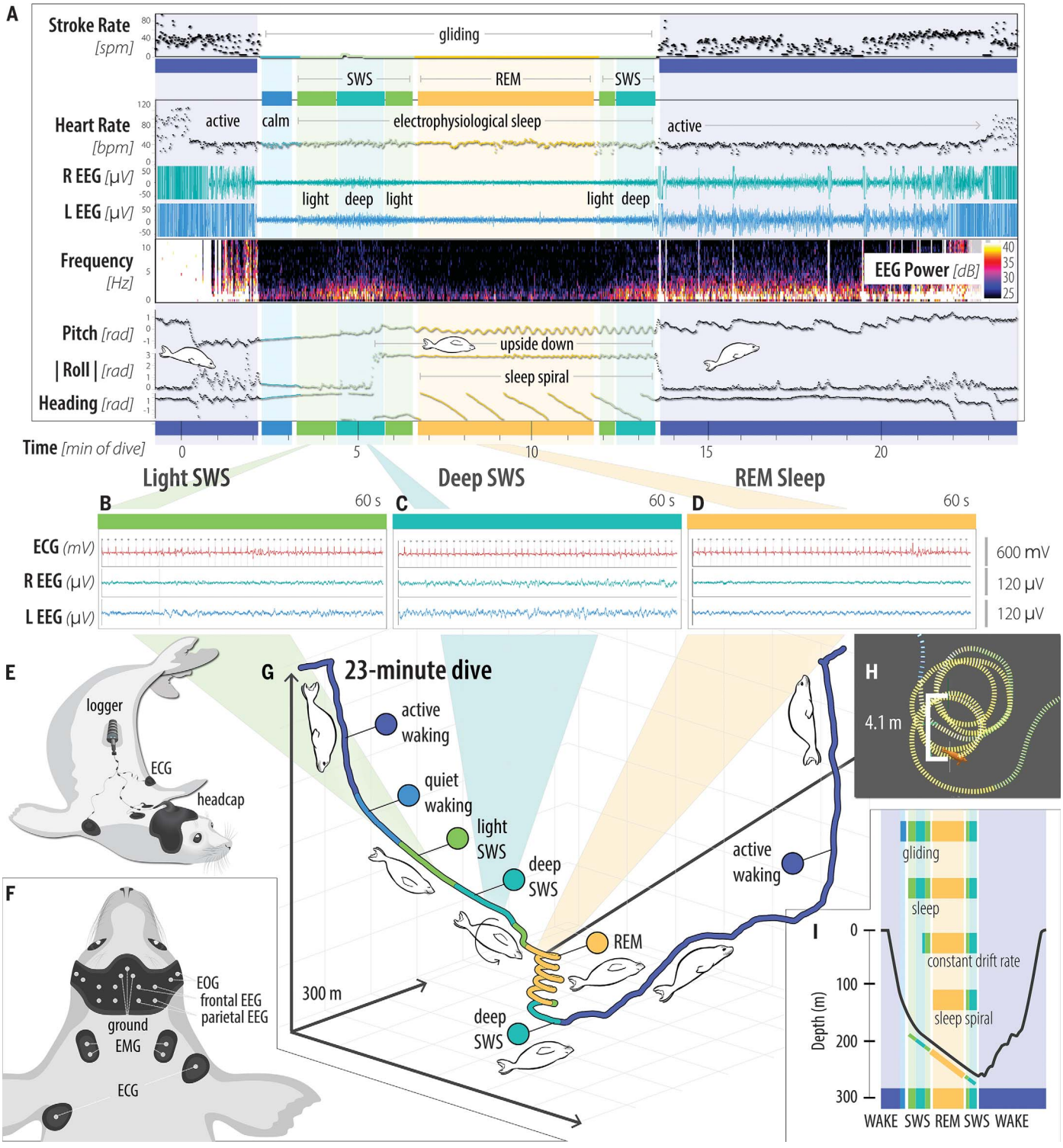


Fig. 1. 3D drift dive sleep patterns. (**A**) A 23-min sleeping dive showing stroke rate [strokes per minute (spm)], heart rate [beats per minute (bpm)], left (L) EEG (μV), right (R) EEG (μV), L EEG spectrogram [power (dB) for frequency (Hz) over time], pitch (radians), roll (absolute value; radians) and heading (radians), and time (minutes of dive). (**B** to **D**) Raw EEG and ECG signals during the transition to light SWS (B), deep SWS (C), and REM (D). During SWS, high-voltage, low-frequency slow waves are present. During REM, low-voltage, high-frequency EEG activity co-occurs with heart rate variability. (**E**) EEG logger configuration demonstrating headcap and logger placement. (**F**) Schematic demonstrating placement of electrodes for electrooculogram (EOG), EEG, ECG, and electromyogram (EMG). (**G**) 3D dive profile color-coded by sleep state: Active waking is shown in dark blue, quiet waking in light blue, light SWS in light green, deep SWS in teal, and REM in yellow. (**H**) Top view of sleep spiral. (**I**) Depth over time showing nested durations of gliding, electrophysiological sleep, constant drift rate, and spiraling.

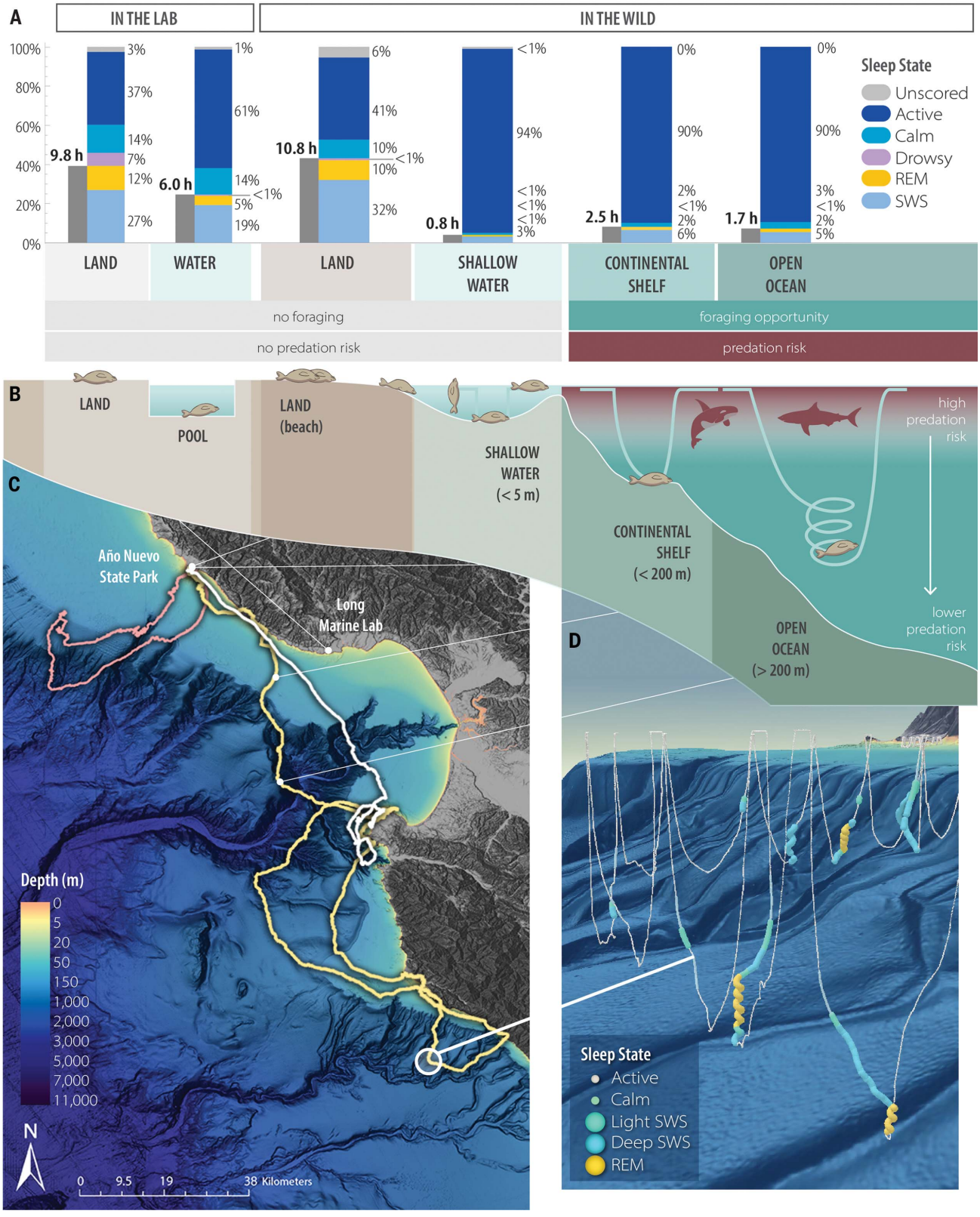


Fig. 2. Sleep patterns from land to sea. (**A**) Daily sleep quotas for seals in the laboratory (on land and in the pool) and in the wild (on land, in shallow water, on the continental shelf, and in the open ocean), including active waking (dark blue), calm (lighter blue), drowsiness (purple), REM sleep (yellow), and SWS (light blue). REM sleep totals include certain and putative REM (see "REM scoring" section in the supplementary materials). (**B**) Schematic showing the resting postures of seals in each habitat, including seals resting on the ocean floor on the continental shelf and drifting in the open ocean. (**C**) 2D map with bathymetry showing georeferenced dead-reckoned tracks for three animals recorded at sea. (**D**) 3D map demonstrating sleeping dive sequence, including the sleeping dive from Fig. 1.

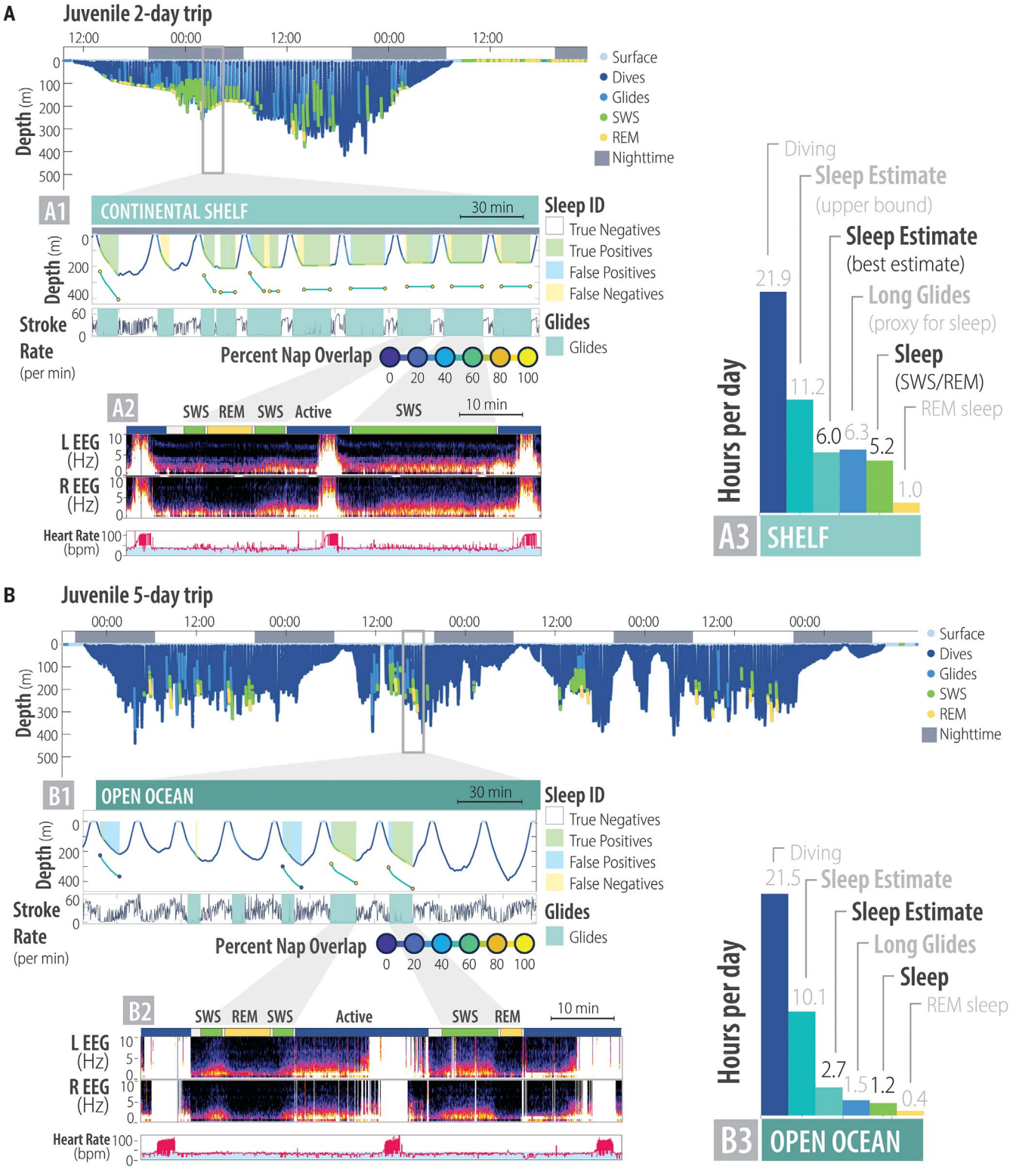


Fig. 3. Sleep identification model performance. Time-depth records for two juvenile seals (**A** and **B**) are colored to indicate surface intervals (light blue), dives (dark blue), glides (blue), SWS (green), and REM sleep (yellow). In panels (A1) and (B1), the identified sleep segments are denoted below the dive profile, where outlined dots at the beginning and end of sleep segments are colored from yellow to dark blue according to overlap with sleep ("percent nap overlap"). Light shaded regions above the dive profile in panels (A1) and (B1) demonstrate sleep identification accuracy (false positives in blue, false negatives in yellow, and true positives in green). Panels (A2) and (B2) display EEG spectrograms and heart rate for two adjacent sleeping dives. Panels (A3) and (B3) quantify daily activity budgets (or provide estimates) in hours per day of diving, sleep estimates (upper bound – unfiltered sleep ID; best estimate/lower bound includes only sleep ID segments that meet filter criteria), gliding (long glides >200 s), sleeping (both SWS and REM), and REM sleep.

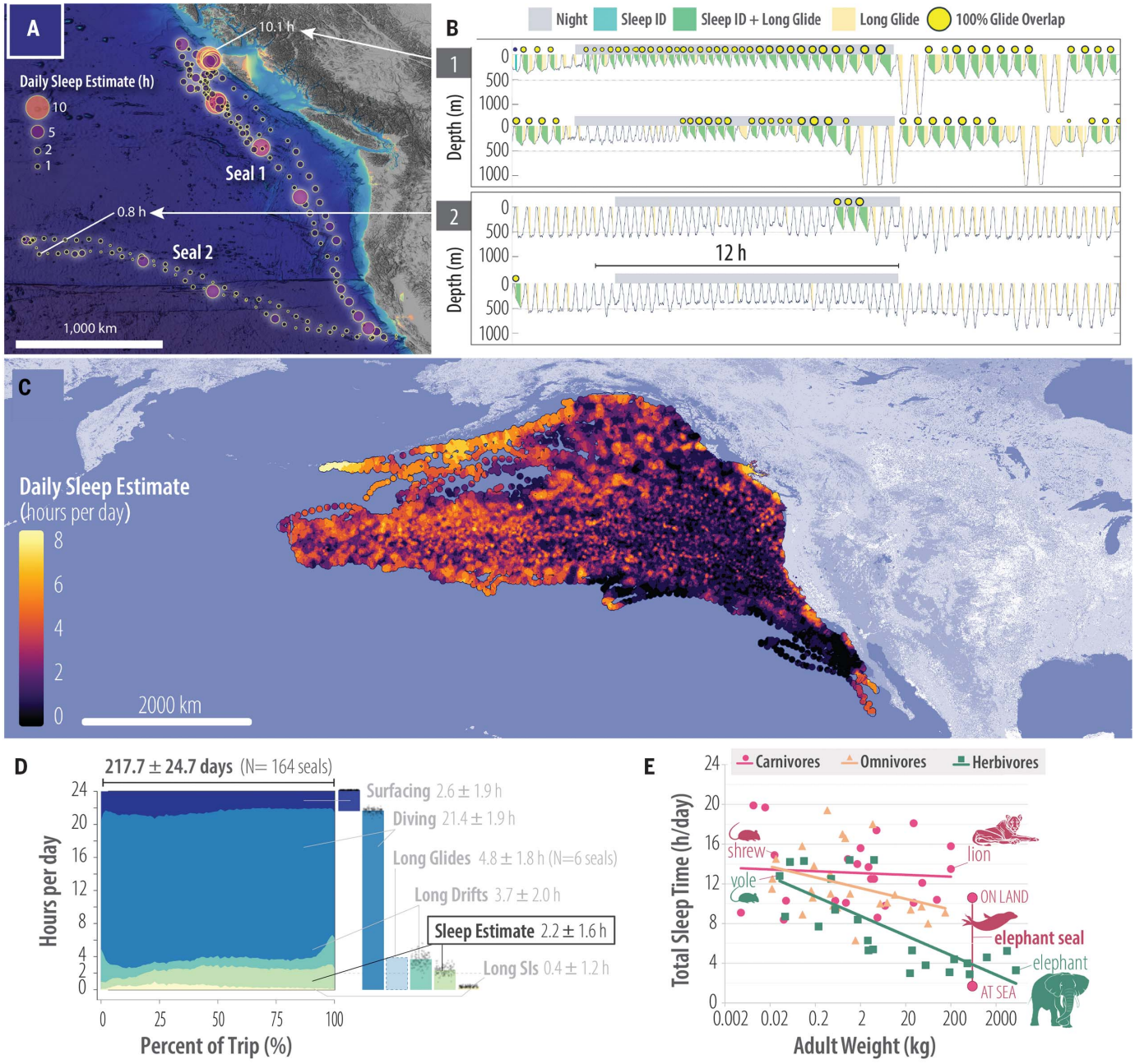


Fig. 4. Estimating daily sleep for 334 adult females. (**A**) Map showing estimated sleep (hours per day) for two of 15 seals instrumented with stroke rate loggers (basemap matches bathymetry legend in Fig. 2). Each circle represents 1 day, with circle size and color reflecting daily sleep (hours per day). (**B1** and **B2**) Time-depth records for two 24-hour days far from shore (B1) and close to Vancouver Island (B2), demonstrating the difference in sleep during pelagic versus coastal foraging. Yellow dots and green shaded regions above dive profiles indicate identified sleep segments that 100% overlap with a long glide segment (yellow and blue shading represent false negatives and false positives, respectively). (**C**) Map with spatially averaged daily sleep estimates clipped to the extent of tracking data (1 point per seal day) across 342 good-quality tracks from 267 adult females. Note the higher sleep time along the coast and foraging grounds. (**D**) Daily activity budgets for long postmolt foraging trips (n = 164 seals) including surface intervals, diving, long glides, long drifts, sleep estimates (filtered long drifts), and long surface intervals. Sleep estimates demonstrate low sleep time throughout the trip. (**E**) Comparative figure showing total sleep time in terrestrial mammalian carnivores, omnivores, and herbivores [reprinted with permission (*1*)]. Extremes of sleep time on land and at sea from EEG recordings in juveniles (Fig. 2) are plotted for comparison. Sleep durations for other mammals are based on behavior and/or EEG in the laboratory and/or wild. Differences in recording location (laboratory versus wild) and sleep identification technique (EEG versus behavior) complicate sleep quantification and direct comparison.

unable to actively transit during sleep to maximize foraging efficiency. Between heightened predation risk and lost foraging opportunities, we expected sleep to be strongly restricted at sea. Supporting this hypothesis, we discovered that seals' daily sleep time was >5 times higher on land than at sea (Fig. 2). Seals slept up to 14 hours/day on land [10.8 ± 3.0 hours/day (mean ± SD)] but as little as 0 hours/day at sea (1.7 ± 0.7 hours/day; tables S3 and S4). After returning from 2 to 3 days at sea, seals remained on land for 18 to 43 hours, sleeping up to 53.3% of each hour before returning to shallow water (fig. S4). This moderate sleep rebound was comparable to the daily patterns

of other seals at the colony (fig. S4), suggesting that this relatively short multiday trip did not incur a notable sleep debt.

Substantial fluctuations in sleep duration allow birds to prioritize migration and breeding for several days (*5*, *10*). As mammals, elephant seals similarly partition strategies over long time scales, sacrificing sleep at sea to support the energy requirements associated with their large size and deferring sleep until they leave the water and are on land with no predators. Seals in laboratory settings show modest differences between total sleep time on land and water (*24*, *25*). Here, we demonstrate greater extremes in total sleep time for wild seals that are necessary to balance sleep with the need to replenish the energy stores of a large, highly mobile predator at sea.

Mapping range-wide sleep patterns at the population level

Using the paired electrophysiological and time-depth signatures of SWS and REM sleep from instrumented seals (Figs. 1 and 2), we developed a high-accuracy sleep identification algorithm that identified segments of inactivity characterized by low vertical speed and acceleration from time-depth data (93% accuracy; Fig. 3 and figs. S5 to S7). This algorithm allowed us to estimate sleep quotas for 334 adult seals from their diving data recorded over several months at sea [n = 170 short trips (74.6 ± 9.5 days) and n = 164 long trips (217.7 ± 24.7 days)] (Fig. 4). These analyses indicated that daily sleep quotas were likely to be universally low (1.1 ± 1.1 hours/day for short 70-day trips and 2.2 ± 1.6 hours/day for >200-day trips) (Fig. 4).

Expanding this analysis to the population level, we can map range-wide sleep patterns to identify critical habitats for protecting wild seals while they sleep at sea. These "sleepscapes," which are based on 342 foraging trips by seals across the North Pacific (Fig. 4C), reveal the same unexpected sleep patterns as in juvenile EEG records. That is, seals slept more while closer to the coastline despite greater predation risk (Fig. 4B1) (*14*). Because coastal foragers consume fewer, larger prey (*17*), our findings suggest that these seals must either expend more energy hunting for larger prey or require more time to rest and process such prey. Although the coastal water column may harbor more predators, the continental shelf may also facilitate sleep by providing shelter from predators and relative proximity to the surface. These findings and the resulting sleepscape aid in identifying critical habitats that may guide coastal conservation efforts for wild animals.

By connecting locomotion with different forms of sleep (SWS versus REM) in northern elephant seals, the present study provides conclusive evidence of sleep during drift dives (*23*, *29*, *30*). Furthermore, these unique recordings of brain activity for a wild, free-ranging marine mammal at sea show that sleeping at depth allows seals to drift safely in and out of sleep paralysis. However, these respites are short, because the large body size [456 to 687 kg (min-max adult female arrival breeding mass); (*31*)] of this elite diver that forages and sleeps in the dark must be sustained by near-constant foraging at sea. Sleep patterns interpreted from the dive records of hundreds of seals revealed only 2 hours/day of sleep for months, rivaling the record for the least sleep among mammals [the African elephant at 2 hours per day; (*32*)]. Both this method (applying sleep signatures from a small sample to reveal population-level patterns) and these findings (a detailed understanding of sleep for a highly mobile, large mammal) provide opportunities for understanding sleep's function, evolution, and pathology across mammals, including in humans.

REFERENCES AND NOTES

1. J. M. Siegel, *Nature* **437**, 1264–1271 (2005).
2. C. V. Senaratna *et al*., *Sleep Med. Rev.* **34**, 70–81 (2017).
3. C. M. Almeida, A. Malheiro, *Sleep Sci.* **9**, 164–168 (2016).
4. C. A. Wyse, A. N. Coogan, C. Selman, D. G. Hazlerigg, J. R. Speakman, *Biol. Lett.* **6**, 696–698 (2010).
5. J. A. Lesku *et al*., *Science* **337**, 1654–1658 (2012).
6. N. C. Rattenborg, S. L. Lima, C. J. Amlaner, *Nature* **397**, 397–398 (1999).
7. E. Ternman, L. Hänninen, M. Pastell, S. Agenäs, P. P. Nielsen, *Appl. Anim. Behav. Sci.* **140**, 25–32 (2012).
8. T. Belling, *Equine Practice* **12**, 2–26 (1990).
9. J. A. Lesku *et al*., *PLOS ONE* **6**, e23203 (2011).
10. N. C. Rattenborg *et al*., *Nat. Commun.* **7**, 12468 (2016).
11. O. I. Lyamin, P. R. Manger, S. H. Ridgway, L. M. Mukhametov, J. M. Siegel, *Neurosci. Biobehav. Rev.* **32**, 1451–1484 (2008).
12. R. A. Martin, N. Hammerschlag, *Mar. Biol. Res.* **8**, 90–94 (2012).
13. M. J. Weise, D. P. Costa, *J. Exp. Biol.* **210**, 278–289 (2007).
14. S. J. Jorgensen *et al*., *Sci. Rep.* **9**, 6153 (2019).
15. P. W. Robinson *et al*., *PLOS ONE* **7**, e36728 (2012).
16. N. Hammerschlag, R. A. Martin, C. Fallows, *Environ. Biol. Fishes* **76**, 341–350 (2006).
17. T. Adachi *et al*., *Sci. Adv.* **7**, eabg3628 (2021).
18. J. M. Kendall-Bar *et al*., *Anim. Biotelem.* **10**, 16 (2022).
19. N. C. Rattenborg, C. J. Amlaner, S. L. Lima, *Neurosci. Biobehav. Rev.* **24**, 817–842 (2000).
20. J. M. Kendall-Bar, A. L. Vyssotski, L. M. Mukhametov, J. M. Siegel, O. I. Lyamin, *PLOS ONE* **14**, e0217025 (2019).
21. O. I. Lyamin *et al*., *Behav. Brain Res.* **129**, 125–129 (2002).
22. O. I. Lyamin, J. M. Siegel, *Handb. Behav. Neurosci.* **30**, 375–393 (2019).
23. Y. Mitani *et al*., *Biol. Lett.* **6**, 163–166 (2010).
24. O. I. Lyamin *et al*., *Curr. Biol.* **28**, 2000–2005.e2 (2018).
25. O. I. Lyamin, *J. Sleep Res.* **2**, 170–174 (1993).
26. O. I. Lyamin *et al*., *Dokl. Biol. Sci.* **444**, 188–191 (2012).
27. O. I. Lyamin, P. O. Kosenko, J. L. Lapierre, L. M. Mukhametov, J. M. Siegel, *J. Neurosci.* **28**, 12614–12621 (2008).
28. B. J. Le Boeuf, D. E. Crocker, in *Great White Sharks, The Biology of* Carcharodon carcharias, A. P. Klimley and D. G. Ainley, Eds. (Academic Press, 1996), pp. 193–205.
29. D. E. Crocker, B. J. L. Boeuf, D. P. Costa, *Can. J. Zool.* **75**, 27–39 (1997).
30. Y. Naito *et al*., *Funct. Ecol.* **27**, 710–717 (2013).
31. R. R. Holser, D. E. Crocker, P. W. Robinson, R. Condit, D. P. Costa, *Proc. Biol. Sci.* **288**, 20211258 (2021).
32. N. Gravett *et al*., *PLOS ONE* **12**, e0171903 (2017).
33. Data for: J. M. Kendall-Bar *et al*., Brain activity of diving seals reveals short sleep cycles at depth, Dryad (2023); https://doi.org/10.7291/D1ZT2B.
34. Code for: J. M. Kendall-Bar *et al*., Brain activity of diving seals reveals short sleep cycles at depth, Zenodo (2023); https://doi.org/10.5281/zenodo.7702650.

ACKNOWLEDGMENTS

We acknowledge the students, volunteers, and researchers contributing to the long-term Año Nuevo elephant seal research program, especially C. Kuhn, J. Hassrick, S. Simmons, M. Fowler, S. Peterson, A. Favilla, S. Kienle, and L. Hückstädt for their assistance in collecting female tracking data. We also thank volunteers C. Lopez and J. Nichols for their help with data collection and analysis; Año Nuevo State Park and the Año Nuevo UC Natural Reserve for their ongoing support, TMMC veterinarians C. Field and S. Johnson for assistance with pilot studies; and Long Marine Lab staff and volunteers for facilitating lab-based studies. Engineers and technicians P. Guerrero, E. Slattery, J. Bielke, and colleagues at Ocean Innovations, Scripps Institution of Oceanography, and the Shorter laboratory at the University of Michigan assisted with the development of the tag housing. **Funding:** This work was supported by the National Ocean Partnership Program (grant N00014-02-1-1012 to D.P.C.); the National Science Foundation (grant N1656282 to D.P.C.); the Strategic Environmental Research and Development Program (SERDP grant RC20-C2-1284 to D.P.C.); the Office of Naval Research (grants N00014-18-1-2822, N00014-00-1-0880, N00014-03-1-0651, and N00014-08-1-1195 to D.P.C. and grant N00014-20-1-2762 to T.M.W.); the Office of Naval Research Defense University Research Instrumentation Program (grant N00014-19-1-2178 to T.M.W. and J.M.K.-B.); the E&P Sound and Marine Life Joint Industry Project (JIP) of the International Association of Oil and Gas Producers (IOGP grant JIP2207-23 to D.P.C.); a National Geographic Early Career Grant (J.M.K.-B.); a Steve & Rebecca Sooy Graduate Research Fellowship (J.M.K.-B.); Achievement Rewards for College Scientists (J.M.K.-B.); the National Science Foundation Graduate Research Fellowship Program (J.M.K.-B.); and a Special Research Grant from the Committee on Research at UC Santa Cruz (D.P.C. and J.M.K.-B.). **Author contributions:** Conceptualization: J.M.K.-B., T.M.W., D.P.C.; Funding acquisition: J.M.K.-B., T.M.W., D.P.C.; Investigation: J.M.K.-B., R.M., D.A.L., J.K.P., R.R.H., T.K., R.S.B., P.W.R., D.E.C., T.A.; Methodology: J.M.K.-B., A.L.V.; Supervision: T.M.W., D.P.C.; Visualization: J.M.K.-B.; Writing – original draft: J.M.K.-B., T.M.W., D.P.C.; Writing – review and editing: J.M.K.-B., T.M.W., R.M., D.A.L., J.K.P., R.R.H., T.K., R.S.B., P.W.R., D.E.C., T.A., O.I.L., A.L.V., D.P.C. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Statistical data are presented primarily in the main text and online supplementary materials, with electronic versions of original sleep data and analysis scripts directly available online through Dryad (*33*) and Zenodo (*34*), respectively.

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

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adf0566
Materials and Methods
Figs. S1 to S7
Tables S1 to S5
References (*35*–*60*)
Movie S1
MDAR Reproducibility Checklist

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

Submitted 30 September 2022; accepted 14 March 2023
10.1126/science.adf0566

PROTEIN DESIGN

Top-down design of protein architectures with reinforcement learning

Isaac D. Lutz[1,2,3]†, Shunzhi Wang[1,2]†*, Christoffer Norn[1,2,4]†, Alexis Courbet[1,2,5], Andrew J. Borst[1,2], Yan Ting Zhao[1,6,7], Annie Dosey[1,2], Longxing Cao[1,2,8], Jinwei Xu[1,2], Elizabeth M. Leaf[1,2], Catherine Treichel[1,2], Patrisia Litvicov[1,6], Zhe Li[1,2], Alexander D. Goodson[1,2], Paula Rivera-Sánchez[4], Ana-Maria Bratovianu[4], Minkyung Baek[1,2,9], Neil P. King[1,2], Hannele Ruohola-Baker[1,3,6,7], David Baker[1,2,3]*

As a result of evolutionary selection, the subunits of naturally occurring protein assemblies often fit together with substantial shape complementarity to generate architectures optimal for function in a manner not achievable by current design approaches. We describe a "top-down" reinforcement learning–based design approach that solves this problem using Monte Carlo tree search to sample protein conformers in the context of an overall architecture and specified functional constraints. Cryo–electron microscopy structures of the designed disk-shaped nanopores and ultracompact icosahedra are very close to the computational models. The icosohedra enable very-high-density display of immunogens and signaling molecules, which potentiates vaccine response and angiogenesis induction. Our approach enables the top-down design of complex protein nanomaterials with desired system properties and demonstrates the power of reinforcement learning in protein design.

Multisubunit protein assemblies play critical roles in biology and are the result of evolutionary selection for function of the entire assembly. Therefore, the subunits in structures such as icosahedral viral capsids often fit together almost perfectly (*1*, *2*). In contrast to direct evolutionary selection on overall system properties, de novo protein design has generated protein architectures using a "bottom-up" hierarchical approach (Fig. 1A, left) in which monomeric structures are first docked into symmetric oligomers (*3*–*6*) and then assembled into closed assemblies with tetrahedral, octahedral, or icosahedral symmetry (*7*–*14*) or open assemblies such as two-dimensional (2D) layers and 3D crystals (*15*–*19*). An advantage of this hierarchical approach is that the multiple interfaces that stabilize the assembly can be validated independently (the first by characterization of the symmetric oligomer and the second by characterization of the nanomaterial assembly from the preformed oligomer), considerably increasing the robustness of the overall design process. Although such designed assemblies are already proving useful for biomedicine in immunobiology and other areas, as highlighted by the recent approval of a de novo–designed COVID vaccine (*20*–*23*), the bottom-up approach does have limitations. The properties of the assembly are limited to what can be generated from the available oligomeric building blocks, at least one of the subunit-subunit interfaces must be strong enough to stabilize a cyclic oligomeric substructure in isolation, and, more generally, there is no way to directly optimize the properties of the overall assembly.

We sought to overcome the limitations of bottom-up protein complex design by developing a top-down approach (Fig. 1A, right) that starts from a specification of the desired properties (overall symmetry, porosity, etc.) of the structure and systematically builds up subunits that pack together to optimize these properties. We reasoned that protein fragment assembly (*24*–*28*), which can generate a wide variety of monomeric protein structures, could provide a suitable mechanism for generating diversity. Previous design approaches such as SEWING have built up proteins from fragments, optimizing for monomer stability at each step (*29*), but we aimed instead to optimize for overall system properties, which could involve trading off monomer stability for increased subunit-subunit interaction strength and other properties. To enable such end state–based optimization, we turned to reinforcement learning (RL), which has achieved considerable success recently in different fields of artificial intelligence, such as self-driving cars (*30*), the AlphaGo program that defeats top human players in the game of Go (*31*, *32*), and algorithm development (*33*). Monte Carlo tree search (MCTS) (*34*, *35*) is an RL algorithm that finds optimal series of choices within a search tree. In M[illegible] choices are selected randomly at each br[illegible] point to find a path down the tree, and after exploring a path, the state is evaluated, and probabilities at each branch point back-propagated up the tree are reweighted accordingly such that subsequent iterations are more likely to lead to optimal paths.

Backbone sampling by MCTS

We sought to develop a MCTS algorithm for generating protein complexes that builds up the monomeric subunits from protein fragments directly optimizing for prespecified global structural properties. We set up the tree search such that at each step in the tree, a short protein fragment is appended at either the N terminus or C terminus of the growing chain. The number of fragments to consider at each step is a trade-off between the rapidity of learning (with a smaller number, weights on each choice can be learned more quickly) and the total diversity of structures that can be generated (which increases with the number of choices at each step). We chose to balance these factors by using as building blocks parametrically generated straight helices, which are fully described by a single parameter (the length, which we allow to vary from nine to 22 residues), followed by short loops clustered into 316 bins (derived from clustering loops in a large helical protein database; see the materials and methods). The search begins with the selection of one of the helix possibilities and then alternates between the addition of a loop or a helix choice at either terminus. Once a loop bin is chosen, we select randomly from the closely related loop backbones within the cluster (Fig. 1B, left). Although this is a far narrower set of local structures than observed in native protein structures, we found in preliminary explorations that a wide variety of compact protein shapes could be readily generated from such building blocks. Building up a 100-residue protein backbone with this approach requires about five helix and four loop additions, yielding a total number of possibilities of ~1×10^{17}, with additional structural diversity from the variation in loop backbones within a bin. The size of the search tree grows exponentially with the number of structural elements, so the space of possibilities is more effectively explored for monomers with fewer helices than for larger monomers.

The search is modulated based on the specific problem specification through geometric constraints that are applied at each step in the search tree and score functions that are evaluated only after full structures are completed. Potential moves consisting of helix or loop fragments are selected at each level of the search tree only if they pass geometric constraints that can be evaluated before the

[1]Department of Biochemistry, University of Washington, Seattle, WA, USA. [2]Institute for Protein Design, University of Washington, Seattle, WA, USA. [3]Department of Bioengineering, University of Washington, Seattle, WA, USA. [4]BioInnovation Institute, DK2200 Copenhagen N, Denmark. [5]Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA. [6]Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, WA, USA. [7]Oral Health Sciences, University of Washington, Seattle, WA, USA. [8]Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China. [9]School of Biological Sciences, Seoul National University, Seoul, Republic of Korea.
***Corresponding author. Email: dabaker@uw.edu (D.B.); swang523@uw.edu (S.W.)**
†These authors contributed equally to this work.

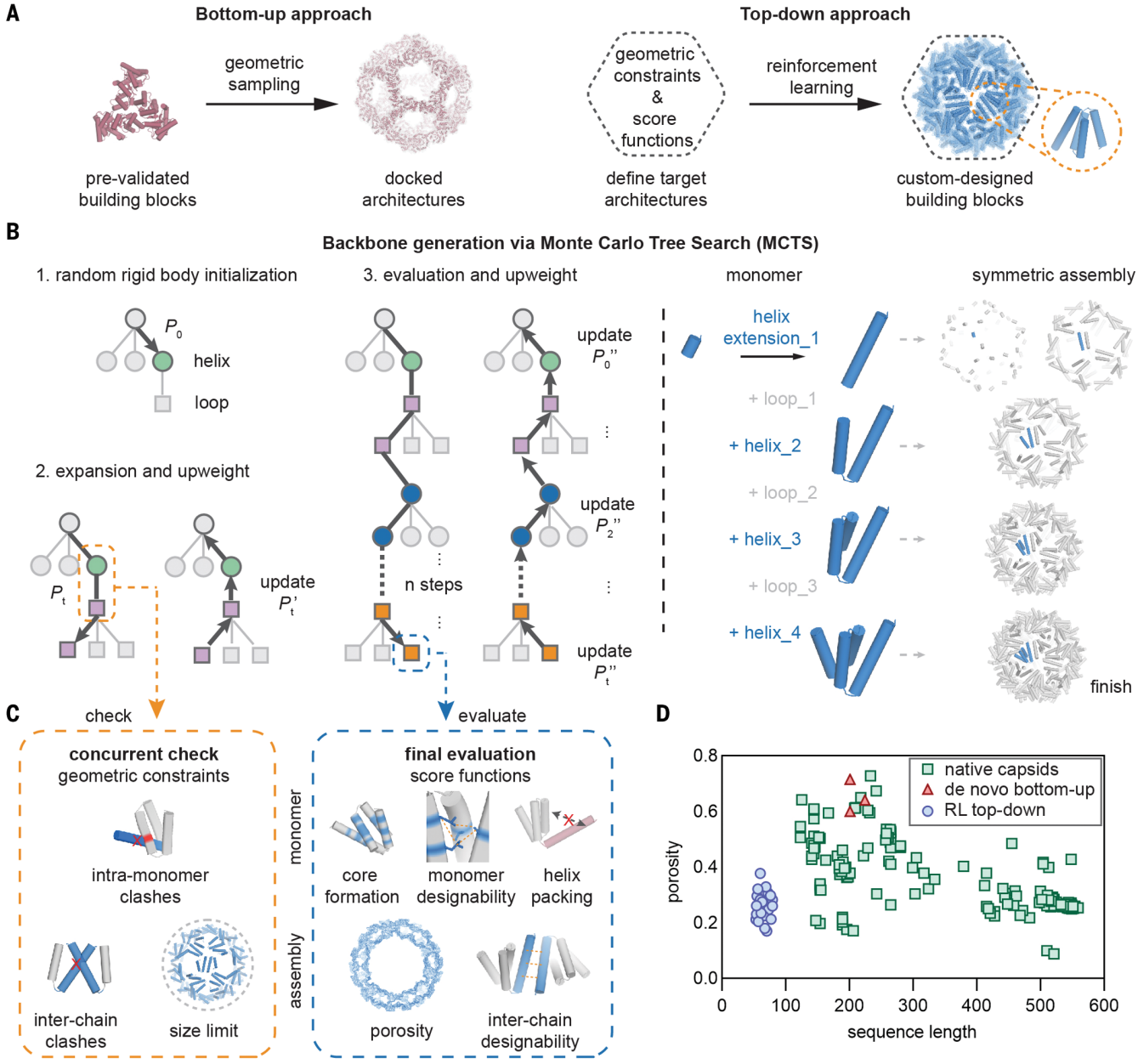


Fig. 1. Top-down design strategy and computational pipeline. (**A**) Bottom-up (left) and top-down (right) strategies to protein assembly design. (**B**) (Left) MCTS architecture for monomer backbone generation. During each simulation, a helix stub is initialized at a random rigid body start position, and different configurations of helices and loops are sampled and constructed sequentially with probability P_t stored in each edge to build the search tree. Each move is checked against a set of predefined geometric constraints during the expansion stage and then updates probabilities P_t' afterward. Upon successful completion of a search tree, the monomer is evaluated by score functions and probabilities P_t'' are back-propagated to update all of the search tree edges. (Right) Symmetric transformations are applied to build an icosahedral capsid in parallel with monomers using the MCTS generative algorithm. (**C**) Concurrent geometric check (left) is performed at every step of the expansion stage and the search tree is terminated if there are violations. Final evaluation (right) with a series of score functions is performed upon completion of a simulation for monomers and assemblies. (**D**) In silico RL-generated capsids (blue) occupy a distinct structural space compared with de novo–designed protein cages (red) and natural capsids (green).

assembly of the entire structure; these include internal clashes and overall shape constraints (see the materials and methods for a full list of geometric constraints). Upon selection of a move passing the geometric constraints, its probability is upweighted, as are the probabilities of all prior moves leading to this point in the search tree. Completed backbones are evaluated using score functions that assess how well the overall generated structure satisfies the user specification of the problem to be solved (Fig. 1C and materials and methods), and the probabilities of selection of each move at each step along the search tree are reweighted accordingly. As individual move weights become increasingly biased after many traversals through the search tree, the generated complete backbones have higher and higher scores (fig. S1). Because each iteration takes on average only tens of milliseconds, high-scoring backbones can be sampled

at scale by searching over tens of thousands of iterations. To address the classical RL problem of balancing exploration with exploitation (*30*–*32*), the search is initialized from many independent trees, and the maximum probability of any one move is capped (see the materials and methods).

We first tested the MCTS approach in silico at the protein monomer level, choosing as a test problem the generation of protein backbones with arbitrarily prespecified overall shapes. To our knowledge, there are no current approaches for addressing this problem. A specified build volume is represented on a grid, and the MCTS is initialized randomly within the volume. At each move, only additions that stay within the specified volume are accepted. For a range of prescribed shapes, including regular polyhedra and letters from the alphabet, the ensembles of generated structures closely fill the specified volumes, and individual backbones have the prespecified shapes (fig. S2). The average sequence length of the solutions increases through the optimization as the choices of moves and combinations of moves that lead to satisfaction of the input constraints are learned, enabling traversal further down the search tree (fig. S1A).

We next sought to generalize the MCTS to the design of symmetric nanomaterials by applying symmetry operators to generate assemblies with the desired symmetry at each step in the search tree. Each move (helix or loop addition) is assessed by considering not only the growing monomer, but also its interactions with all nearby symmetry mates, computed using transformation matrices specifying each symmetry operator; moves that introduce steric clashes are discarded (Fig. 1, B and C). We tested these capabilities in silico by designing cyclic assemblies with symmetries C5 through C12, as well as tetrahedral, octahedral, icosahedral, and quasisymmetric icosahedral assemblies of up to 240 subunits (figs. S3 and S4). We found that by providing different geometric constraints and score functions to guide the search, we could control properties such as shape, size, porosity, and termini position from the top down (figs. S3 to S6).

Nanopore construction using constrained symmetric MCTS

As a first experimental test of the MCTS approach, we applied it to the highly constrained design challenge of filling the space between two previously designed cyclic protein rings (*6*, *36*) to generate disk-shaped structures with a central nanopore (Fig. 2A). Filling this substantial but irregularly shaped space such that there are no large voids between the two rings is not straightforward with previously described protein design methods. We approached this challenge with MCTS by geometrically constraining the search to the space between the two rings, requiring dense packing such that the only large void in the resulting assembly is the pore of the inner C6 ring. Both the inner and the outer ring have C6 symmetry, and the search tree was initialized to start at the N termini of the outer ring and simultaneously build six subunits that collectively fill the empty space. We performed the MCTS for each of 2000 placements of a set of different inner rings with a range of inner pore sizes inside a constant outer ring (for each inner ring, we sampled rotations around and translations along the common cyclic symmetry axis). We selected backbones that fully filled the space between the two rings, designed sequences with ProteinMPNN (*37*), and selected for experimental characterization 32 designs predicted to assemble into the designed assemblies by AlphaFold (AF) (*38*). Of these, we found that 28 were soluble and could be purified and 11 formed particles with the expected size and shape by negative-stain electron microscopy (nsEM). nsEM 3D reconstructions for two designs had an overall shape closely consistent with that of the design models (Fig. 2B; some C7 2D class averages were also obtained; fig. S7). We obtained a cryo–electron microscopy (cryo-EM) map of a third design at 5.1-Å resolution and found it to be closely consistent with the design model: The alpha helices of the model are clearly within the contours of the density (Fig. 2C and fig. S8). The MCTS solution effectively satisfies the design criteria: The space between the two original rings is completely filled in, generating a disk-like structure with a narrow circular pore in the center. We are not aware of any previously designed or naturally occurring proteins that have this overall shape, which could be very useful for downstream nanopore-based sensing applications. More generally, these results demonstrate that the MCTS approach can solve highly constrained protein design problems.

Top-down design of mini-icosahedra

We next explored the use of MCTS to generate icosahedral assemblies by using 59 transformation matrices to compute symmetry mates for a growing monomer. We sought to design very small, closely packed capsids inaccessible by other design methods, and developed geometric constraints and score functions to specifically favor such structures (Fig. 1 and materials and methods). The end state–based score functions include measures of cage porosity and interface designability, as well as external placement of at least one terminus to enable fusion constructs (Fig. 1C). Given a specification of the length and number of helices in the monomer and the size of the overall assembly, we initialized millions of MCTS trajectories starting from a short helical fragment randomly placed within a specified upper distance bound of the origin in a random orientation and performed 10,000 iterations for each to generate a large set of diverse structures. The MCTS generated closely packed icosahedral assemblies in silico, which span a structural space distinct from that of native and previous de novo icosahedra, with shorter sequence lengths than any previously described protein icosahedra and porosities comparable to the densely packed capsids generated by evolution (Fig. 1D).

The MCTS method rapidly generates tens of thousands of candidate icosahedral assemblies, and we experimented with approaches for rapidly designing sequences that stabilize these assemblies in a manner compatible with our overall top-down approach. In previous bottom-up nanocage design studies, the sequences and backbones of the oligomeric building blocks are pre-optimized, so only the new interface formed between the building blocks in the cage is designed, and the overall backbone is kept largely fixed (*11*). By contrast, with the top-down MCTS approach, the entire sequence must be designed, with backbone relaxation to optimize sequence-structure compatibility both within and between the monomers and to increase interface shape complementarity. A deep neural network trained to learn the sequence and structure relationships of native proteins was used to generate amino acid sequence profiles for each position in the newly generated backbones, which were used in turn to bias amino acid selection in the sequence design stage using Rosetta design (materials and methods and figs. S9 to S15). The resulting designs were filtered on the basis of interface contact molecular surface area (*38*), shape complementarity, predicted binding energy, exposed surface hydrophobicity, and AF (*39*) prediction similarity to the design model (see the materials and methods). The rigid body and internal degrees of freedom of the selected icosahedral assemblies were then optimized by Rosetta symmetric relaxation (*40*, *41*), starting from both the Rosetta design model of the assembly and the AF-predicted structure of the monomer mapped back onto the assembly. To further increase sequence-structure compatibility, we repeated this design-predict-relax cycle three times, at each iteration performing sequence design on the full assemblies generated in the previous iteration, mapping back the predicted monomer structures into the assemblies, and relaxing the full structure in Rosetta. We applied this sequence design and backbone refinement procedure to 220,000 of the MCTS-generated backbones and selected 368 designs for experimental characterization (detailed filtering processes are described in the materials and methods and figs. S11 to S13).

Linear gene fragments encoding each design with hexahistidine purification tags were cloned

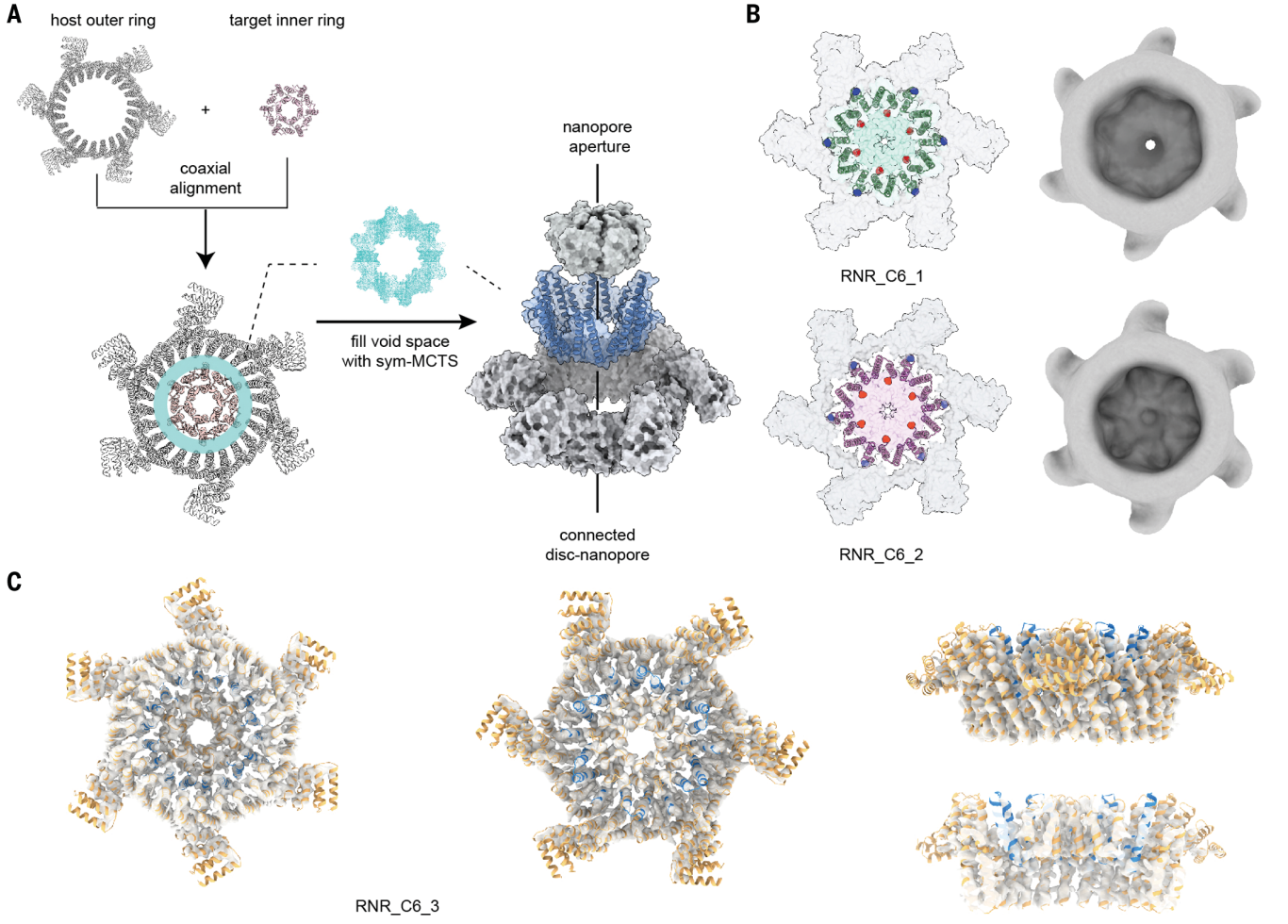


Fig. 2. Disk-nanopore design with symmetric MCTS. (**A**) Schematic illustration of MCTS-based sampling to build space-filling connectors between two concentric rings to generate disk-like structures with different nanopore inner diameters. The inner ring was placed in the center of a host outer ring, varying the rotation and vertical offset, which generates different void volumes (teal; middle panel above arrows). MCTS was then performed to densely fill these void volumes (blue). (**B**) Design models (left column) and nsEM 3D ab initio reconstruction maps (right column) of two connected disk-nanopores (RNR_C6_1 and RNR_C6_2). The symmetric MCTS sampling built helices to connect the inner ring C terminus and outer ring N terminus (highlighted in red and blue, respectively, in the left column). (**C**) The cryo-EM map at 5.1-Å resolution for design RNR_C6_3 viewed from the top, bottom, and side is very close to the design model, with a narrow circular pore in the center of an otherwise nonporous disk-like structure.

into an *Escherichia coli* expression vector, and the proteins produced in *E. coli* in a 96-well format were purified by immobilized metal affinity chromatography (IMAC) pull-down. A total of 208 of the 368 designs were expressed and soluble as assessed by SDS–polyacrylamide gel electrophoresis. To evaluate particle formation, we performed nsEM on the IMAC elution fraction for each soluble sample. Two designs (RC_I_1 and RC_I_2, RL capsid with I symmetry, design 1 and 2) formed uniform particles with the expected size and shape (Fig. 3, A and B). Size-exclusion chromatography (SEC) of both designs yielded single peaks with an apparent molecular weight in the range expected for these assemblies (Fig. 3, C and D). The designed assemblies had the expected alpha-helical circular dichroism (CD) spectra and apparent melting temperatures above 65°C. nsEM analysis showed that assembly morphologies were retained after 1 hour of treatment at 95°C and subsequent cooling to 25°C (Fig. 3, F and H, and fig. S17).

To evaluate the accuracy of our design strategy, we determined the structures of SEC-purified RC_I_1 and RC_I_2 capsid particles using cryo-EM (Fig. 4 and fig. S18). For RC_I_1, 3D reconstruction yielded a 2.5-Å-resolution cryo-EM atomic model that closely matched the computational design (Fig. 4, A and B, and fig. S19). The N-terminal helices of two monomers pack in an antiparallel fashion to form the primarily hydrophobic C2 interface, whereas the two helices near the C terminus form the C5 interface with their neighbors (Fig. 4, B and C). Small apertures (diameter ~13 Å) present at the C3 axes of the capsid make the N termini available for genetic fusion (Fig. 4C). Over the designed monomer, the root mean square deviation (RMSD) between the cryo-EM structure and the design model is 0.76 Å (Fig. 4D); a single rotamer flip (Phe^{63}) and tilting of the C-terminal helix results in a slight expansion of the overall cage diameter, resulting in an RMSD over all 60 subunits of 3.72 Å (Fig. 4E). For RC_I_2, the 2.9-Å cryo-EM structure of design RC_I_2 was even closer to the design model (Fig. 4, F and G, and fig. S20), with RMSDs at the C2 and C5 interfaces of 0.66 and 0.27 Å, respectively (Fig. 4H). The RC_I_2 monomer adopts the designed three-helical bundle fold with a 0.59-Å RMSD to the design model (Fig. 4I), and the overall assembly is almost identical to the design model with a 1.39-Å RMSD over all 60 subunits (Fig. 4J). The C2 interface is situated near the extended C terminus of the

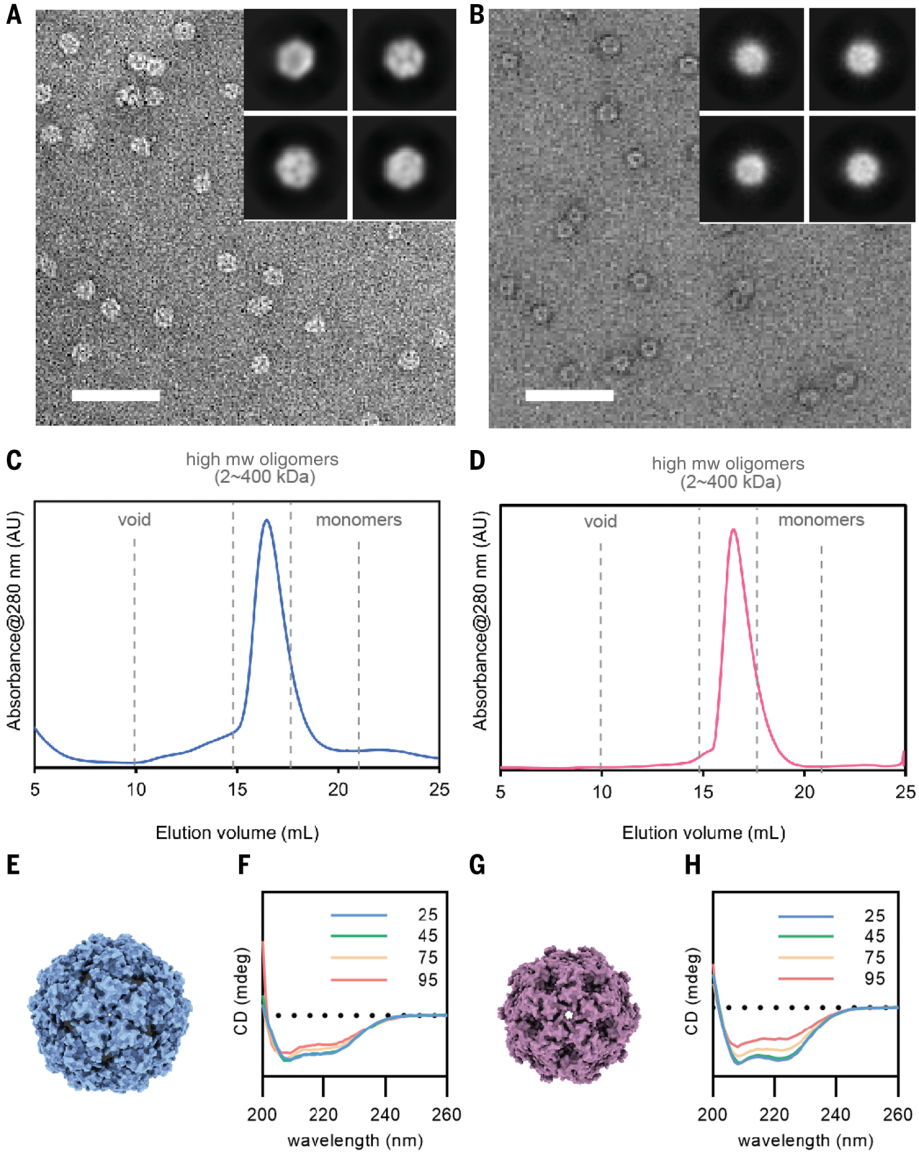


Fig. 3. Experimental characterization of designed capsids RC_I_1 and RC_I_2. (**A** and **B**) Representative nsEM micrographs and reference-free 2D class averages (inset) for RC_I_1 (left) and RC_I_2 (right). Scale bar, 200 nm. (**C** and **D**) A single peak was observed for each SEC elution profile near the expected elution volumes for the target complexes. (**E** and **G**) Capsid computational design models. (**F** and **H**) Circular dichroism spectra measured at different temperatures (°C).

monomer, allowing for potential monomeric or dimeric genetic fusions. The C5 pentameric interface is mediated by interactions between the N-terminal helices, which point inward and enable functionalization of the interior of the capsid. With diameters of 13 and 10 nm for RC_I_1 and RC_I_2, respectively, and associated monomer lengths of 67 and 54 residues, the designed mini-capsids are considerably smaller than most viral capsids.

Applications of top-down–designed capsids

The compact size and corresponding small exterior surface area of the designed particles enables the display of 60 or 120 copies of N- and/or C-terminal fused proteins with exceptionally high density: six or more times higher than previously designed icosahedral cages. We set out to explore whether this higher density could lead to greater biological efficacy in signaling and vaccine applications. We began by exploring the robustness of the designs to substantial sequence changes and to fusion of proteins to their outward-facing termini.

To evaluate robustness to sequence changes, we used ProteinMPNN (*37*) to generate diverse sequences for the RC_I_1 capsid backbone, and the designs were filtered using the AF and Rosetta metrics described above. Two of six experimentally tested ProteinMPNN designs, RC_I_1-H9 and RC_I_1-H11 (the former designed by ProteinMPNN using the working capsid backbone Cα coordinates as input, the latter the idealized polyA backbone without any backbone optimization and relaxation), assembled into the designed I1 symmetric capsid as evidenced by IMAC, SEC, and nsEM. A 3-Å cryo-EM structure of RC_I_1-H11 was almost identical to the design model, with a monomeric RMSD of 0.60 Å (Fig. 5A and fig. S21) and a very low full-cage RMSD over all 60 subunits of only 0.96 Å. RC_I_1-H9 and RC_I_1-H11 have on average 46% sequence divergence from the parent capsid and 30% sequence difference from each other, including highly diverse interface residue selections (fig. S22; for example, the errant Phe^{63} of the parent capsid was redesigned to Glu^{63} in RC_I_1-H11, likely accounting at least in part for the closer agreement of RC_I_1-11 with the design model). These results demonstrate that the RL approach can generate directly designable protein backbone geometries with a high degree of accuracy.

We evaluated the robustness of the designs to genetic fusion by fusing SpyTag, SpyCatcher (*42*), and green fluorescent protein (GFP) proteins to the RC_I_1-H11 capsid with an N-terminal $(GGS)_n$ linker (Fig. 5B and figs. S23 to S25). In all cases, SEC elution profiles and nsEM micrographs showed monodisperse particles of the expected size and shape (see the materials and methods). The 2D class averages (inset) revealed spherical structures similar to that of the original icosahedral capsid, with additional density at the periphery of the particles, consistent with fused proteins connected to scaffolds through a flexible linker. Unlike a larger cage, nuclear localization sequence–tagged capsids fused to GFP are efficiently translocated into the nucleus, opening the door to nuclear delivery of high-valency protein and DNA-organizing constructs (fig. S26).

To assess the efficacy of the designed capsids in activating cellular signaling pathways by clustering cell surface receptors, we fused 60 copies of the angiopoietin 1 (Ang1) F domain (Fd), which binds the Tie2 receptor, to RC_I_1-H11 using SpyTag-SpyCatcher conjugation (*14*, *18*, *43*) (see the materials and methods and fig. S27). We found that the F domain–displaying capsids had very high potency in driving FOXO1 exclusion from the nucleus (Fig. 5, C and D), activating the AKT pathway (Fig. 5D and fig. S28, A to C) and stabilizing nascent blood vessels formed from human umbilical vein endothelial cells (HUVECs; Fig. 5E) (*43*–*49*). The Fd-displaying capsids (0.16 nM RC_I_1-H11-Fd) elicited stronger responses than a 10-fold greater concentration of a much

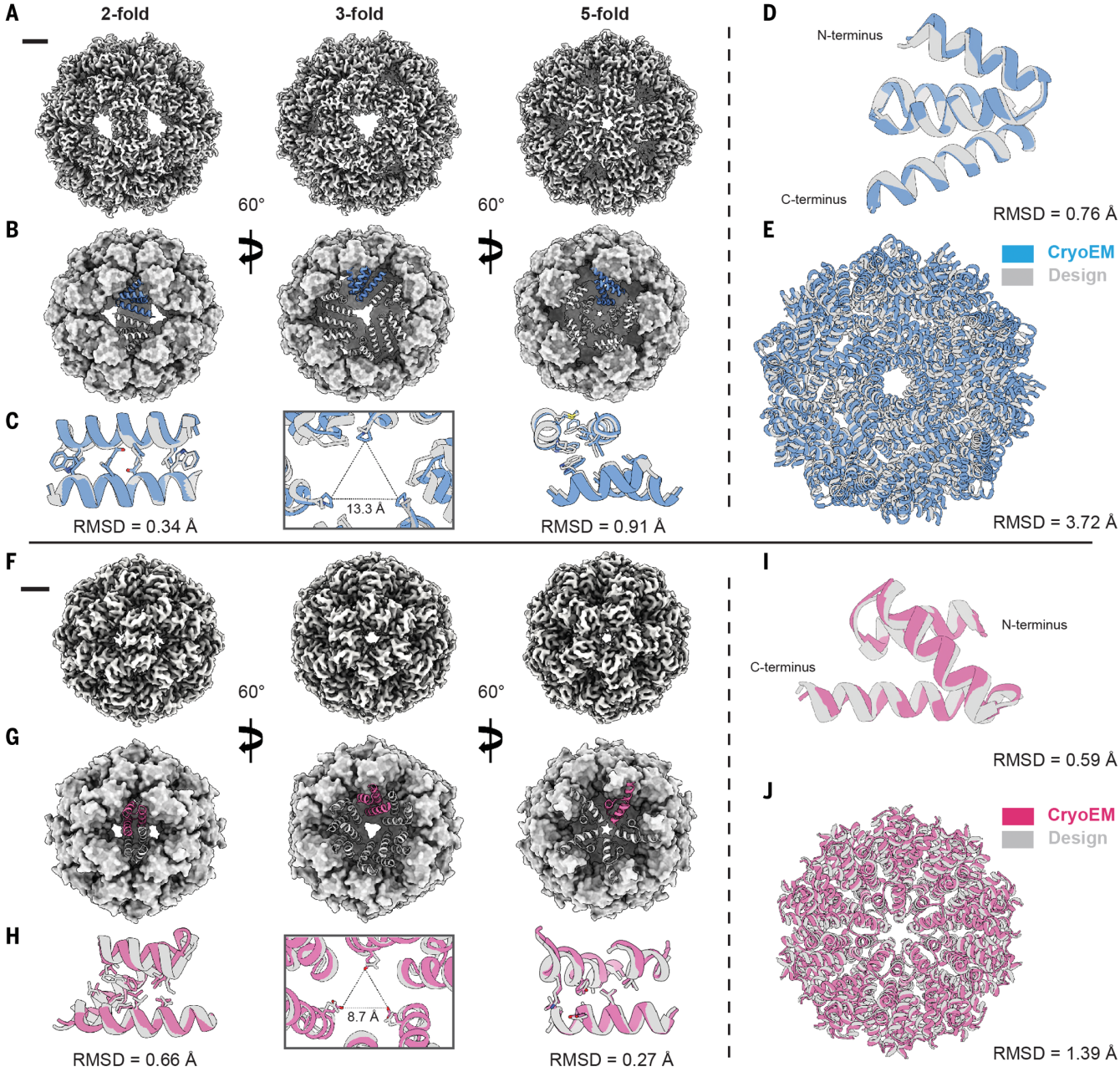


Fig. 4. Near-atomic resolution cryo-EM structures of designed capsids match design models. (**A**) A 2.5-Å cryo-EM reconstruction of RC_I_1 viewed along the three symmetry axes. Scale bar, 20 Å. (**B**) Cryo-EM structure of RC_I_1 highlighting monomer packing and interfaces along each symmetry axis. (**C**) Overlay and RMSD calculations for RC_I_1 compared with the design model for each symmetry interface (cryo-EM is shown in blue; design is shown in gray). (**D**) Overlay and RMSD calculation for a single monomer of RC_I_1. (**E**) Overlay and RMSD calculation for the entire 60-mer RC_I_1 capsid. (**F**) A 2.9-Å cryo-EM reconstruction of RC_I_2 viewed along the three symmetry axes. Scale bar, 20 Å. (**G**) Cryo-EM structure of RC_I_2 highlighting monomer packing and interfaces along each symmetry axis. (**H**) Overlay and RMSD calculations for RC_I_2 compared with the design model for each symmetry interface (cryoEM is shown in pink; design is shown in gray). (**I**) Overlay and RMSD calculation for a single monomer of RC_I_2. (**J**) Overlay and RMSD calculation for the entire RC_I_2 capsid.

larger F-domain–presenting icosahedral nanoparticle (I53-50) (*12*, *43*); the elevated potency likely results from the higher surface display density [to facilitate comparison, concentrations at the bottom of Fig. 5D are in terms of Fd monomer (0.16 nM capsid × 60 Fd copies per capsid = 10 nM Fd)]. The 0.16 nM (10 nM Fd) capsid also elicited stronger responses than 100 nM Ang1. The F domain–displaying capsid is thus an exceptionally potent Tie2-activating ligand. The designed capsid is also far easier to produce and much more stable than Ang1 and thus could be useful in stimulating differentiation and regeneration.

The high surface presentation density enabled by the designed scaffolds provides a route to investigating the effect of packing density on the elicitation of immune responses by nanoparticle-based immunogens. As a first step in this direction, we fused trimeric influenza hemagglutinin (HA) to the N terminus of I1-capsid RC_I_1 using a $(GS)_6$ linker. The fusion protein was expressed and secreted from mammalian cells and clearly forms

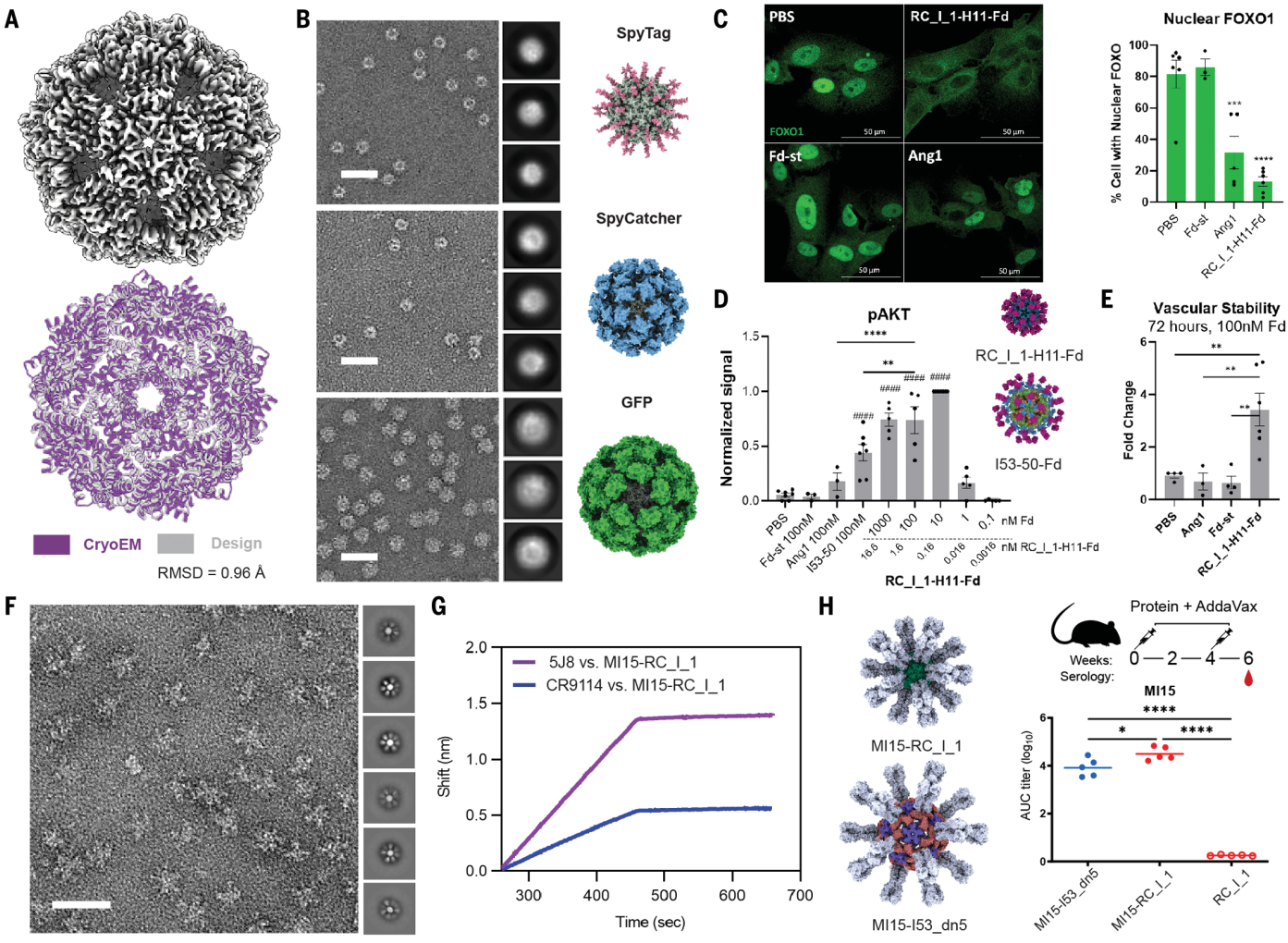


Fig. 5. Applications of designed capsids. (**A**) Robustness of RC_I_1 capsid to sequence redesign using ProteinMPNN. The 3-Å-resolution cryo-EM reconstruction of RC_I_1-H11 (top) reveals a close agreement between the experimental structure (purple) and the design model (gray) with a RMSD of 0.96 Å. The RC_I_1-H11 structure is nearly identical to RC_I_1 despite considerable sequence differences [bottom; residue differences are highlighted in red (RC_I_1-H11) and teal (RC_I_1)]. (**B**) From top to bottom, models and representative nsEM images of spyTag-, spyCatcher-, and GFP-fused (to N terminus) RC_I_1-H11 with 2D class averages. Scale bar, 50 nm. (**C** and **D**) RC_I_1-H11-Fd activates Tie2 downstream Akt phosphorylation and FOXO1 translocation. Serum-starved HUVECs were treated with serially diluted RC_I_1-H11-Fd (1000-0.1 nM), Fd-st (100 nM), Ang1 (100 nM), I53-50 (100 nM), or phosphate-buffered saline (PBS) control for 15 min before protein lysate collection for Western blot analysis, or cells were fixed for FOXO1 antibody stain. (C) Left, representative confocal images of HUVECs immunofluorescence stained with FOXO1 antibody. Right, quantification showing the percentage of cells with nuclear FOXO1; 100 cells were counted in each biological replicate. Levels of significance were compared with PBS control in the FOXO1 graph. (D) Quantification of Western blot showing pAKT signal normalized to RC_I_1-H11-Fd at 10 nM. RC_I_1-H11-Fd induces a significantly higher signal than the previously characterized I53-50-Fd (inset) at 100 nM Fd equivalent. (**E**) Quantification of vascular stability by averaging the number of nodes, meshes, and tubes calculated at the 72-hour time point using the Angiogenesis Analyzer plug-in in ImageJ (fig. S28D). In (C) to (E), *P* values were calculated using one-way ANOVA with Bonferroni's multiple-comparisons test in Prism for comparing groups of two or more; **P* < 0.05; ***P* < 0.01; ****P* < 0.001; *****P* < 0.0001; significance over PBS control is noted as # in (D). (**F**) Representative nsEM micrograph and 2D class averages (inset) of mammalian cell secreted RC_I_1 particle flexibly fused with M15 influenza HA (MI15-RC_I_1). Scale bar, 50 nm. (**G**) The RC_I_1 displayed HA is antigenically intact, reacting with both head (5J8) and stem (CR9114) anti-HA antibodies in biolayer interferometry experiments. (**H**) Models of RC_I_1 (top) and I53_dn5 (bottom) displaying MI15 influenza HA (left); the presentation is considerably denser in the former. Top right: Mouse immunization schedule. Bottom right: HA-specific antibody titers in immune sera. Statistical significance was determined using one-way ANOVA with Tukey's multiple-comparisons test; **P* < 0.05; *****P* < 0.0001. The RC_I_1 display format produces a higher antibody titer than the I53_dn5 nanoparticle currently in clinical trials.

HA-displaying particles according to SEC and nsEM (fig. S29 and Fig. 5F). Biolayer interferometry showed binding of both 5J8 [anti-HA head antibody (*50*)] and CR9114 [anti-HA stem antibody (*51*)] immunoglobulin G to HA capsids (Fig. 5G), indicating that the HA remains antigenically intact when displayed on the surface of the capsids. We immunized mice with HA-displaying RC_I_1, as well as a much larger icosahedral immunogen, HA-I53_dn5 (*52*), which has previously been shown to elicit protective responses against influenza and is currently being evaluated in clinical trials (*53*). We found that HA-displaying RC_I_1 elicited a strong antibody response against vaccine-matched HA that was greater than that produced by the clinical vaccine candidate by a small but statistically significant

amount (Fig. 5H). These results indicate that the high antigen presentation density enabled by top-down design can yield robust immune responses.

Conclusion

Our top-down RL approach enables the solution of design challenges inaccessible to previous bottom-up design methods. Cryo-EM structures confirm the design of 54- and 67-residue proteins that assemble into 60-subunit icosahedra with both internal monomer and overall assembly structure nearly identical to the computational models, and of disk-shaped nanopores generated by densely filling the space between cyclic protein rings with different diameters. Both the icosahedra and the disk designs are distinct from any previously designed or naturally occurring structures; the former have smaller subunits, smaller radii, and lower porosities, and the latter have narrow central pores within large, circular, otherwise nonporous structures. These structures could not have been built with previous bottom-up approaches. For the icosahedra, generating the shape complementarity of the interfaces requires the context of the full capsid structure, possible only through a top-down approach, and for the disks, densely filling a prescribed volume from preexisting building blocks is generally not possible. The density of protein chains and termini available for fusion to the icosahedra is considerably greater than the most compact previously designed assembly, enabling fusion to functional protein domains to generate bioactive nanoparticles. The Ang1 F domain–displaying capsids are potent activators of angiogenesis, and the influenza HA–displaying capsids elicit strong anti-HA antibody responses in mice. The capability of the MCTS approach to optimize any set of specified geometric criteria in a top-down fashion provides a route to potent, multivalent cellular receptor agonists and vaccines that are custom designed to rigidly scaffold immunogen or receptor-binding monomers and precisely position them relative to one another. More generally, our results demonstrate the power of RL for protein design, which we expect can be increased further by the incorporation of policy and value networks (*30*–*32*, *54*) to further guide the search.

REFERENCES AND NOTES

1. R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, J. Rudnick, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 15556–15560 (2004).
2. T. Douglas, M. Young, *Science* **312**, 873–875 (2006).
3. J. E. Padilla, C. Colovos, T. O. Yeates, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2217–2221 (2001).
4. J. A. Fallas *et al*., *Nat. Chem.* **9**, 353–360 (2017).
5. J. Zhu *et al*., *Chem. Rev.* **121**, 13701–13796 (2021).
6. B. I. M. Wicky *et al*., *Science* **378**, 56–61 (2022).
7. N. P. King *et al*., *Science* **336**, 1171–1174 (2012).
8. Y.-T. Lai, D. Cascio, T. O. Yeates, *Science* **336**, 1129 (2012).
9. Y.-T. Lai *et al*., *Nat. Chem.* **6**, 1065–1071 (2014).
10. N. P. King *et al*., *Nature* **510**, 103–108 (2014).
11. Y. Hsia *et al*., *Nature* **540**, 150 (2016).
12. J. B. Bale *et al*., *Science* **353**, 389–394 (2016).
13. Y. Hsia *et al*., *Nat. Commun.* **12**, 2294 (2021).
14. R. Divine *et al*., *Science* **372**, eabd9994 (2021).
15. C. J. Lanci *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7304–7309 (2012).
16. J. C. Sinclair, K. M. Davies, C. Vénien-Bryan, M. E. M. Noble, *Nat. Nanotechnol.* **6**, 558–562 (2011).
17. S. Gonen, F. DiMaio, T. Gonen, D. Baker, *Science* **348**, 1365–1368 (2015).
18. A. J. Ben-Sasson *et al*., *Nature* **589**, 468–473 (2021).
19. Z. Li *et al*., bioRxiv 2022.11.18.517014 [Preprint] (2022).
20. A. C. Walls *et al*., *Cell* **183**, 1367–1382.e17 (2020).
21. P. S. Arunachalam *et al*., *Nature* **594**, 253–258 (2021).
22. P. S. Arunachalam *et al*., bioRxiv [Preprint] (2022).
23. J. Y. Song *et al*., *EClinicalMedicine* **51**, 101569 (2022).
24. K. T. Simons, R. Bonneau, I. Ruczinski, D. Baker, *Proteins* **37**, 171–176 (1999).
25. D. Gront, D. W. Kulp, R. M. Vernon, C. E. M. Strauss, D. Baker, *PLOS ONE* **6**, e23294 (2011).
26. E. Verschueren *et al*., *Curr. Opin. Struct. Biol.* **21**, 452–459 (2011).
27. N. Koga *et al*., *Nature* **491**, 222–227 (2012).
28. R. Pearce, X. Huang, G. S. Omenn, Y. Zhang, *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2208275120 (2023).
29. T. M. Jacobs *et al*., *Science* **352**, 687–690 (2016).
30. B. R. Kiran *et al*., *IEEE Trans. Intell. Transp. Syst.* **23**, 4909–4926 (2022).
31. V. Mnih *et al*., *Nature* **518**, 529–533 (2015).
32. D. Silver *et al*., *Nature* **529**, 484–489 (2016).
33. A. Fawzi *et al*., *Nature* **610**, 47–53 (2022).
34. R. Coulom, in *Computers and Games*, H. J. van den Herik, P. Ciancarini, H. H. L. M. Donkers, Eds. (Springer, 2007), vol. 4630 of *Lecture Notes in Computer Science*, pp. 72–83; http://link.springer.com/10.1007/978-3-540-75538-8_7.
35. L. Kocsis, C. Szepesvári, in *Machine Learning: ECML 2006*, J. Fürnkranz, T. Scheffer, M. Spiliopoulou, Eds. (Springer, 2006), vol. 4212 of *Lecture Notes in Computer Science*, pp. 282–293; http://link.springer.com/10.1007/11871842_29.
36. J. P. Hallinan *et al*., *Commun. Biol.* **4**, 1240 (2021).
37. J. Dauparas *et al*., *Science* **378**, 49–56 (2022).
38. L. Cao *et al*., *Nature* **605**, 551–560 (2022).
39. J. Jumper *et al*., *Nature* **596**, 583–589 (2021).
40. L. G. Nivón, R. Moretti, D. Baker, *PLOS ONE* **8**, e59004 (2013).
41. P. Conway, M. D. Tyka, F. DiMaio, D. E. Konerding, D. Baker, *Protein Sci.* **23**, 47–55 (2014).
42. B. Zakeri *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **109**, E690–E697 (2012).
43. Y. T. Zhao *et al*., *EMBO Rep.* **22**, e53471 (2021).
44. S. Han *et al*., *Sci. Transl. Med.* **8**, 335ra55 (2016).
45. I. Kim *et al*., *Circ. Res.* **86**, 24–29 (2000).
46. M. Kim *et al*., *J. Clin. Invest.* **126**, 3511–3525 (2016).
47. C. Daly *et al*., *Genes Dev.* **18**, 1060–1071 (2004).
48. K. L. DeCicco-Skinner *et al*., *J. Vis. Exp.* **91**, e51312 (2014).
49. N. P. J. Brindle, P. Saharinen, K. Alitalo, *Circ. Res.* **98**, 1014–1023 (2006).
50. J. C. Krause *et al*., *J. Virol.* **85**, 10905–10908 (2011).
51. C. Dreyfus *et al*., *Science* **337**, 1343–1348 (2012).
52. G. Ueda *et al*., *eLife* **9**, e57659 (2020).
53. S. Boyoglu-Barnum *et al*., *Nature* **592**, 623–628 (2021).
54. D. Silver *et al*., *Nature* **550**, 354–359 (2017).
55. I. Lutz, Protein backbone MCTS, Zenodo, (2023); https://doi.org/10.5281/zenodo.7709840.
56. S. Wang, C. Norn, I. Lutz, 2023 RL capsid design, Zenodo (2023); https://doi.org/10.5281/zenodo.7758067.

ACKNOWLEDGMENTS

We thank F. Dimaio, J. Dauparas, B. Coventry, T. Huddy, N. Woodall, R. Kibler, J. Watson, and I. Haydon for help with computational design and discussion; R. Kibler, D. Feldman, L. Milles, and N. Ennist for help with the experiments; and S. Dickinson and J. Quipse for help in maintaining and operating the electron microscopes used. **Funding:** This work was conducted at the Advanced Light Source (ALS), a national user facility operated by Lawrence Berkeley National Laboratory on behalf of the Department of Energy, Office of Basic Energy Sciences, through the Integrated Diffraction Analysis Technologies (IDAT) program, supported by DOE Office of Biological and Environmental Research. Additional support comes from the National Institutes of Health project ALS-ENABLE (NIH grant P30 GM124169) and a High-End Instrumentation Grant S10OD018483. This work was also supported by the National Institute on Aging grant (grant 1U19AG065156-01 to I.D.L. and D.B.); Amgen (S.W.), the Audacious Project at the Institute for Protein Design (A.J.B., Z.L., L.C., H.R.-B., Y.T.Z., D.B., and N.P.K.); the NIH/National Institute of Dental and Craniofacial Research (NIDCR) (grant T90 DE021984 to Y.T.Z); the National Institute of Allergy and Infectious Diseases (NIAID grant 1P01AI167966 to N.P.K.); Novo Nordisk (foundation grant NNF17OC0030446 to C.N.); the Open Philanthropy Project Universal Flu Vaccine and Improving Rosetta Design (A.D., N.P.K., and D.B.); a Microsoft gift (M.B.); and the Department of Defense Peer Reviewed Medical Research Program (award W81XWH-21-1-0006 to H.R.-B. and D.B.). **Author contributions:** I.D.L. and S.W. contributed equally and the author order was chosen arbitrarily; citations on CVs, etc., will be adjusted accordingly. S.W., I.D.L., C.N., and D.B. conceptualized the research. I.D.L. developed the RL backbone generation method. S.W. and C.N. developed the sequence design pipeline. L.C. and M.B. developed the DL sequence profile prediction method. S.W. designed the original capsids and performed the screening, expression, and characterization experiments. A.C. and J.X. designed the nanopore ring connector and performed cryo-EM characterization. A.J.B. designed cryo-EM experiments and optimization of sample purification conditions, performed the initial cryo-EM screening experiments, and optimized the cryo-EM freezing conditions. A.J.B., S.W., and Z.L. prepared additional cryo-EM grids and collected cryo-EM data. A.J.B. processed the cryo-EM data and built and solved the structures for each designed capsid. S.W. and A.D. designed and characterized the fusion capsids. A.D., C.T., E.M.L., and N.P.K. designed and performed the immunization studies for the HA-capsid fusions. A.D.G. produced and characterized HA-capsid proteins. Y.T.Z., P.L., and H.R.-B. designed and performed the cell signaling assays. C.N., P.R.S., and A.M.B. designed and performed the nuclear localization experiments. All authors analyzed the data. D.B. supervised the research. S.W., I.D.L., and D.B. wrote the manuscript with the input from the other authors. All authors revised the manuscript. **Competing interests:** D.B., S.W., I.D.L., C.N., A.D., N.P.K., and A.J.B. are inventors on a provisional patent application (63/383,700) submitted by the University of Washington for the design, composition, and applications of the protein assemblies described in this work. The remaining authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. MCTS backbone sampling code and structural coordinates of the in silico–generated examples are available at https://github.com/idlutz/protein-backbone-MCTS and archived at Zenodo (*55*). Code for the sequence design and filtering of assemblies is available at https://files.ipd.uw.edu/pub/2023_RL_capsid_design/sequence_design_pipeline.tar and archived at Zenodo (*56*). For RC_I_1, RC_I_2, and RC_I_1-H11 capsid structures, coordinates are deposited in the Protein Data Bank with the accession codes 8F54, 8F53, and 8F4X; cryo-EM density maps are deposited in the Electron Microscopy Data Bank (EMDB) with the accession codes EMD-28860, EMD-28859, and EMD-28858. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adf6591
Materials and Methods
Figs. S1 to S29
Tables S1 to S5
References (*57*–*78*)
MDAR Reproducibility Checklist

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

Submitted 5 November 2022; accepted 21 March 2023
10.1126/science.adf6591

QUANTUM MECHANICS

Schrödinger cat states of a 16-microgram mechanical oscillator

Marius Bild[1,2]†, Matteo Fadel[1,2]†*, Yu Yang[1,2]†, Uwe von Lüpke[1,2], Phillip Martin[1,2], Alessandro Bruno[1,2], Yiwen Chu[1,2]*

According to quantum mechanics, a physical system can be in any linear superposition of its possible states. Although the validity of this principle is routinely validated for microscopic systems, it is still unclear why we do not observe macroscopic objects to be in superpositions of states that can be distinguished by some classical property. Here we demonstrate the preparation of a mechanical resonator in Schrödinger cat states of motion, where the ~10^{17} constituent atoms are in a superposition of two opposite-phase oscillations. We control the size and phase of the superpositions and investigate their decoherence dynamics. Our results offer the possibility of exploring the boundary between the quantum and classical worlds and may find applications in continuous-variable quantum information processing and metrology with mechanical resonators.

Quantum mechanics is one of the most successful scientific theories ever formulated. However, from the early days of quantum mechanics until now, it has been unclear why quantum phenomena, such as state superpositions, are never observed in the macroscopic world. In his 1935 work (*1*), Erwin Schrödinger imagined a device able to poison a cat as a consequence of a radioactive decay, concluding that the superposition of an atom being "decayed" and "not decayed" could be mapped onto a superposition of the cat being simultaneously "dead" and "alive." There are two aspects of this hypothetical scenario that make it seem absurd and counterintuitive: First, a cat is a macroscopic, everyday object; and second, "dead" and "alive" are states that are mutually exclusive within our classical experience.

Many explanations have been proposed as to why we may never encounter a cat in such an unfortunate situation. Macroscopic objects may simply be too complex and subject to too many sources of decoherence to sustain a superposition of classically distinct states. Other theories introduce additional effects beyond standard quantum mechanics, such as wave function collapse due to intrinsic stochastic noise or gravitational decoherence (*2*). These effects are typically expected to scale with the mass of the system and the distinctness of the states that are superposed. Therefore, observing state superpositions in massive objects is of key importance for exploring the validity range of quantum mechanics as we know it. Beyond its fundamental interest, preparing and detecting Schrödinger's cat states is essential for applications in quantum technologies. Main examples include Heisenberg-limited parameter estimation protocols (*3*, *4*) and error-protected quantum information processing (*5*, *6*).

There have been many experimental demonstrations of Schrödinger cat states (which we will call "cat states" from here on). These include superpositions of internal and motional degrees of freedom in trapped ions (*7*, *8*), phase-space superpositions of electromagnetic waves in both the optical (*9*, *10*) and microwave domains (*11*–*13*), Greenberger–Horne–Zeilinger states (*14*, *15*), current superpositions in superconducting quantum interference devices (*16*), and spatial superpositions of atomic clouds (*17*) and large molecules (*18*). We experimentally demonstrate the preparation of cat states in the motional degree of freedom of a solid-state mechanical resonator. Given the variety of definitions found in previous works, we define a cat state of a harmonic oscillator as a coherent superposition of two or more st… with well-separated phase-space distributi…

Experimental setup

Our device, which we call a $\hbar$BAR, as in previous works (*19*), consists of a high-overtone bulk acoustic-wave resonator (HBAR) coupled to a superconducting transmon qubit. The transmon qubit allows us to create, control, and read out phonon states in the HBAR. Qubit and HBAR are fabricated on separate sapphire chips, which are subsequently flip chip bonded into the final device (Fig. 1A). A dome of piezoelectric aluminum nitride on the HBAR chip coherently couples the electric field of the qubit with the strain field of the resonator modes. The device is placed inside a three-dimensional Al cavity, which allows us to control the qubit with the standard circuit quantum electrodynamics toolbox (*20*). The acoustic free spectral range is ~12 MHz, and frequency tuning the qubit via a microwave drive–induced Stark shift allows us to address several longitudinal phononic modes. The acoustic lattice oscillations are localized within a Gaussian mode with waist $w_0 = 27$ μm and length $L = 435$ μm, giving a mode volume of $\pi w_0^2 L \approx 0.001$ mm^3 [see supplementary text (*21*) section B for details]. More details about this circuit quantum acoustodynamics (cQAD) system (*19*, *22*) and the device (*23*) can be found in previous works.

In the classical picture, one can imagine a coherent state $|\alpha\rangle$ in the phonon mode as a coherent displacement of the atomic lattice with an amplitude proportional to α. In the quantum picture, an example of a cat state is a quantum superposition of two coherent states with opposite displacement amplitudes, leading to the physical interpretation of such a state as the superposition of two oscillations

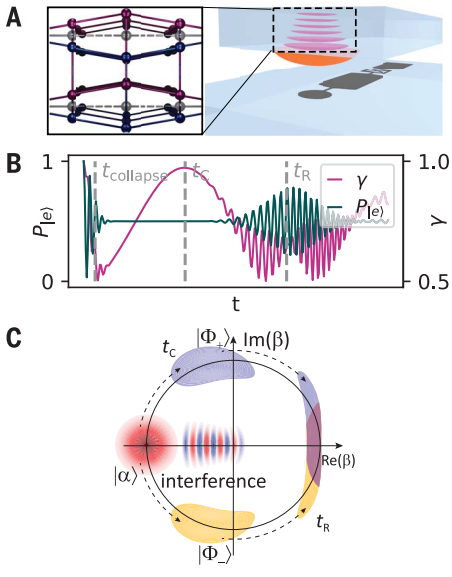


Fig. 1. Illustration of the $\hbar$BAR device and system evolution. (**A**) Schematics of the $\hbar$BAR device. The HBAR chip (top) has a layer of piezoelectric aluminum nitride (orange) and supports standing acoustic waves (pink). The transmon qubit on the lower chip has a circular antenna to couple with the HBAR. The inset shows the superposition of two opposite-phase oscillations of atoms in the crystal lattice. (**B**) Simulated evolution without decoherence of the qubit $|e\rangle$ state population $P_{|e\rangle}$ and purity γ under the JC interaction when the qubit is initialized in $|-Z\rangle$ and the phonon in a coherent state. (**C**) Illustration of the evolution of an initial phonon coherent state (red circle on the left) in phase space. The blue (yellow) crescent shapes indicate the state $|\Phi_+\rangle$ ($|\Phi_-\rangle$), which is the phonon state when the qubit is initialized in $|+X\rangle$ ($|-X\rangle$). Interference fringes appear around time t_C when the qubit is prepared in a superposition of $|+X\rangle$ and $|-X\rangle$. Around the revival time t_R, the two phonon states again overlap (purple).

[1]Department of Physics, ETH Zürich, 8093 Zürich, Switzerland. [2]Quantum Center, ETH Zürich, 8093 Zürich, Switzerland.
*Corresponding author. Email: fadelm@phys.ethz.ch (M.F.); yiwen.chu@phys.ethz.ch (Y.C.)
†These authors contributed equally to this work.

of the atomic lattice with the same frequency ω_p and relative phase π. Considering a snapshot in time where both oscillations are at their displacement maximum, Schrödinger's cat being in a superposition of dead and alive is analogous to a superposition of atoms in the HBAR being in two distinct positions in space, as illustrated in the inset of Fig. 1A. Here, we define the positions as distinct when their separation is larger than the fluctuations due to quantum, thermal, or other sources of noise.

Generating cat states

To realize a cat state in our system, we use the Jaynes-Cummings (JC) interaction with the qubit and phonon on resonance (*11*, *24*). The interaction Hamiltonian is

$$H/\hbar = g_0\left(\sigma^+ a + \sigma^- a^\dagger\right) \qquad (1)$$

where g_0 is the coupling strength between qubit and phonon mode, σ^+ is the raising operator for the qubit, and $a^\dagger$ is the raising operator for the phonon mode. This results in Rabi oscillations between the states $|e, n-1\rangle$ and $|g, n\rangle$ at a rate $g_0\sqrt{n}$, where $|g\rangle$ ($|e\rangle$) is the qubit ground (excited) state and $|n\rangle$ is the Fock state of n phonons. As a consequence of this $\sqrt{n}$ scaling, if the phonon mode is prepared in a coherent state with large enough amplitude and the qubit is prepared in $|g\rangle$ or $|e\rangle$, their coherent interaction rapidly dephases. Hence, the oscillations of the qubit population "collapse" (Fig. 1B) with a decaying amplitude proportional to (*25*, *26*) $\exp\left(-(t/t_{\text{collapse}})^2\right)$, where $t_{\text{collapse}} = \sqrt{2}/g_0$ is the collapse time in the limit of $\alpha \gg 1$. At this time, the qubit and phonon states are entangled. This can be seen in Fig. 1B as a minimum in the qubit state purity $\gamma(t) = \text{Tr}\left(\rho_q(t)^2\right)$ around t_{collapse}, where $\rho_q(t)$ is the reduced density matrix of the qubit. Notably, owing to the quantized phonon energy and the consequent discrete oscillation frequency spectrum, the oscillations revive in finite time (*24*). For $\alpha \gg 1$, this revival occurs at $t_R = 2\pi\alpha/g_0$ (*27*). Between the collapse and revival, at time $t_R/2$, the qubit and phonon disentangle from each other. The state being separable at $t_R/2$ coincides with the occurrence of a superposition of two distinct states in phase space, realizing a cat state in the phonon mode (*11*, *21*, *24*, *28*).

A more intuitive explanation for the origin of the cat state comes from the time evolution of the reduced phonon state in phase space. As illustrated in Fig. 1C and shown in supplementary text section A, if the qubit is initialized in the state $|\pm X\rangle \equiv (|e\rangle \pm |g\rangle)/\sqrt{2}$ and in the limit of large α, the evolution leads to a rotation in phase space with an angular velocity $\mp|g_0/2\alpha|$ and a distortion of the coherent states. We call the resulting states $|\Phi_\pm(t)\rangle$, whose full expressions are given in the supplementary text (eq. S15). If the the qubit is initialized in the state $|\pm Z\rangle \equiv (|+X\rangle \pm |-X\rangle)/\sqrt{2}$, the phonon state will evolve into $|\Phi_+(t)\rangle \pm |\Phi_-(t)\rangle$ as shown in Fig. 1C. At time $t_R/2$, the two state components $|\Phi_\pm\rangle$ have covered a rotation angle of $\mp\pi/2$ around a circle of radius α, maximizing their separation in phase space and forming a cat state (*24*). Finally, at the revival time t_R, the two phonon state components $|\Phi_\pm\rangle$ have both rotated by a phase of π and approximately recombine in phase space.

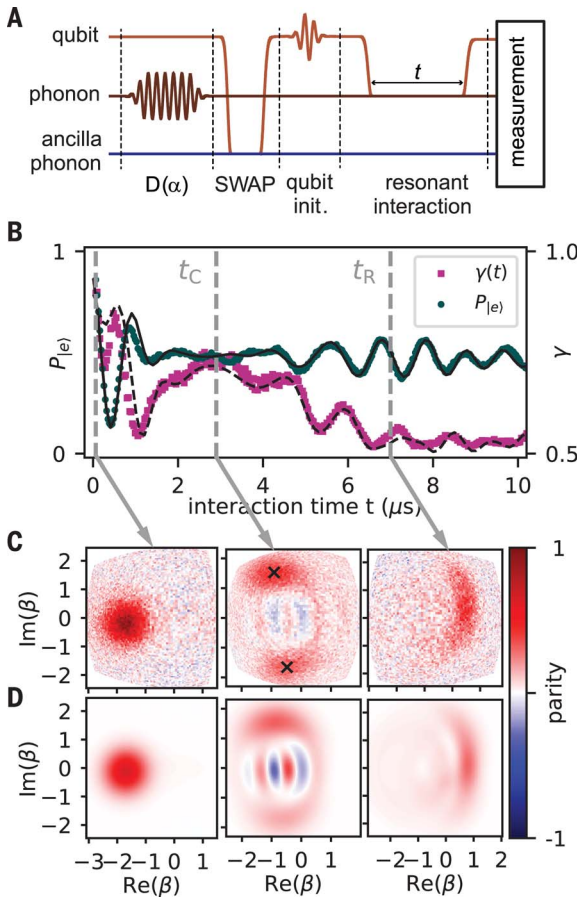


Fig. 2. Collapse and revival dynamics. (**A**) Experimental sequence for observing collapse and revivals dynamics and for preparing cat states (details in the main text). (**B**) Measured qubit population and state purity. The solid and dashed black lines are the simulation results of the qubit population and purity, respectively. Three time points of particular interest are highlighted (dashed lines): initial state time, cat state time (t_C), and revival state time (t_R). (**C**) Measured Wigner function of the phonon state at the three time points. Axes are the real and imaginary parts of the complex displacement amplitude β used during Wigner tomography (*23*). The black crosses indicate the positions of the two coherent states composing the fitted CSS state Eq. 2) (**D**) Corresponding simulated Wigner functions.

Experimental results

We experimentally confirm both the predicted collapse and revival of Rabi oscillations and the creation of mechanical cat states in the phonon mode. The basic sequence used in the experimental demonstration of the JC dynamics described above can be seen in Fig. 2A. We displace the phonon mode with a resonant drive of amplitude A to a coherent state with amplitude α. To mitigate any effect of the drive on the qubit state, we then cool the qubit with an ancillary phonon mode (*22*, *23*). The qubit is subsequently prepared in its initial state by applying a drive pulse with variable phase and amplitude. To induce the resonant interaction, we tune the qubit to the phonon mode frequency for a variable interaction time t. Depending on which of the subsystems we want to characterize, we choose a measurement sequence that implements the appropriate measurement operator. First, we simply measure the qubit excited state population. The resulting data are shown in Fig. 2B for $A = 0.35$. Here the value for A is a scaling factor for the amplitude of a microwave drive, which we calibrate to find a corresponding initial coherent state size of $\alpha = 1.75$ (supplementary text section E). As expected, we observe oscillations that collapse after a time $t_{\text{collapse}} \approx 0.9$ μs and revive at $t_R \approx 6.7$ μs. This revival indicates the coherent exchange of energy quanta between the qubit and phonon mode during the resonant interaction. By performing full qubit tomography after the resonant interaction, we can also reconstruct the reduced density matrix of the qubit subsystem ρ_q and calculate the purity of the qubit state $\gamma(t)$. We confirm a local minimum of $\gamma(t)$ around the predicted collapse time t_{collapse}, followed by a local maximum around $t_R/2$ (Fig. 2B).

We now focus on the time evolution of the phonon subsystem by performing full Wigner tomography of the phonon state after the resonant interaction times $t = 0$, 2.9 and 7.0 μs. To this end, we use the parity measurement technique established in a previous work (*23*). To compensate for the effect of qubit dephasing during the parity measurement, we normalize

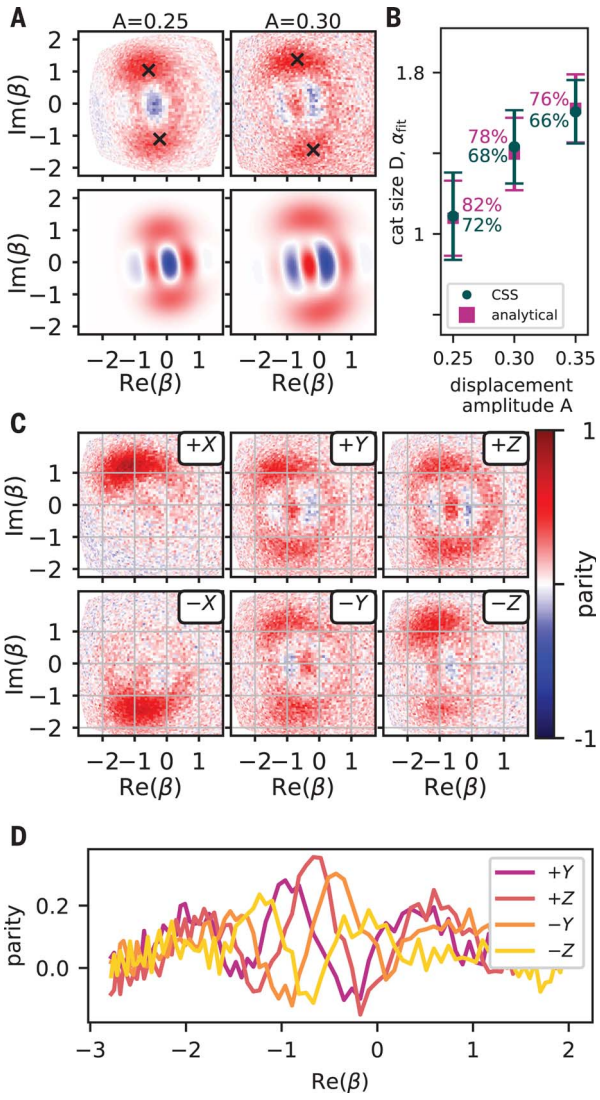


Fig. 3. Cat state amplitude and phase control. (**A**) Cat states prepared with different displacement pulse amplitudes *A*. Top row: Measured Wigner functions; bottom row: analytical state $\rho(t_C)$ that best fits the data. The fitted CSS states, with coherent state positions indicated by black crosses, have $D = 1.09$ (1.43) for $A = 0.25$ (0.30). (**B**) Cat state sizes as a function of the displacement amplitude *A*, obtained from fitting the data in (A) to analytical states $\rho(t_C)$ and CSS states $|C\rangle$. Numbers are the fidelity with respect to the fitted state. Error bars show cat state sizes resulting in 1 % deviation in the fidelity (supplementary text section G) (**C**) Cat states resulting from the initial qubit states indicated by the respective labels. (**D**) Cross-cuts of interference fringes from (C) for four initial qubit states, obtained by averaging the data between $-0.11 < \mathrm{Im}(\beta) < 0.16$.

all measured parity values to that of the Fock $|0\rangle$ phonon state (supplementary text section D). The measured Wigner functions are shown in Fig. 2C, where axes in phase space are normalized by measuring the distribution of populations in the phonon Fock states for coherent states created with different drive amplitudes (*22*) (supplementary text section E).

From the measured data, we confirm the evolution of the initial coherent state (Fig. 2C, left) into a cat state at $t_C = 2.9$ μs (Fig. 2C, center), showing two state components clearly distinct in phase space and interference fringes located between them. We choose this value of t_C because it corresponds to the measured maximum in the qubit state purity. It deviates somewhat from the value predicted by using the large α limit, which is $t_R/2 \approx 3.3$ μs. For the evolution time $t = 7.0$ μs, the predicted refocusing into a crescent-shaped overlap between the counter-rotating state components can be observed (Fig. 2C, right).

To benchmark the cat state and obtain an estimate of its size, we implement a maximum-likelihood reconstruction (*29*) of the phonon state ρ_p from the measured state with $A = 0.35$ and $t_C = 2.9$ μs. We then fit the reconstructed state to an analytical expression of the expected phonon state $\rho(t_C)$ in the absence of decoherence and after tracing out the qubit (supplementary text section A). Fixing the interaction time to t_C from the experiment, the fit maximizes the fidelity between ρ_p and $\rho(t_C)$ by varying the initial coherent state size α_{fit} of $\rho(t_C)$. The result yields a fidelity of $\mathcal{F} \approx 76\%$ to an analytical state with initial coherent state size $\alpha_{fit} = 1.62$, which is smaller than the initial displacement $\alpha = 1.75$ because the expression for $\rho(t_C)$ does not include phonon losses. We attribute the infidelity to a combination of decoherence and measurement imperfections that lead to additional artifacts in the Wigner function (*23*). To further confirm that the phonon state behaves as expected, Fig. 2D shows the results of a master equation simulation of the full experimental protocol with independently measured system parameters, showing good agreement with the measurements in Fig. 2C.

The state we obtained resembles the two-component coherent state superpositions (CSS)

$$|C\rangle = \mathcal{N}\left(|\alpha_1\rangle + e^{i\vartheta}|\alpha_2\rangle\right) \qquad (2)$$

a type of cat state that is often invoked in quantum information (*5*, *6*) and parameter estimation protocols (*3*, *4*). Here $|\alpha_{1,2}\rangle$ are two coherent states and $\mathcal{N}$ is the appropriate normalization constant. We can fit our reconstructed state to Eq. 2 by optimizing α_1, α_2, and ϑ for maximum fidelity $\mathcal{F}\left(\rho_p, |C\rangle\langle C|\right)$. Because our state is not centered around the origin in phase space, we use half the phase space distance between the coherent state components $D = |\alpha_1 - \alpha_2|/2$ as a measure of the cat state size. This choice is motivated by considering a coherent state superposition centered around the origin in phase space, such that $\alpha_1 = -\alpha_2$. Then $D = |\alpha_{1,2}| = \sqrt{\bar{n}}$, where $\bar{n}$ is the average phonon population of the state created. For the state in Fig. 2C, we obtain a state size $D = 1.61$, corresponding to $\bar{n} = D^2 = 2.60$, with a fidelity of $\mathcal{F} \approx 66\%$. The smaller state size D compared to the initial coherent displacement is a combination of decoherence and the choice of interaction time $t_C < t_R/2$, resulting in the two counter-rotating state components not reaching their maximum separation in phase space. The fidelity is lower compared to the fitted analytical state $\rho(t_C)$, because $\rho(t_C)$ itself has a finite infidelity to the CSS state $|C\rangle$.

We can now translate the parameters of the measured cat state into physical properties of the phonon mode, such as the spatial separation between atoms. A state size of $D = 1.61$ corresponds to a maximal delocalization of $7.0 \cdot x_{ZPF}$, where x_{ZPF} is the zero point motion of an equivalent one-dimensional (1D) quantum harmonic oscillator (supplementary text section B). Because we are not considering a center-of-mass mode, there is some freedom in choosing x_{ZPF}, which is then associated with an effective oscillating mass of the mode. If we choose the root-mean-square (RMS) value of the atomic displacements, we find an effective mass of $M_{eff}^{RMS} = 16.2$ μg, corresponding to ~10^{17} atoms, delocalized over a distance of 2.1×10^{-18} m (supplementary text section B).

In applications such as bosonic encodings of a qubit state, full control over the phase and amplitude of the created cat state is required (*5*, *6*, *30*). In the following, we demonstrate this level of control in our experiment. By varying the amplitude *A* of the phonon displacement drive, we can control the amplitude of the initial coherent state and the size of the resulting cat state. For displacement amplitudes $A = 0.25$ and 0.30, we create cat states with $D = 1.09$

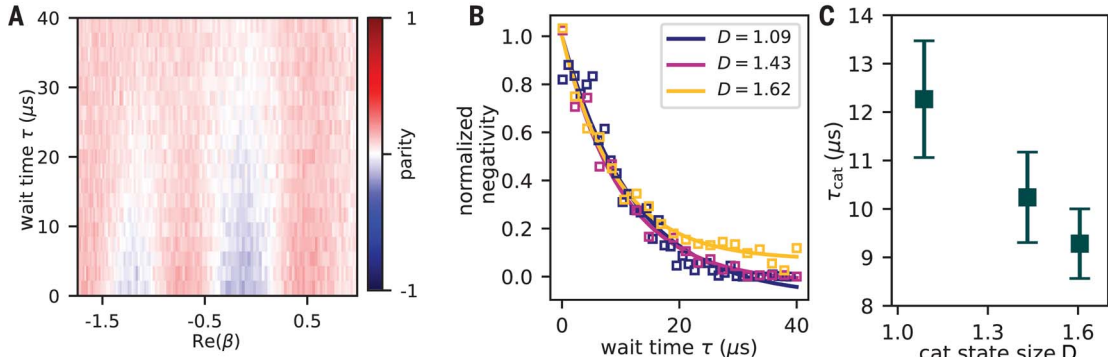


Fig. 4. Decoherence of cat states. (**A**) Measured 1D cuts through the interference fringes of the $D = 1.43$ cat state for a range of wait times between state creation and measurement. (**B**) Extracted negativities (squares) from each cut versus wait times for three cat state sizes, together with fitted exponential decays (solid lines). Both data and fitted curves are normalized to the fitted value at $\tau = 0$. (**C**) Characteristic decay times τ_{cat} extracted from the fits in (B) for all three cat state sizes. Error bars are uncertainties extracted from the fit.

and 1.43, respectively (Fig. 3A). The fidelities of reconstructions of the measured states to both a CSS state and $\rho(t_C)$ are given in Fig. 3B. The best fit $\rho(t_C)$ is plotted in the lower row of Fig. 3A, showing good qualitative agreement with the data. As before, the finite fidelities and lower-contrast fringes of the measured states as compared to the best-fit $\rho(t_C)$ arise mainly from decoherence of the state during measurement.

In the two-component cat state encoding of a qubit, the six cardinal points of the Bloch sphere are given by two coherent states $|\alpha_1\rangle$ and $|\alpha_2\rangle$, along with their four superpositions with $\pi/2$ difference in the phase ϑ (Eq. 2). We can prepare similar states by initializing the transmon qubit state in all six cardinal points $|\pm X\rangle, |\pm Y\rangle, |\pm Z\rangle$ of its Bloch sphere before performing the cat state generation protocol. The preparation of $|\pm X\rangle$ and $|\pm Y\rangle$ is calibrated using collapse and revival measurements as a function of the qubit drive phase (supplementary text section A). The results for $A = 0.35$ and $t_C = 2.10$ μs are shown in Fig. 3C. We observe a distorted coherent state located on the upper (lower) half of phase space for the qubit initially in $|\pm X\rangle$. This separation in phase space is expected from the opposite rotation directions between the phonon states when the qubit is prepared in $|\pm X\rangle$ (Fig. 1C). The $|\pm Y\rangle, |\pm Z\rangle$ states then give rise to four cat states that differ in phase by $\pi/2$, as can be observed in the phases of the interference fringes in Fig. 3 D. The initial energy of the qubit, and thus of the total system, is not the same for all six scenarios, resulting in slightly different sizes for the cat states. In the limit of large cat state size, this difference becomes negligible, and the phonon subspace maps onto that of the cat state encoding.

Superposition states are nonclassical states that are notoriously prone to decoherence. We now investigate the quantum-to-classical transition of different-sized cat states by letting them evolve freely for a varying wait time τ before performing Wigner tomography. During this evolution, the qubit is far detuned from the phonon mode. In particular, we focus on a slice through the Wigner function's interference fringes at Im(β) = 0, which highlights the nonclassical features of the superposition. Figure 4A shows the time evolution of this slice for the $D = 1.43$ cat state. We observe that the negative features disappear on a time scale much faster than $T_1^{ph} \approx 84$ μs, the energy relaxation time of the phonon mode.

As a measure for the nonclassicality of the state, we extract the-time dependent negativity (*31*), defined as $\delta(t) \equiv \int(|W(\beta,t)| - W(\beta,t))d\beta$. Here, $W(\beta,t)$ is the measured Wigner function of the cat state at time t, and the integration is over the 1D slice in phase space parameterized by the complex displacement amplitude β. Figure 4B shows the resulting $\delta(t)$ for the three different cat state sizes of Fig. 3B. We fit each dataset to an exponential decay plus a constant offset. The offset in the measured Wigner values arises from the fact that our Wigner tomography is not performed in the ideal dispersive limit (*23*). The extracted decay time scales τ_{cat} are plotted in Fig. 4C. We show in section H of the supplementary text that, in the limit of large $|\alpha|$, $\delta(\tau)$ decays exponentially with a time constant $\tau_{cat} = T_1^{ph}/(2|\alpha|^2)$. However, for small $|\alpha|$, τ_{cat} deviates from this expression and is in fact dependent on properties of the exact state, such as the phase of the superposition. The data in Fig. 4C show the expected qualitative behavior of faster-decaying negativity for larger-sized cat states, and we present a more detailed quantitative analysis in the supplementary text (*21*).

Concluding remarks

Our results show the generation of cat states in a microgram-mass solid-state mechanical mode using the tools of cQAD and pave the way toward using such systems for tests of wave function collapse models (*32*). These tests would benefit from larger-sized cat states, resonators with higher masses (*33*), and longer phonon lifetimes. To facilitate comparison with other mechanical resonators and theoretical models that often consider center-of-mass motion, we note that the HBAR mode is a standing wave in which a half-wavelength section approximates a center-of-mass mode where all atoms oscillate in the same direction. Such a section has a mass on the order of 30 ng, obtained by dividing the total effective mass of 16 μg by the longitudinal mode number of ~500.

The maximum size of the cat state that we can prepare is currently limited by our device parameters, including both the qubit and phonon decoherence rates. The latter is especially important given that, in general, the decoherence rate of the cat state is proportional to the square of the cat state size D. Furthermore, additional improvements to the properties of qubit and phonon resonator would enable alternative cat state generation protocols that can in principle lead to states with a higher fidelity to, for example, a CSS state (*12*, *13*). We point out, however, that although CSS states represent a useful benchmark (because they have been extensively studied for applications such as quantum information and quantum metrology), many of their salient features are present already in the states that we have demonstrated. These include the phase-space separation of state components, which is important for error protection of encoded qubits (*30*, *34*, *35*), and the presence of interference fringes with high Fisher information, which is useful for quantum-enhanced sensing (*4*, *36*, *37*).

REFERENCES AND NOTES

1. E. Schrödinger, *Naturwissenschaften* **23**, 807–812 (1935).
2. A. Bassi, K. Lochan, S. Satin, T. P. Singh, H. Ulbricht, *Rev. Mod. Phys.* **85**, 471–527 (2013).
3. L. Pezzè, A. Smerzi, *Atom Interferometry, Proceedings of the International School of Physics "Enrico Fermi," Varenna*, G. M. Tino, M. A. Kasevich, eds. (IOS Press, Amsterdam, 2014).
4. W. J. Munro, K. Nemoto, G. J. Milburn, S. L. Braunstein, *Phys. Rev. A* **66**, 023819 (2002).
5. P. T. Cochrane, G. J. Milburn, W. J. Munro, *Phys. Rev. A* **59**, 2631–2634 (1999).
6. M. Mirrahimi *et al.*, *New J. Phys.* **16**, 045014 (2014).
7. C. Monroe, D. M. Meekhof, B. E. King, D. J. Wineland, *Science* **272**, 1131–1136 (1996).
8. H.-Y. Lo *et al.*, *Nature* **521**, 336–339 (2015).
9. A. Ourjoumtsev, H. Jeong, R. Tualle-Brouri, P. Grangier, *Nature* **448**, 784–786 (2007).
10. K. Huang *et al.*, *Phys. Rev. Lett.* **115**, 023602 (2015).
11. A. Auffeves *et al.*, *Phys. Rev. Lett.* **91**, 230405 (2003).
12. S. Deléglise *et al.*, *Nature* **455**, 510–514 (2008).
13. B. Vlastakis *et al.*, *Science* **342**, 607–610 (2013).
14. D. Leibfried *et al.*, *Nature* **438**, 639–642 (2005).
15. W.-B. Gao *et al.*, *Nat. Phys.* **6**, 331–335 (2010).

16. J. R. Friedman, V. Patel, W. Chen, S. K. Tolpygo, J. E. Lukens, *Nature* **406**, 43–46 (2000).
17. T. Kovachy *et al*., *Nature* **528**, 530–533 (2015).
18. Y. Y. Fein *et al*., *Nat. Phys.* **15**, 1242–1245 (2019).
19. Y. Chu *et al*., *Science* **358**, 199–202 (2017).
20. A. Blais, A. L. Grimsmo, S. M. Girvin, A. Wallraff, *Rev. Mod. Phys.* **93**, 025005 (2021).
21. See supplementary materials.
22. Y. Chu *et al*., *Nature* **563**, 666–670 (2018).
23. U. von Lüpke *et al*., *Nat. Phys.* **18**, 794–799 (2022).
24. V. Bužek, H. Moya-Cessa, P. L. Knight, S. J. D. Phoenix, *Phys. Rev. A* **45**, 8190–8203 (1992).
25. F. W. Cummings, *Phys. Rev.* **140**, A1051–A1056 (1965).
26. G. Rempe, H. Walther, N. Klein, *Phys. Rev. Lett.* **58**, 353–356 (1987).
27. J. H. Eberly, N. B. Narozhny, J. J. Sanchez-Mondragon, *Phys. Rev. Lett.* **44**, 1323–1326 (1980).
28. J. Gea-Banacloche, *Phys. Rev. Lett.* **65**, 3385–3388 (1990).
29. K. S. Chou *et al*., *Nature* **561**, 368–373 (2018).
30. A. Grimm *et al*., *Nature* **584**, 205–209 (2020).
31. A. Kenfack, K. Zyczkowski, *J. Opt. B Quantum Semiclassical Opt.* **6**, 396–404 (2004).
32. B. Schrinski *et al*., *Phys. Rev. Lett.* **130**, 133604 (2023).
33. C. Whittle *et al*., *Science* **372**, 1333–1336 (2021).
34. N. Ofek *et al*., *Nature* **536**, 441–445 (2016).
35. C. Flühmann *et al*., *Nature* **566**, 513–517 (2019).
36. J. Joo, W. J. Munro, T. P. Spiller, *Phys. Rev. Lett.* **107**, 083601 (2011).
37. A. Facon *et al*., *Nature* **535**, 262–265 (2016).
38. M. Bild *et al*., Schrödinger cat states of a 16-microgram mechanical oscillator, Zenodo (2023); https://doi.org/10.5281/zenodo.7701912.

ACKNOWLEDGMENTS

We thank O. Romero-Isart, A. Grimm, and I. C. Rodrigues for useful discussions, and A. Brooks for help with figure making. Fabrication of devices was performed at the FIRST cleanroom of ETH Zürich and the BRNC cleanroom of IBM Zürich. **Funding:** M.B. was supported by the QuantERA II Program that has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no 101017733, and with the Swiss National Science Foundation. M.F. was supported by The Branco Weiss Fellowship–Society in Science, administered by the ETH Zürich. P.M. was supported by the ThinkSwiss fellowship. **Author contributions:** U.v.L. fabricated the device. M.B., M.F., Y.Y., and U.v.L. performed the experiments and analyzed the data. All authors performed theoretical calculations and simulations. Y.C. conceived of the project and supervised the work. M.B., M.F., Y.Y., U.v.L., and Y.C. wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data and software are available in the manuscript or the supplementary material or are deposited at Zenodo (*38*).
License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adf7553
Supplementary Text
Figs. S1 to S8
Tables S1 and S2
References (*39*–*46*)

Submitted 10 November 2022; accepted 20 March 2023
10.1126/science.adf7553

OPTICAL FIBERS

Scaling information pathways in optical fibers by topological confinement

Zelin Ma[1], Poul Kristensen[2], Siddharth Ramachandran[1]*

Spatial mode-count scalability in optical fibers is of paramount importance for addressing the upcoming information-capacity crunch, reducing energy consumption per bit, and for enabling advanced quantum computing networks, but this scalability is severely limited by perturbative mode mixing. We show an alternative means of light guidance, in which light's orbital angular momentum creates a centrifugal barrier for itself, thereby enabling low-loss transmission of light in a conventionally forbidden regime wherein the mode mixing can be naturally curtailed. This enables kilometer-length-scale transmission of a record ~50 low-loss modes with cross-talk as low as −45 decibels/kilometer and mode areas of ~800 square micrometers over a 130-nanometer telecommunications spectral window. This distinctive light-guidance regime promises to substantially increase the information content per photon for quantum or classical networks.

Light transport with optical fibers in multiple orthogonal dimensions of a photon is of great utility for increasing the information capacity of classical communications networks, reducing energy consumption per bit, as well as for facilitating development of advanced quantum computing networks in high-dimensional Hilbert spaces (*1*, *2*). With degrees of freedom such as wavelength, amplitude, phase, and polarization having been exhausted, spatial multiplexing remains the last available dimension to be considered (*3*). Two primary approaches for exploiting space comprise the use of multimode fibers (MMFs), whose spatial modes are mutually orthogonal, and multicore fibers (MCFs), which consist of multiple single-mode waveguide cores. MCFs are appealing, given their deployment simplicity and backward compatibility, but their individual cores must be sufficiently spatially separated to avoid cross-talk, a requirement that limits channel-count scalability because reliability considerations place an upper limit on overall fiber dimensions (*3*). MMFs, by contrast, feature several benefits, such as high spatial efficiency—because all spatial modes occupy the same waveguiding core—and cost and efficiency savings through shared-pump optical amplification (*4*). However, modes in conventional MMFs stochastically couple with each other because of ever-present ambient perturbations, a problem that can be addressed through the use of multi-input–multi-output digital signal processing (MIMO-DSP) in classical communications systems (*3*), wherein mode mixing may actually be preferred. But this necessarily increases the complexity and power consumption of the receiver, and, crucially, this solution cannot be applied to single photons in quantum networks. An ideal solution would be an MMF whose modes behave like the separated cores of an MCF, because it would deliver the benefits of both approaches. However, MMFs designed to limit mode mixing over kilometer-length scales have achieved only up to 12 modes (*5*), and even so, with cross-talk considerably worse than those of MCFs. Further scalability has eluded the scientific community for almost a decade.

Here, we demonstrate a mechanism to transmit light over long lengths of optical fibers by exploiting a topological effect wherein the photon remains confined to a fiber as a result of a centrifugal barrier that light's orbital angular momentum (OAM) creates for itself. This form of light guidance is forbidden by the principles of total internal reflection (TIR), the primary means for light transport demonstrated by Colladon, Babinet, and Tyndall more than 150 years ago. Not restricted by the TIR condition, this regime of light transport results in propagating eigenstates that are naturally immune to perturbative mode mixing. We validate the practical utility of these modes, hereafter called topologically confined modes (TCMs), by demonstrating light transmission over a record number of low-loss spatial modes with negligible mode coupling over a kilometer of fiber, featuring mode areas that are an order of magnitude greater than that available from single-mode fibers (SMFs) or MCFs, even when they are tightly bent. We show that these beneficial metrics can be obtained over multiple spectral bands in use for telecommunications today.

Concept and demonstration of topological confinement

A heuristic ray picture of the concept of topological confinement is shown in Fig. 1A (*6*). It schematically depicts a step-index fiber comprising a high-index core, n_{co}, surrounded by a low-index cladding, n_{cl}, and three exemplary eigenmodes with similar effective indices, $n_{eff} < n_{cl}$. These modes, by definition, do not satisfy the TIR condition, and hence are normally expected to have high confinement loss because they radiate away from the waveguide; hence they acquired the common nomenclature of "cutoff" or "leaky" modes (*7*). However, these properties have a strong dependence on mode symmetry. Fiber eigenmodes are quantized by two transverse indices: L, denoting the azimuthal index resulting in the mode carrying OAM, and a radial index, m (manifesting in oscillatory field amplitude in the radial direction). Including polarization [left- or right-handed circular polarization (LCP or RCP)] orthogonality, a given $|L|$ comprises four modes: a pair of spin-orbit aligned (SOa) modes ($+|L|$, LCP and $-|L|$, RCP) with n_{eff} slightly different from a pair of spin-orbit antialigned (SOaa) modes ($-|L|$, LCP and $+|L|$, RCP) (*8*). As Fig. 1A shows, the transverse momentum k_T is progressively more oriented in the azimuthal rather than the radial direction as $|L|$ increases. Because the "escape" rate of a mode tends to increase as k_T is progressively more radially oriented, it follows that the confinement loss of unbound, radiative modes decreases as the mode's azimuthal index increases.

A more rigorous illustration of topological confinement is obtained by considering the waveguide eigenvalue equation (*9*)

$$\frac{d^2F(r)}{dr^2} + \frac{1}{r}\frac{dF(r)}{dr} + \left[k_0^2 n^2(r) - \frac{L^2}{k_0^2 r^2} - \beta^2\right]F(r) = 0 \quad (1)$$

where r is the radial coordinate, $F(r)$ is the radial amplitude of the electrical field, $n(r)$ is the refractive index profile of the fiber, $k_0 = 2\pi/\lambda$ is the free-space wave vector of light of wavelength λ, and β ($= k_0 \cdot n_{eff}$) is the propagation constant of an eigenmode. The attractive potential that enables TIR-induced bound modes is $n^2(r)$, but in the presence of OAM, a modified effective potential may be defined as

$$n^2_{centrifugal}(r) = n^2(r) - \frac{L^2}{k_0^2 r^2} \quad (2)$$

Figure 1B is a plot of this topologically modified profile, $n_{centrifugal}(r)$, experienced by cutoff modes with three nonzero L's, calculated at different wavelengths such that they have identical $n_{eff} < n_{cl}$. As is evident, cutoff modes encounter refractive index barriers (index trenches) of progressively greater magnitudes as $|L|$ increases. Hence, although the mode is no longer strictly bound by TIR (so that it no longer experiences zero confinement loss), its confinement loss strongly depends on $|L|$ because of the centrifugal barrier that light's OAM creates for itself.

Figure 1C shows the simulated confinement loss versus $|L|$ for modes with radial orders

[1]Boston University, Boston, MA 02215, USA. [2]OFS-Fitel, 2605 Brøndby, Denmark.
*Corresponding author. Email: sidr@bu.edu

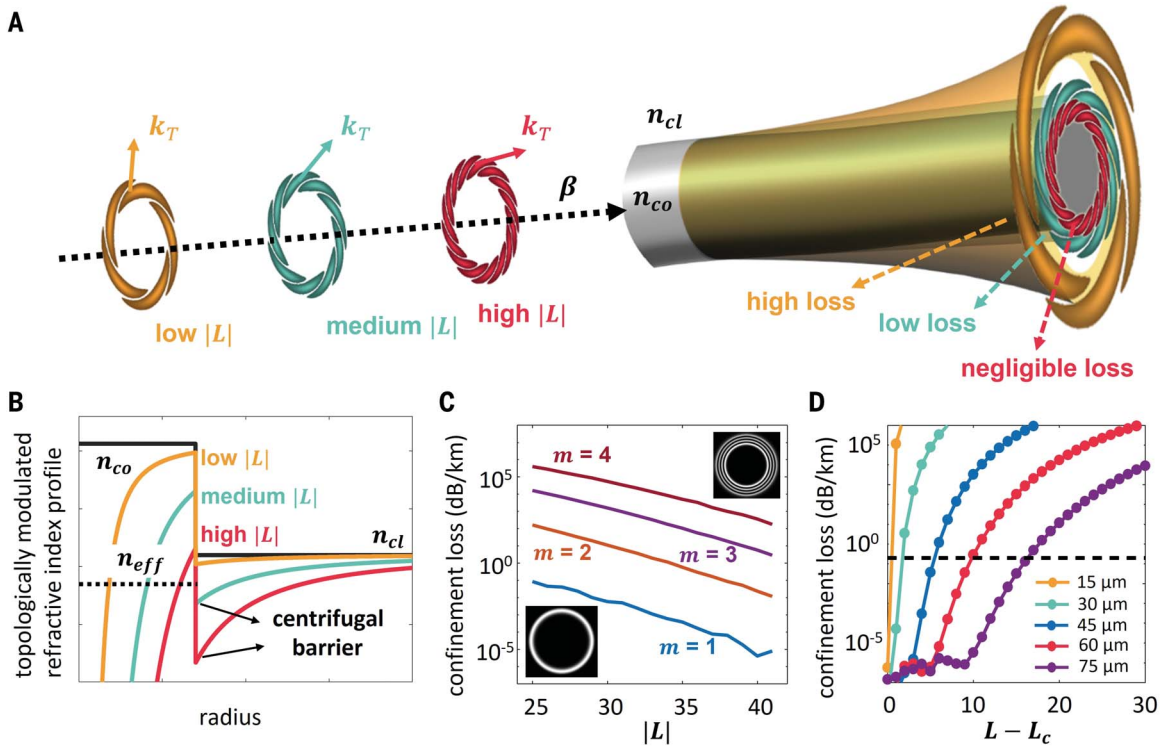


Fig. 1. The principle of light transport with TCMs. (**A**) The propagation of three cutoff OAM modes with different L's but identical n_{eff} in an exemplary step-index fiber. Select OAM modes in this fiber are illustrated as spiral patterns for visual clarity, with the number of spiral arms equal to $|L|$. Confinement loss of these modes decreases as $|L|$ increases because their transverse wave vectors, k_T, become progressively more azimuthally oriented. (**B**) Topologically modulated refractive index profile for the three cutoff OAM modes shown in (A). Higher $|L|$ induces a deeper index trench, creating a centrifugal barrier that prevents the cutoff modes from leaking out. (**C**) Simulated confinement loss versus mode orders L and m in a step-index fiber with 60-μm ring diameter. The n_{eff}'s of all the modes are held 1% below n_{cl}. Insets are simulated modal images of m = 1 and m = 4 modes. (**D**) Simulated confinement loss at 1550 nm versus relative OAM order $L - L_c$, in five step-index fibers with the same index contrast but different core sizes. L_c is the OAM order of the last TIR bound mode (L_c = 7, 16, 25, 35, and 45 for fiber core sizes of 15, 30, 45, 60, and 75 μm, respectively). Dashed line represents the overall measured loss (0.2 dB/km) of SMF-28.

$m \in [1,4]$ of a step-index fiber with $n_{co} - n_{cl}$ = 0.04 and core diameter of 60 μm, using a standard-perfectly-matched-layer model (*10*) (data points are simulated at different λ's such that n_{eff}/n_{cl} is identical for all modes). Modes of low $|L|$s, and modes of high m regardless of $|L|$, are highly lossy, justifying their long-hld nomenclature as "cutoff" or "leaky" modes, but for first radial order (m = 1) modes of sufficiently high $|L|$, confinement loss decreases substantially, down to as low as ~10^{-6} dB/km. Hence, although TCMs with high $|L|$ and m = 1 do not satisfy TIR, their confinement loss suggests that their behavior mirrors that of conventional TIR bound modes.

Figure 1D shows the simulated confinement loss versus relative topological charge ($L - L_c$, where L_c is the topological charge of the last bound mode, guided conventionally by TIR, before cutoff) at 1550 nm in five different step-index fibers of identical index contrasts (as in Fig. 1C), but with core diameters ranging from 15 to 75 μm. Also shown is the overall measured loss (~0.2 dB/km) for conventional transmission fibers (e.g., SMF-28). The number of modes that have confinement loss substantially lower than SMF-28 loss increases with $L - L_c$. This yields a crucial design criterion for realizing fibers that scale mode count with TCMs. Any fiber with a discrete index step at its outer core boundary (required for the existence of a centrifugal barrier; Fig. 1B) that guides, through conventional TIR, sufficiently high-$|L|$ modes (high L_c), has a large ensemble of even higher-$|L|$ cutoff modes that also experience negligible confinement loss, and the number of available TCMs simply increases with TIR bound-mode count. This rule implies that TCMs exist in any highly multimoded waveguide, comprising either high index contrasts, large core sizes, or both.

We tested this concept using a ring-core fiber with a ring outer diameter of 56.6 μm because it is functionally similar to a step-index fiber (fig. S1) (*11*–*13*). Figure 2A shows the cutback loss measurements on all SOa and SOaa modes with m = 1, $|L| \in [25,43]$. Conventional TIR bound modes have an average loss of ~1.4 dB/km, whereas TCM losses are ~5.1 dB/km (ignoring high-loss outliers that are not included in mode counts discussed in the next section). As is evident, TCM losses are orders of magnitude lower than conventional wisdom posits for cutoff modes, validating the idea that centrifugal barriers greatly aid light transmission for modes that are supposed to be radiated away. Barring a few transition TCMs that interact with parasitic high-m modes (fig. S4), the measured loss scales adiabatically, and only slightly, with mode order $|L|$. Moreover, we find that TCM losses decrease considerably with increasing fiber-draw tension (fig. S3). These findings suggest that interfacial scattering loss at the core-cladding boundary plays a dominant role in the loss measured in the current fiber (*14*, *15*), and that further glass viscosity and drawing-parameter optimizations promise substantially lower losses for TCMs. Given that simulated confinement losses for the TCMs (Fig. 2A, black dashed curve) are as low as ~10^{-6} dB/km, and that low-loss telecommunications-grade fibers with similar index contrasts and index gradients are regularly available (*16*), we expect manufacturing optimizations to substantially reduce TCM losses down to those of commercial high-index-contrast SMFs.

This topological effect is analogous to the OAM-dependent, above-threshold ionization of electrons in atomic orbitals (*17*) and to the metastability of Feshbach molecules in high rotational states (*18*). The effect is also well known in the context of nuclear reactions (*19*).

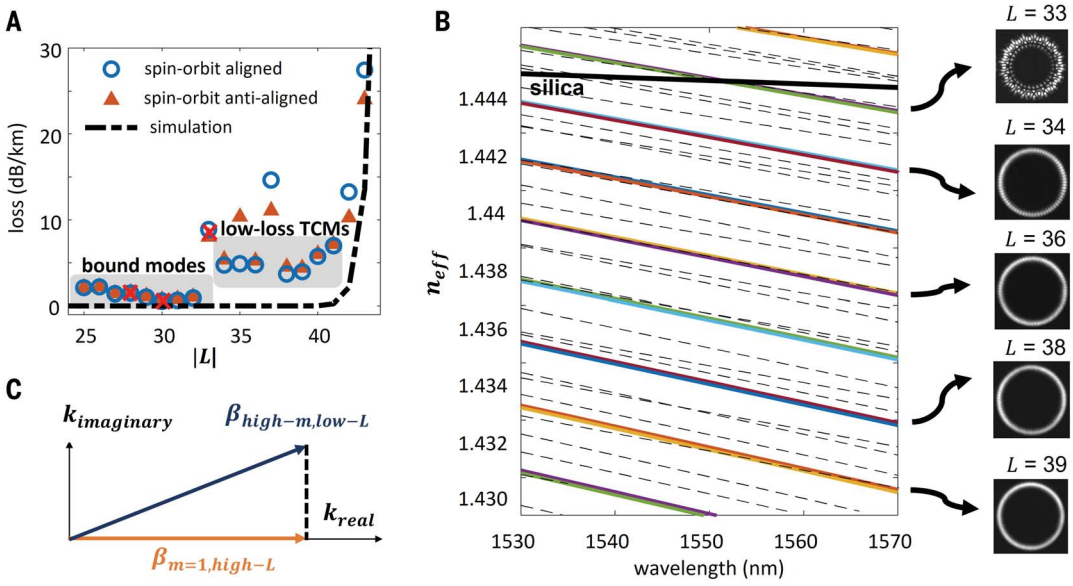


Fig. 2. Guidance beyond cutoff and natural distortion immunity of TCMs. (**A**) Experimental and simulated confinement loss at 1550 nm. (**B**) (Left) n_{eff} versus λ for select modes. Solid colored lines are desired m = 1 modes of different $|L|$; dashed black lines represent undesired high-m modes. The solid black line shows the index of the silica cladding (conventional boundary between TIR bound and cutoff modes). (Right) Experimentally measured output images of LCP modes at 1550 nm out of a 480-m-long ring-core fiber. (**C**) Schematic illustration of the mismatch between the complex β of desired m = 1 and undesired high-m cutoff modes, showing that coupling can be frustrated even when n_{eff} of modes match because their loss differences result in large imaginary wave-vector mismatch.

For light, centrifugal barriers increase the lifetimes of photons in whispering-gallery cavities (*20*, *21*).

Natural immunity to mode mixing

The key benefit of this regime of light transport is not just in the availability of more eigenstates—that result could potentially have been achieved with larger waveguides or higher-index cores for conventional TIR bound modes; rather, it is that topological confinement radically changes the density of states of low-loss propagating modes in multimode fibers, with implications for modal fidelity, as demonstrated below.

Figure 2B shows the simulated n_{eff} versus λ of select modes in the same ring-core fiber. Solid colored lines represent m = 1 modes of various $|L|$, and the dashed black lines represent high-m modes. It is immediately apparent that in a highly multimoded fiber, the different dispersion relationships for modes of different radial and azimuthal orders result in a plethora of degeneracies, making them especially susceptible to mode mixing. This is evident from a transmission experiment (fig. S2A) over a 480-m-long fiber, in which an attempt at sending a signal in a conventional TIR bound mode (L = 33, m = 1) results in a completely distorted output mode image (because it is uncontrollably coupled with the degenerate L = 16, m = 4 mode). Such modal degeneracies are the primary reason behind the current lack of progress in scaling cross-talk–minimized mode counts using conventional TIR bound modes.

By contrast, markedly distinct behavior is evident for TCMs (whose n_{eff} lie below the silica cladding index, depicted as a solid black curve in Fig. 2B). Although the L = 39, m = 1 mode is degenerate with an L = 24, m = 4 mode, the output image still shows a clear single ring. Modal degeneracy does not cause modal distortions in this case because the two modes have orders-of-magnitude-different confinement losses (Fig. 1C), leading to frustration of phase-matched coupling (*22*) (Fig. 2C). As such, all TCMs ($|L| \geq 34$, m = 1) show clean measured modal output images even though some transition TCMs reveal extra total loss (fig. S4). Hence, operation in the TCM regime allows scaling mode count while decreasing the modal density of states that cause unwanted perturbative mixing.

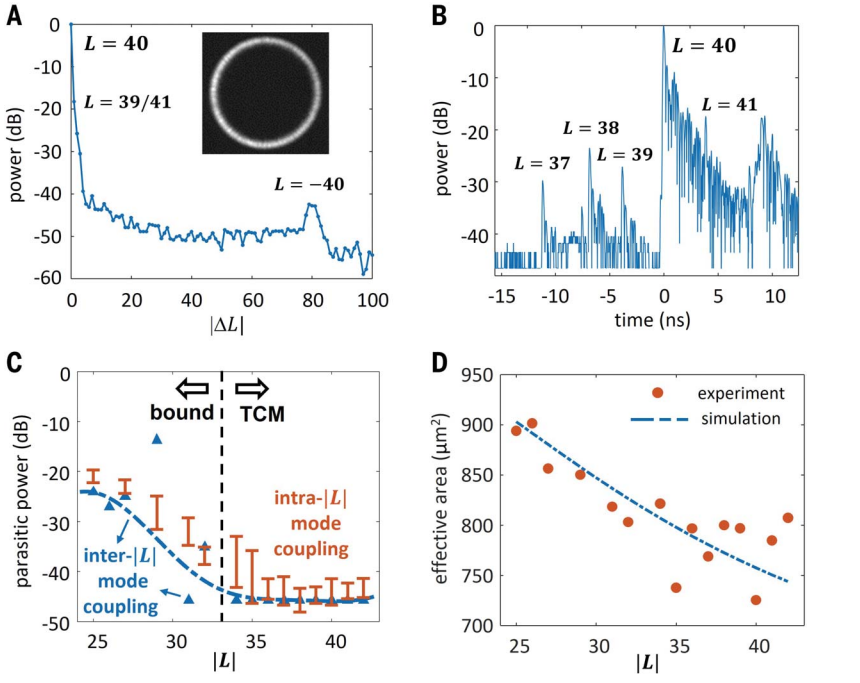


Fig. 3. Modal characterization of the TCM-supporting ring-core fiber at 1550 nm. (**A**) Parasitic mode power spectrum for L = 40, RCP launched mode measured by spatial interferometry. (Inset) Launched mode image at fiber output. (**B**) Time-of-flight trace of L = 40, RCP launched mode. (**C**) Parasitic power for inter-$|L|$ and intra-$|L|$ coupling. (**D**) Measured and simulated A_{eff} at 1550 nm versus $|L|$.

Fiber properties and transmission characteristics

We quantified the cross-talk of desired m = 1 modes at 1550 nm over this 480-m-long fiber using two techniques: (i) spatial interferometry (*23*), which yields power coupled in all modes for a given launched mode, and (ii) impulse-response measurements (*5*), which help disambiguate fundamental cross-talk due to in-fiber coupling from "technical," discrete cross-talk from multiplexers. The latter is of

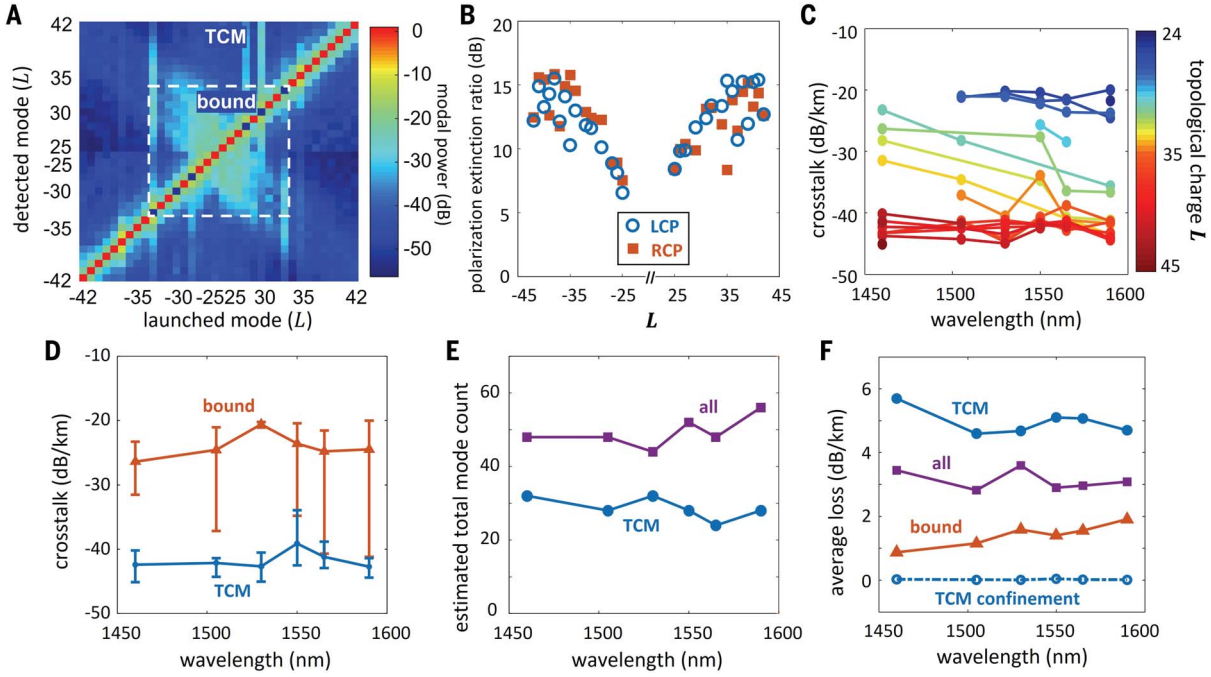


Fig. 4. Information capacity potential of TCM-supporting ring-core fibers. (**A**) Transfer matrices of all the modes with RCP. The red markers along diagonal indicate good modes; deep-blue markers indicate unstable, mixed modes. The box enclosed by the white dashed line denotes transition between TIR bound modes and TCMs. (**B**) PER of all available modes with LCP and RCP. (**C**) Cross-talk versus λ of modes with +|*L*|, RCP in S, C, and L telecommunications bands. (**D**) Cross-talk versus λ for TIR bound modes and TCMs. The color curves are the average cross-talk; error bars indicate the maximum and minimum values. (**E**) Estimated total mode count versus λ for TCMs and the sum of TCMs and TIR bound modes. (**F**) Average loss versus λ of good bound modes, TCMs, and their sum. The dashed curve represents the average simulated confinement loss of TCMs.

minimal concern in our analysis because it could be substantially minimized by using compact, fiber-compatible integrated mode transformers (*24*) in practical applications. Figure 3A shows an exemplary spatial interferometry measurement when the L = +40, RCP mode is launched. Parasitic power at the output primarily resides in two modes (supplementary text): in the L = −40, RCP mode at the −43-dB level, and in modes separated from the launched mode by ΔL = ±1, at the −20-dB level. The former (intra-|L| mode coupling) arises from in-fiber coupling between nearest neighbors (in n_{eff}), whereas the latter (inter-|L| mode coupling) arises from experimental mode-launch imperfections, as well as in-fiber, bend-induced coupling. A representative impulse-response measurement for the same launched mode reveals two distinct features of inter-|L| coupling: discrete coupling resulting in a spike in temporal trace that is due to mode launching, and distributed coupling between this spike from the L = 39 mode and the main peak for the L = 40 mode. The integrated power of this distributed coupling is found to be lower than the noise floor of the measurement (−46 dB). Figure 3C is concerned only with coupling of the fundamental (in-fiber, distributed) kind and plots the parasitic power of these two main sources of mode coupling for launched modes of interest in this fiber (supplementary text). The first finding of interest is that intra-|L| coupling dominates over bend-induced inter-|L| coupling, justifying using the intra-|L| coupling to represent the cross-talk (supplementary text). The second, more notable finding is that TCMs are more resistant to mode mixing, featuring cross-talk < −40 dB/km for almost all TCMs, barring a few outliers primarily at the transition between bound and cutoff modes. Some TCMs reach cross-talk as low as −45 dB/km. By contrast, the cross-talk for conventional TIR bound modes remains ~−20 dB/km, the same as that found in previous investigations of MMFs (*5*). Figure 3D shows that the effective areas (A_{eff}) of these modes, measured from recorded intensity profiles (*12*), are an order of magnitude larger than those for SMFs or MCFs. Furthermore, we find that all these attractive attributes of TCMs are maintained even when the fiber is bent down to radii as small as 6 mm (fig. S5).

To characterize this fiber over kilometer lengths of practical utility in data center applications, we used a Sagnac reflector to equivalently double the length of our 480-m-long fiber (*25*) (fig. S2B). Figure 4A shows the full transfer matrix for all the modes |L| = 25 – 42, RCP (Identical performance for LCP modes is shown in fig. S6.). Except for the outlier |L| = 28, 30, 33 modes (deep-blue square markers along the diagonal indicate TIR bound modes that experience mode mixing), and not counting the outliers in the TCM regime that exhibit high loss (Fig. 2A), we obtained 50 modes (red square markers) that exhibit cross-talk of ~−40 dB/km for TCMs and ~−20 dB/km for conventional TIR bound modes (based on measured power in the antidiagonals). The majority of the power in the two off-diagonals adjacent to the launched mode are from mode-launch imperfections, whereas fundamental in-fiber, bend-induced cross-talk in these modes is immeasurably low. We also measured the polarization extinction ratios (PER) (Fig. 4B) (*12*), which increased with |L| from ~10 to ~15 dB because of the fundamental property of OAM conservation (*26*). Figure 4C shows dominant (intra-|L|) in-fiber cross-talk for measurements repeated for λ ∈ [1460,1590] nm, covering the telecommunications S and C band and part of the L band (as before, outlier modes with excessive mode mixing or loss were excluded). Red and blue curves of various shades denote TCM and TIR bound-mode cross-talk, respectively. The cross-talk of modes depicted by other colors are typically for TIR bound modes that are close to cutoff, and hence modes that resemble TCMs as wavelength increases. Again, the average cross-talk of all TCMs lies between −40 and −45 dB/km, whereas most bound modes appear pinned at a cross-talk of −20 dB/km, regardless of wavelength. Especially apparent is the cross-talk

of TIR bound modes close to cutoff, demonstrating that, as the guidance mechanism for light evolves from TIR to topological confinement, the cross-talk also improves by two orders of magnitude. In addition, Fig. 4, D to F depicts the cross-talk, mode count, and average loss, respectively, demonstrating a 1-km-long MMF with record-high mode counts of ~50, comprising ~20 conventional TIR bound modes with < −20 dB/km cross-talk and ~30 TCMs with cross-talk < −40 dB/km across three telecommunications spectral bands of interest. The measured cross-talk and mode counts are on par with those of MCFs (*27*), and the broad bandwidth promises the compatibility with wavelength-division multiplexing. The 10× larger A_{eff} and chromatic dispersion (fig. S7C), as well as the relatively large differential modal delays (fig. S7, A and B), suggest that transmission in TCMs may be helpful in managing nonlinear signal distortions (*28*, *29*) in telecommunications applications. In addition, a plethora of low-cross-talk modes potentially enables low cross-talk per superpositions of modes as well (*30*), as required for quantum transport, effectively allowing high-dimensional encoding (*2*) and sources through intermodal interaction (*31*). Average losses for TCMs (~5.0 dB/km) and for TIR bound modes (~1.4 dB/km) are, at their current maturity, reasonable for amplifier and inter–data center applications, but because confinement losses are predicted to be much lower, manufacturing optimizations (fig. S3) would make these fibers suitable for substantially longer telecommunications links.

Summary and conclusions

The physics of frustrated coupling in TCMs yields an improvement of over two orders of magnitude in mode-coupling resistance over any MMF demonstrated to date. This fundamental feature has enabled cross-talk and channel counts that are on par with MCFs, while providing additional crucial benefits, such as order-of-magnitude-greater mode areas and extreme bend tolerance. By contrast, TCMs are propagating modes of an MMF, and as such, share other favorable attributes including high spatial efficiency and the prospect of efficient, gain-equalized optical amplification of all channels. Thus, topological confinement represents a regime of light transport that combines the benefits of both MCFs and conventional MMFs, yielding a platform that is not only suitable for scaling the capacity of classical telecommunications and quantum networks but also serves as a distinctive fiber-laser and nonlinear signal-processing medium on account of their large chromatic dispersions and record-large mode areas.

REFERENCES AND NOTES

1. P. J. Winzer, D. T. Neilson, *J. Lightwave Technol.* **35**, 1099–1115 (2017).
2. M. Erhard, M. Krenn, A. Zeilinger, *Nat. Rev. Phys.* **2**, 365–381 (2020).
3. B. J. Puttnam, G. Rademacher, R. S. Luís, *Optica* **8**, 1186–1203 (2021).
4. Y. Jung, S. Alam, D. J. Richardson, S. Ramachandran, K. S. Abedin, in *Optical Fiber Telecommunications VII*, A. E. Willner, Ed. (Academic Press, 2020), pp. 301–333.
5. P. Gregg, P. Kristensen, S. Ramachandran, *Opt. Express* **24**, 18938–18947 (2016).
6. Z. Ma, P. Kristensen, S. Ramachandran, in *2020 Conference on Lasers and Electro-Optics (CLEO)*, San Jose, CA, 10 to 15 May 2020, paper SF1P.2 (Optica Publishing Group, 2020).
7. R. Sammut, A. W. Snyder, *Appl. Opt.* **15**, 477–482 (1976).
8. Z. Ma, S. Ramachandran, *Nanophotonics* **10**, 209–224 (2020).
9. A. K. Ghatak, K. Thyagarajan, *Introduction to Fiber Optics* (Cambridge Univ. Press, 1998).
10. J. P. Berenger, *J. Comput. Phys.* **114**, 185–200 (1994).
11. Z. Ma, P. Kristensen, S. Ramachandran, in *2021 Conference on Lasers and Electro-Optics (CLEO)*, San Jose, CA, 9 to 14 May 2021, paper SM1F.4 (2021).
12. Materials and methods are available as supplementary materials.
13. Z. Ma, P. Kristensen, S. Ramachandran, in *2021 European Conference on Optical Communication (ECOC)*, Bordeaux, France, 13 to 16 September 2021, paper TU4A.3 (2021).
14. M. E. Lines, *J. Appl. Phys.* **55**, 4052–4057 (1984).
15. M. Wandel, "Attenuation in silica-based optical fibers," thesis, Technical Univ. of Denmark (2005).
16. M. Wandel, P. Kristensen, in *Fiber Based Dispersion Compensation*, S. Ramachandran, Ed. (Springer, 2007), pp. 7–42.
17. R. R. Freeman, P. H. Bucksbaum, *J. Phys. At. Mol. Opt. Phys.* **24**, 325–347 (1991).
18. C. Chin, R. Grimm, P. Julienne, E. Tiesinga, *Rev. Mod. Phys.* **82**, 1225–1286 (2010).
19. J. M. Blatt, V. F. Weisskopf, *Theoretical Nuclear Physics* (Wiley, 1952).
20. A. Chiasera *et al.*, *Laser Photonics Rev.* **4**, 457–482 (2010).
21. M. Sumetsky, *Phys. Rev. Lett.* **111**, 163901 (2013).
22. D. B. Stegall, T. Erdogan, *IEEE Photonics Technol. Lett.* **11**, 343–345 (1999).
23. N. Bozinovic, S. Golowich, P. Kristensen, S. Ramachandran, *Opt. Lett.* **37**, 2451–2453 (2012).
24. S. Bade *et al.*, in *Optical Fiber Communications Conference 2018 (OFC)*, San Diego, CA, 11 to 15 March 2018, paper Th4B.3 (Optica Publishing Group, 2018).
25. Z. Ma, M. W. Khalid, S. Ramachandran, in *2022 Conference on Lasers and Electro-Optics (CLEO)*, San Jose, CA, 15 to 20 May 2022, paper STu4P.3 (Optica Publishing Group, 2022).
26. P. Gregg, P. Kristensen, S. Ramachandran, *Optica* **2**, 267–270 (2015).
27. Y. Sasaki, in *2017 Optical Fiber Communications Conference (OFC)*, Los Angeles, CA, 19 to 23 March 2017, paper Th1H.2 (2017).
28. A. D. Ellis, N. Mac Suibhne, F. C. Gunning, S. Sygletos, *Opt. Express* **21**, 22834–22846 (2013).
29. G. Rademacher, K. Petermann, *J. Lightwave Technol.* **34**, 2280–2287 (2016).
30. A. P. Greenberg, G. Prabhakar, S. Ramachandran, *Nat. Commun.* **11**, 5257 (2020).
31. X. Liu, D. B. Kim, V. O. Lorenz, S. Ramachandran, in *Quantum 2.0*, Boston, MA, 13 to 16 June 2022, paper QTh4B.1 (Optica Publishing Group, 2022).

ACKNOWLEDGMENTS

We thank A. P. McCall and M. W. Khalid for help with instrumentation in the experiments. **Funding:** This study was supported by the Vannevar Bush Faculty Fellowship (N00014-19-1-2632), the Office of Naval Research (ONR) Multi University Research Initiative (MURI) (Grant N00014-20-1-2450), Brookhaven National Labs (Contract 354281), and the Air Force Office of Scientific Research (AFOSR) BRI program (Grant FA9550-14-1-0165). **Author contributions:** Conceptualization: Z.M. and S.R. Methodology: Z.M., P.K., and S.R. Investigation: Z.M. Visualization: Z.M. Funding acquisition: S.R. Project administration: S.R. Supervision: S.R. Writing – Z.M., P.K., and S.R. **Competing interests:** S.R. and Z.M. are inventors on a patent (US11506841B2) submitted by Boston University that covers the topological guidance of light described in this paper. P.K. declares no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add1874
Materials and Methods
Supplementary Text
Figs. S1 to S7
Reference (*32*)

Submitted 27 May 2022; resubmitted 23 December 2022
Accepted 22 March 2023
10.1126/science.add1874

CANCER EVOLUTION

The evolution of two transmissible cancers in Tasmanian devils

Maximilian R. Stammnitz[1]†, Kevin Gori[1], Young Mi Kwon[1], Edward Harry[2], Fergal J. Martin[3], Konstantinos Billis[3], Yuanyuan Cheng[4], Adrian Baez-Ortega[1], William Chow[2], Sebastien Comte[5,6], Hannes Eggertsson[7], Samantha Fox[8,9], Rodrigo Hamede[5,10], Menna Jones[5], Billie Lazenby[8], Sarah Peck[8], Ruth Pye[11], Michael A. Quail[2], Kate Swift[12], Jinhong Wang[1], Jonathan Wood[2], Kerstin Howe[2], Michael R. Stratton[2], Zemin Ning[2], Elizabeth P. Murchison[1]*

Tasmanian devils have spawned two transmissible cancer lineages, named devil facial tumor 1 (DFT1) and devil facial tumor 2 (DFT2). We investigated the genetic diversity and evolution of these clones by analyzing 78 DFT1 and 41 DFT2 genomes relative to a newly assembled, chromosome-level reference. Time-resolved phylogenetic trees reveal that DFT1 first emerged in 1986 (1982 to 1989) and DFT2 in 2011 (2009 to 2012). Subclone analysis documents transmission of heterogeneous cell populations. DFT2 has faster mutation rates than DFT1 across all variant classes, including substitutions, indels, rearrangements, transposable element insertions, and copy number alterations, and we identify a hypermutated DFT1 lineage with defective DNA mismatch repair. Several loci show plausible evidence of positive selection in DFT1 or DFT2, including loss of chromosome Y and inactivation of *MGA*, but none are common to both cancers. This study reveals the parallel long-term evolution of two transmissible cancers inhabiting a common niche in Tasmanian devils.

Transmissible cancers are contagious somatic cell lineages that spread through populations by the physical transfer of living cancer cells. Although few such diseases are known in nature, Tasmanian devils (*Sarcophilus harrisii*), marsupial carnivores endemic to the Australian island of Tasmania, host at least two transmissible cancer clones. These cancers, known as devil facial tumor 1 (DFT1) and devil facial tumor 2 (DFT2), both primarily cause malignant facial and oral tumors that are spread by biting (Fig. 1A) (*1*–*3*). DFT1 was first observed in 1996 in northeastern Tasmania and has subsequently spread widely (*4*, *5*); DFT2, however, was discovered in 2014 on the D'Entrecasteaux Channel Peninsula in Tasmania's southeast and is believed to remain confined to this area (*3*, *6*, *7*). Both DFT1 and DFT2 are usually fatal, and rapid Tasmanian devil population declines associated with DFT1 have led to concern for conservation of the species (*4*, *5*, *8*).

[1]Transmissible Cancer Group, Department of Veterinary Medicine, University of Cambridge, Cambridge, UK. [2]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK. [3]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK. [4]School of Life and Environmental Sciences, University of Sydney, Sydney, NSW, Australia. [5]School of Natural Sciences, University of Tasmania, Hobart, TAS, Australia. [6]Vertebrate Pest Research Unit, NSW Department of Primary Industries, Orange, NSW, Australia. [7]deCODE Genetics Inc., Reykjavik, Iceland. [8]Save the Tasmanian Devil Program, Tasmanian Department of Natural Resources and Environment, Hobart, TAS, Australia. [9]Toledo Zoo, Toledo, OH, USA. [10]CANCEV, Centre de Recherches Ecologiques et Evolutives sur le Cancer, Montpellier, France. [11]Menzies Institute for Medical Research, University of Tasmania, Hobart, TAS, Australia. [12]Mount Pleasant Laboratories, Tasmanian Department of Natural Resources and Environment, Prospect, TAS, Australia.
***Corresponding author. Email: epm27@cam.ac.uk**
†Present address: Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Barcelona, Spain.

The emergence of two transmissible cancers in Tasmanian devils suggests that the species is particularly susceptible to this type of disease. Indeed, DFT1 and DFT2 appear to be independent occurrences of the same pathological process, and their comparison may illuminate the constraints of the biological niche that they inhabit. DFT1 and DFT2 are both undifferentiated Schwann cell cancers with similar dependence on receptor tyrosine kinase signaling (*9*–*12*). DFT1 first arose from the cells of a female founder devil and equally affects male and female devil hosts (*2*, *13*–*15*); DFT2, however, originated from a male devil and shows preference for male hosts, perhaps because of immunogenicity of chromosome Y–derived antigens in female hosts (*3*, *7*, *10*). Both cancers escape the allogeneic immune system, and, in DFT1, this is mediated by transcriptional repression of major histocompatibility complex (MHC) class I genes (*16*). In DFT2, however, cell surface MHC class I molecules are usually detectable, and the high similarity between expressed tumor and host MHC class I alleles may underlie the lack of immune rejection (*17*). The genomes of DFT1 and DFT2 show comparable mutational patterns, but no common positively selected driver mutations have been detected (*10*). Furthermore, whereas DFT1 has split into several spatially defined sublineages during its spread through Tasmania (*18*), little is known about the clonal diversity of DFT2.

In addition to their importance as threats to animal health and their intrinsic interest as unusual pathogens, transmissible cancers provide an opportunity to study how mutations in cancer accumulate with time. Most human cancer studies involve the analysis of tumor biopsies collected either at a single session or at time points separated by short inter
The long-term survival of DFT1 and DFT2 p
mits repeated sampling of the same cancer lineages through decades, which enables direct investigation of variation in mutation rates, together with those of their constitutive mutational signatures, within and between clones.

Here, we describe high-coverage whole-genome sequences of 78 DFT1 and 41 DFT2 tumors, as well as that of a single nontransmissible carcinoma and a panel of 80 normal Tasmanian devil genomes, that were analyzed relative to a newly assembled, highly contiguous Tasmanian devil reference genome. By capturing the somatic genetic diversity present within the DFT1 and DFT2 lineages, our goal was to understand the dynamics of these diseases' emergence and spread, to estimate their mutation rates, and to characterize their long-term patterns of evolution. By intersecting findings from different Tasmanian devil cancers, we identify genomic events that underpin transmissible cancer in this species. Our analysis provides detailed insight into the evolution and diversification of two parallel cancer clones that have survived in a transmissible niche.

An improved reference genome for the Tasmanian devil

Previous Tasmanian devil genome assemblies were highly fragmented (*13*, *19*, *20*). To produce an improved genome assembly for the species, we extracted high molecular weight DNA from the female fibroblast cell line used in an earlier assembly (*13*). We sequenced this to 76-fold and 12-fold coverage using long-read (fragment N50: 9.05 kilobases, kb) and ultralong-read (N50: 57.13 kb) sequencing technology (*21*). In addition, DNA was analyzed with optical mapping, linked-read sequencing, and high-dimension conformation capture (Hi-C). A reference genome assembly, mSarHar1.11, was generated by combining these data (Table 1, table S1, and fig. S1). Notably, 99.8% of bases were placed on one of seven scaffolds, which correspond to the six devil autosomes and chromosome X. Genome annotation was performed with the Ensembl gene annotation pipeline (*22*), guided by a newly sequenced Tasmanian devil multitissue transcriptome atlas, and yielded 19,228 protein-coding gene models (table S1).

DFT1 and DFT2 phylogenies

To investigate genetic variation within Tasmanian devil transmissible cancers, we sequenced the whole genomes of 63 DFT1s and 39 DFT2s (Fig. 1A) to a median depth of 83× and analyzed these alongside 15 DFT1 and 2 DFT2 publicly available genomes (table S2). The DFT1s were primarily selected to capture genetic and spatiotemporal diversity in this clone (Fig. 1B and table S2). These included representatives of the six major clades (A1,

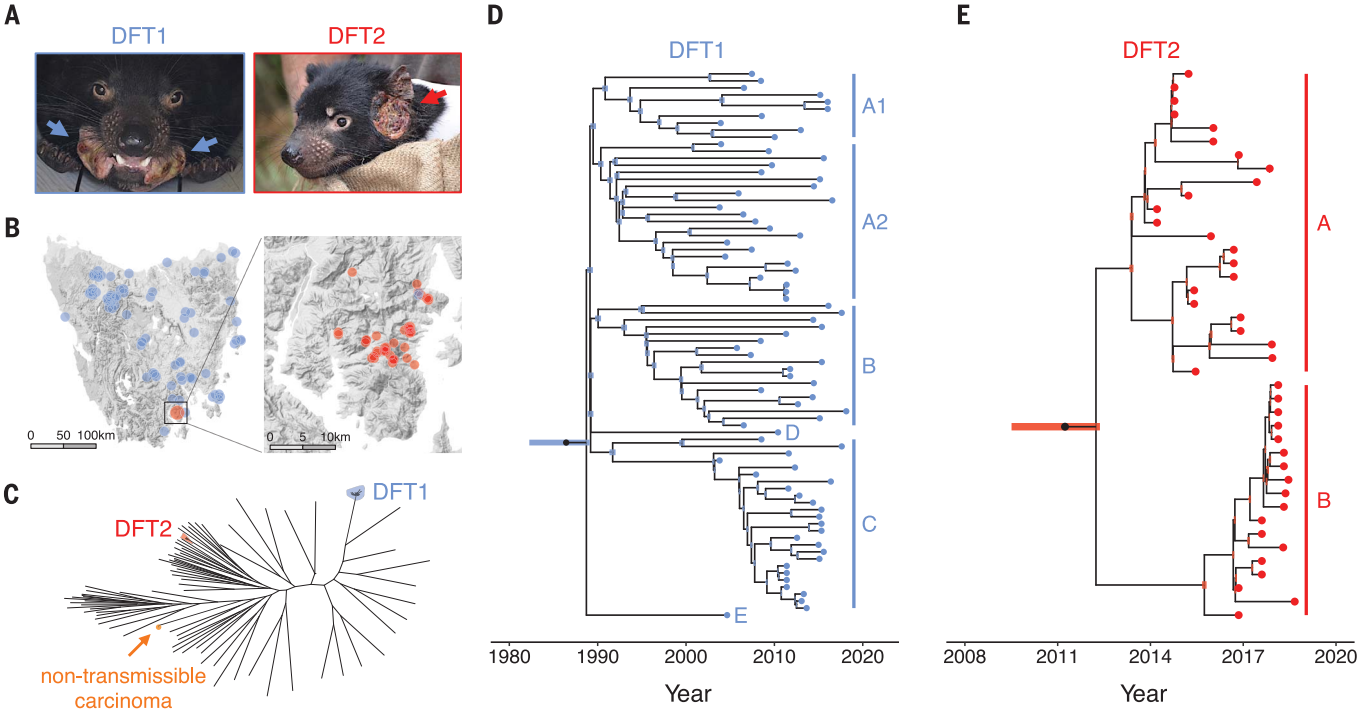


Fig. 1. DFT1 and DFT2 phylogenies. (**A**) Representative photographs of animals infected with DFT1 and DFT2. (**B**) Sampling locations of 78 DFT1 and 41 DFT2 tumors included in the study. (**C**) Maximum likelihood phylogenetic tree constructed with 104,799 somatic and 1,070,436 germline substitutions from 38 DFT1s, 12 DFT2s, a single nontransmissible carcinoma, and 79 Tasmanian devils; only the subset of DFT1s and DFT2s with tumor purity ≥75% were included. Black unlabeled tips represent Tasmanian devils, and shaded tips represent those belonging to DFT1, DFT2, or the nontransmissible carcinoma. Branch lengths are uninformative. High-resolution labeled tree with bootstrap support is available in fig. S2. (**D** and **E**) Time-resolved phylogenetic trees for DFT1 and DFT2 constructed with 171,283 and 21,252 somatic substitution mutations, respectively. Tumor clades are labeled (A1, A2, B, C, D, and E in DFT1; A and B in DFT2). Bars at internal nodes represent 95% Bayesian credible intervals around date estimates. Bars at root nodes represent 95% Bayesian credible intervals around date estimates and incorporate uncertainty in somatic or germline assignment of substitutions shared by all tumors within a clone and absent from all normal Tasmanian devils. Dating is based on tumor sampling dates and does not account for the pretransmission interval, the offset between date of clone emergence, and date of sampling; this is of relevance because bulk tissue sequencing captures only clonal mutations or those present in sizeable subclones (*55*). High-resolution labeled DFT1 and DFT2 trees with node posterior probability are available in figs. S3 and S5.

A2, B, C, D, and E) (*18*) and were collected from 38 locations between 2003 and 2018. For DFT2, we sequenced all available tumors sampled between 2014 and 2018, which all occurred within DFT2's known range on the D'Entrecasteaux Channel Peninsula (Fig. 1B). Some subsets of DFT1 and DFT2 tumors were derived from the same individual hosts. These included sets of matched primary facial tumors and internal metastases as well as samples from distinct facial or body tumors occurring in single hosts (table S2). In addition, we sequenced a nontransmissible anal sac carcinoma sampled from a captive Tasmanian devil and analyzed genomes from 80 normal Tasmanian devils, including matched hosts (71 newly sequenced, 9 publicly available) (table S2).

Single-base substitutions were called in each sample, and normal Tasmanian devil genomes were used to identify and exclude germline substitutions from tumor sequences. This yielded 205,890, 23,152, and 5764 somatic substitutions in DFT1, DFT2, and the nontransmissible anal sac carcinoma, respectively, as well as 1,458,776 germline variants (table S3). Analysis of the latter revealed a median of 0.132 heterozygous sites per kilobase (range 0.083 to 0.153) in the sampled population of Tasmanian devils, with the DFT1 and DFT2 founder devils both falling within this range (table S3).

We confirmed the independent clonal origins of DFT1 and DFT2 by constructing a maximum likelihood tree using substitutions from both tumor and normal samples. As expected, DFT1 and DFT2 tumors each clustered into distinct groups whose positions relative to normal animals are consistent with the notion that these clones' founder devils originated in northeastern Tasmania (DFT1) or on the D'Entrecasteaux Channel Peninsula (DFT2) (Fig. 1C and fig. S2) (*3*, *4*, *10*).

Table 1. mSarHar1.11 Tasmanian devil reference genome assembly and annotation metrics.

Metric	Value
Contigs (N50)	445 (63.34 Mb)
Largest contig	195.48 Mb
Chromosomal scaffolds (N50)	7 (611.35 Mb)
Largest chromosomal scaffold	716.41 Mb
Repeat-masked genome	44.3%
Coding genes	19,228
Noncoding genes	4336

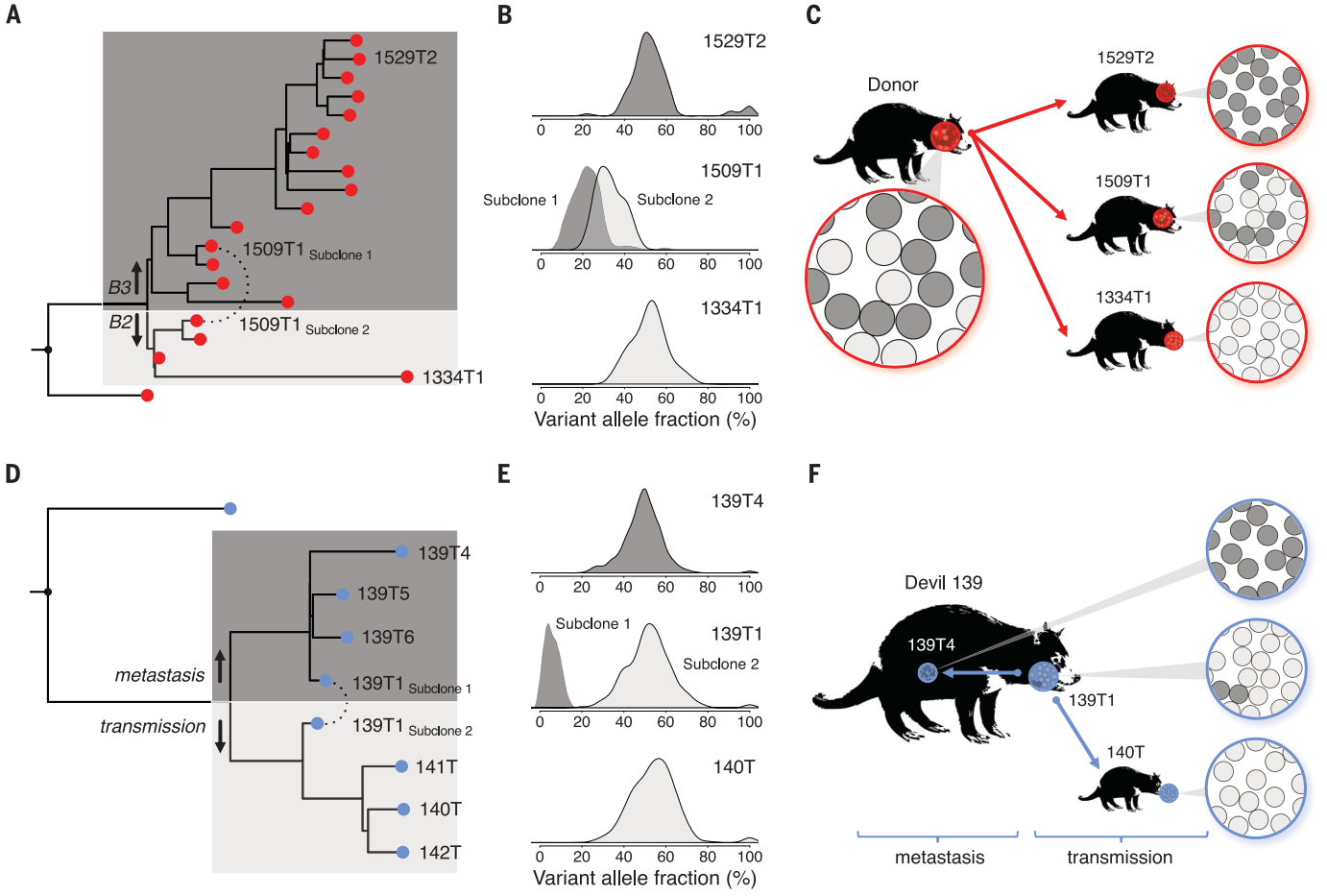


Fig. 2. Intratumor genetic heterogeneity in DFT1 and DFT2. (**A** to **C**) Example of heterogeneous cell transmission in DFT2. 1509T1 is a DFT2 clade B tumor composed of two detectable subclonal cell populations, $1509T1_{Subclone1}$ and $1509T1_{Subclone2}$. (A) Computational separation of $1509T1_{Subclone1}$ and $1509T1_{Subclone2}$ and inclusion on a phylogenetic tree revealed subclone membership of distinct clade subgroups, DFT2-B3 and DFT2-B2. Branch lengths are proportional to number of substitution variants. (B) Variant allele distribution of 1509T1, together with those of representative DFT2-B3 (1529T2) and DFT2-B2 (1334T1) tumors; only variants occurring after the split between DFT2-B3 (dark gray) and DFT2-B2 (light gray) are included. Because tumors are diploid, most mutations occur in the heterozygous state and would be expected to be found at 50% proportion. (C) Model illustrating transmission of DFT2 from an earlier donor devil, which carried both DFT2-B3 (dark gray) and DFT2-B2 (light gray) cells, to recipient devils. Recipient tumors are composed either of clonal populations of DFT2-B3 (upper, dark gray, 1529T2), clonal populations of DFT2-B2 (lower, light gray, 1334T1) or a subclonal mixture of DFT2-B3 and DFT2-B2 (middle, mixture of light gray and dark gray, 1509T1). Arrows do not necessarily represent direct transmission. (**D** to **F**) Example of differential transmission and metastasis of subclones in DFT1. 139T1 is a DFT1 facial tumor composed of two detectable subclonal cell populations, $139T1_{Subclone1}$ and $139T1_{Subclone2}$, represented by dark gray and light gray shading, respectively. (D) $139T1_{Subclone1}$ clusters phylogenetically with a set of internal metastases sampled from the same individual (139T4, 139T5, 139T6), and $139T1_{Subclone2}$ clusters with facial tumors sampled from three devils (140T, 141T, 142T) involved in a DFT transmission chain (fig. S6). Branch lengths are proportional to number of substitution variants. (E) Variant allele distribution of 139T1, together with those of a representative metastasis involving the same host (139T4) and a representative tumor secondary to transmission involving a different host (140T); only variants occurring after the split between the metastases (dark gray) and transmission (light gray) are included. Because tumors are diploid, most mutations occur in the heterozygous state and would be expected to be found at 50% proportion. (F) Model illustrating differential spread of subclones. $139T1_{Subclone1}$ (dark gray) and $139T1_{Subclone2}$ (light gray) are both represented in tumor 139T1. Cells belonging to $139T1_{Subclone1}$ seeded internal metastases (represented by 139T4), whereas cells from $139T1_{Subclone2}$ were transmitted onwards to recipient devils (represented by 140T). Further details available in fig. S6. The Tasmanian devil silhouette used throughout this figure is adapted from Nilsson *et al*. (*56*).

Time-resolved phylogenetic trees were generated for DFT1 and DFT2 with substitution mutation rates inferred by using tumor sampling dates (Fig. 1, D and E). By assuming a constant mutation rate, DFT1 was estimated to have arisen in 1986 (95% Bayesian credible interval 1982 to 1989), which implies a substantial delay from its emergence until its first observation in 1996 (Fig. 1D and fig. S3) (*4*). The DFT1 tree showed the expected arrangement of the six identified tumor clades (*18*) and revealed that these split from one another very early in DFT1 evolution in a rapid diversification event that almost certainly involved a single tumor donor (fig. S4). DFT2, however, is estimated to have first emerged in 2011 (95% Bayesian credible interval 2009 to 2012). It subsequently split into two major sympatric groups, which we term DFT2 clades A and B (Fig. 1E and fig. S5). The potential for individual devils to be coinfected with distinct lineages of DFT1 (*18*), DFT2, or both (*10*) is apparent. The presence of true or near-polytomies evident on both the DFT1 and DFT2 phylogenetic trees, which are defined by very short internal branches (Fig. 1, D and E), suggests that it may not be uncommon for infectious devils to transmit

their tumor to more than two secondary hosts. Such events may, however, be enriched at early time points in the trees because of survivorship bias (*23*).

Intratumor genetic heterogeneity in DFT1 and DFT2

Bulk sequencing of tumor tissue, as performed here, will capture only clonal mutations or those present in sizeable subclones. When present, however, the distribution of subclones among tumors could be informative about the clonality of transmission in DFT1 and DFT2.

We screened tumors for subclones by searching for mutation populations that showed unexpected allele fractions. One DFT2 tumor, 1509T1, was found to be composed of two subclonal cell populations that were represented at ~60 and 40% frequency, respectively. We computationally isolated these subclones, and inspection of their positions on the DFT2 phylogenetic tree revealed that they belonged to separate DFT2 clade B sublineages, which we term DFT2-B2 and DFT2-B3 (Fig. 2, A and B). Indeed, mutations defining each subclone were observed clonally in related contemporaneous tumors from different hosts. These data are compatible with a model in which an earlier donor tumor contained cells belonging to both DFT2-B2 and DFT2-B3; onward transmission founded descendent tumors composed of either DFT2-B2 or DFT2-B3 cells or, in the case of 1509T1, a mixture of both DFT2-B2 and DFT2-B3 cells (Fig. 2C).

We similarly investigated intratumor heterogeneity in DFT1 using a closely related set of tumors that were part of a series of direct transmission events (Fig. 2D and fig. S6). This case involved a female devil with a facial tumor and several metastases. Cells were transmitted from this female's facial tumor to her unweaned male offspring, who, once weaned, further transmitted his tumor to two additional hosts while the group was housed together in captivity (fig. S6). The index female's facial tumor was composed of two detectable subclones at ~90 and 10% proportions, which clustered with the tumor of the offspring and with her metastases, respectively (Fig. 2, D and E). This suggests that two distinct cell lineages, both represented within the index facial tumor, differentially contributed to metastatic dissemination and onward transmission (Fig. 2F).

These case studies hint at the genetic heterogeneity present within individual DFT tumors and, in the DFT2 example, imply that this diversity can be maintained across transmission bottlenecks. Thus, at least in some cases, DFT tumors are seeded by more than one cell.

DFT1 and DFT2 substitutions and indels

To obtain an overview of the mutational processes operating in Tasmanian devil cancers, we inspected each tumor's mutational spectrum, a representation of the distribution of mutations across the six base substitution classes, displayed together with their immediate 5′ and 3′ base contexts. Such spectra can be decomposed into their constituent mutational signatures, patterns of co-occurring mutation types that reflect the activities of underlying endogenous or exogenous mutational processes (*24*). As expected, DFT1 and DFT2, as well as the single nontransmissible anal sac carcinoma, showed evidence for the presence of two known mutational signatures, single-base substitution signatures 1 (SBS1) and 5 (SBS5), which are found almost universally in human cancer (*25*) and have been described previously in Tasmanian devil tumors (Fig. 3A and fig. S7) (*10*). SBS1 is characterized by C>T mutations at CpG dinucleotide contexts and is believed to primarily arise because of spontaneous deamination of 5′-methylcytosine (*24*). SBS5, however, shows little base specificity, and its etiology is poorly understood (*25*, *26*). Consistent with a previous report (*10*), no evidence of ultraviolet light mutagenesis was detectable in DFT1 or DFT2 mutation patterns, which indicates that the cells that transmit DFT are not usually exposed to sunlight. Patterns of short insertions and deletions (indels) in DFT1 and DFT2 revealed imprints of indel signatures 1 (ID1) and 2 (ID2) in both cancers (*25*), although ID1 dominated in DFT1 (66% ID1, 34% ID2) whereas ID1 and ID2 were present at similar proportions in DFT2 (47% ID1, 53% ID2; Fig. 3B and fig. S8). These signatures are defined by the accumulation of insertions (ID1) or deletions (ID2) of single thymine or adenine bases that occur at mononucleotide tracts and arise through polymerase slippage that involves the nascent (ID1) or the template (ID2) DNA strand (*25*).

Mutational signatures SBS1, SBS5, ID1, and ID2 all present clock-like properties in human cells, defined by linear correlation with donor age (*25*, *27*, *28*). Their rates vary widely among tissues, and, whereas the rates of SBS1, ID1, and ID2 correlate with one another and are believed to reflect the number of mitoses that a cell has experienced, SBS5 rate is independent of these (*25*). We characterized overall substitution and indel rates, as well as rates of SBS1, SBS5, ID1, and ID2 in DFT1 and DFT2 by regressing the number of mutations attributable to each signature in each tumor against sampling date (Fig. 3, C to F). These analyses revealed that overall substitution and indel mutation rates in DFT2 were 3.0 and 3.9 times higher, respectively, than those of DFT1 (Table 2). SBS1 and ID1 accumulate only moderately faster in DFT2 than in DFT1, but rates of SBS5 and ID2 are both considerably higher in DFT2 than in DFT1 (Table 2; Fig. 3, E and F; and table S3).

The relationship between substitution burden and sampling date is linear in both DFT1 and DFT2. Nevertheless, a group of DFT1 tumors can be observed with fewer substitutions attributable to both SBS1 and SBS5 than expected (Fig. 3G). These tumors belong to a single branch of the phylogenetic tree, clade C2/3, which corresponds to the group of clade C tumors sampled in northwest Tasmania (Fig. 3G). The mutation rate we inferred when considering only these tumors (179 mutations per year, 95% confidence interval 131 to 227) is similar to that of the remaining

Table 2. Summary of DFT1 and DFT2 mutation rates. Mutation rates were estimated by using linear regression except for Substitutions (BEAST), which was estimated by using a Bayesian phylogenetic approach (*57*). Rates represent mutation count per genome per year. These can be converted to mutation count per nucleotide per genome per year by dividing by callable genome size (2,983,750,195 nucleotides). Rate ranges represent 95% confidence interval of the linear fit except for Substitutions (BEAST), for which range represents 95% Bayesian credible interval. DFT1 hypermutator clade E was excluded from substitution and indel rate calculations. Ratio ranges represent error-propagated 95% confidence intervals. Level of significance of *F* test for linear fit is shown, ratios of mutation classes which did not show significant linear fits are not displayed. **F* test $P < 0.01$; ***F* test $P < 1 \times 10^{-8}$; ns, not significant.

Mutation class	DFT1 rate, per year	DFT2 rate, per year	DFT2:DFT1 rate ratio
Substitutions (BEAST)	215.5 [212.2 to 218.5]	516.7 [505.8 to 527.7]	2.397 [2.336 to 2.457]
Substitutions	163.2 [119.6 to 206.8]**	496.3 [436.0 to 556.5]**	3.041 [2.753 to 3.329]
SBS1 substitutions	12.6 [10.5 to 14.7]**	17.0 [13.1 to 20.9]**	1.351 [1.074 to 1.629]
SBS5 substitutions	150.6 [108.8 to 192.4]**	479.2 [422.1 to 536.4]**	3.182 [2.886 to 3.478]
Indels	22.2 [12.6 to 31.8]**	86.5 [73.0 to 100.0]**	3.897 [3.445 to 4.348]
ID1 indels	9.8 [3.5 to 16.1]*	13.0 [6.7 to 19.3]*	1.326 [0.539 to 2.113]
ID2 indels	1.4 [−0.43 to 3.25], ns	27.1 [20.5 to 33.6]**	–
LINE-1 insertions	−0.2 [−0.35 to 0.04], ns	24.1 [18.5 to 29.6]**	–
Rearrangement events	1.1 [0.3 to 1.8]*	3.6 [2.7 to 4.5]**	3.394 [2.646 to 4.142]
Copy number events	0.7 [0.2 to 1.1]*	5.6 [4.7 to 6.5]**	8.391 [7.701 to 9.080]

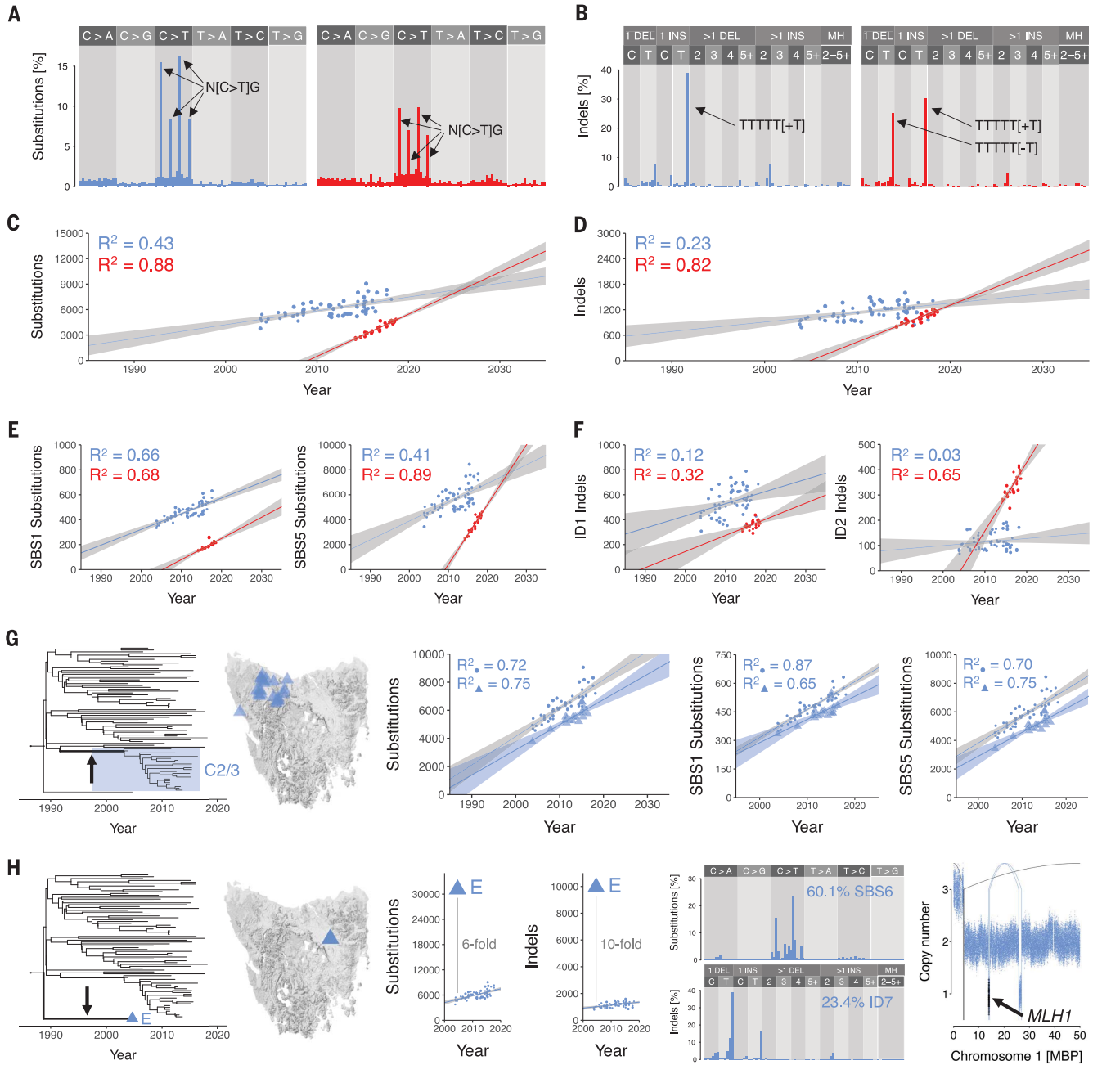


Fig. 3. DFT1 and DFT2 substitutions and indels. (**A** and **B**) Mutational spectra for somatic substitutions (A) and indels (B) in DFT1 (blue, *n* = 176,428 substitutions, *n* = 22,479 indels; variants specific to DFT1 clade E were excluded) and DFT2 (red, *n* = 23,152 substitutions, *n* = 4054 indels). Fully labeled plots available in figs. S7 and S8. (**C** and **D**) Rate of accumulation of substitutions (C) and indels (D) in DFT1 excluding clade E (blue) and DFT2 (red). Each point represents a tumor, plotted by sampling date. Lines represent linear regression, gray shading 95% confidence interval. R^2, coefficient of determination. (**E**) Rate of accumulation of substitution mutations corresponding to mutational signatures SBS1 (left) and SBS5 (right) in DFT1 excluding clade E (blue) and DFT2 (red). Each point represents a tumor, plotted by sampling date. Lines represent linear regression, gray shading 95% confidence interval. (**F**) Rate of accumulation of substitution mutations corresponding to mutational signatures ID1 (left) and ID2 (right) in DFT1 excluding clade E (blue) and DFT2 (red). Each point represents a tumor, plotted by sampling date. Lines represent linear regression, gray shading 95% confidence interval. (**G**) Transient reduction in DFT1 substitution mutation rate occurring within phylogenetic branch leading to DFT1 clade C2/3 (arrow and shading, left); tumors in DFT1 clade C2/3 occur in Tasmania's northwest (map). Overall mutation rate reduction (center) is attributable to both mutational signatures SBS1 and SBS5 (second from right, right); each point represents a tumor, plotted by sampling date, with clade C2/3 tumors represented as triangles. Lines represent linear regression, gray shading 95% confidence interval. (**H**) The single representative of DFT1 clade E, sampled in northeast Tasmania (tree, map) has elevated numbers of substitution and indel mutations; central plots show numbers of substitutions (left) and indels (right) in all DFT1 tumors plotted by sampling date, with clade E tumor represented by triangle. Clade E has distinctive substitution and indel mutational spectra, with at least 60% of the spectrum explained by signature SBS6 (second from right; fully labeled plots available in fig. S9). Clade E carries a deletion encompassing the *MLH1* locus (right; dots represent normalized read coverage within 1-kb genomic windows, with windows including *MLH1* shaded in black; MBP, mega–base pairs; connecting arcs represent rearrangements). High-resolution images and source data available in figs. S7 to S9 and table S3.

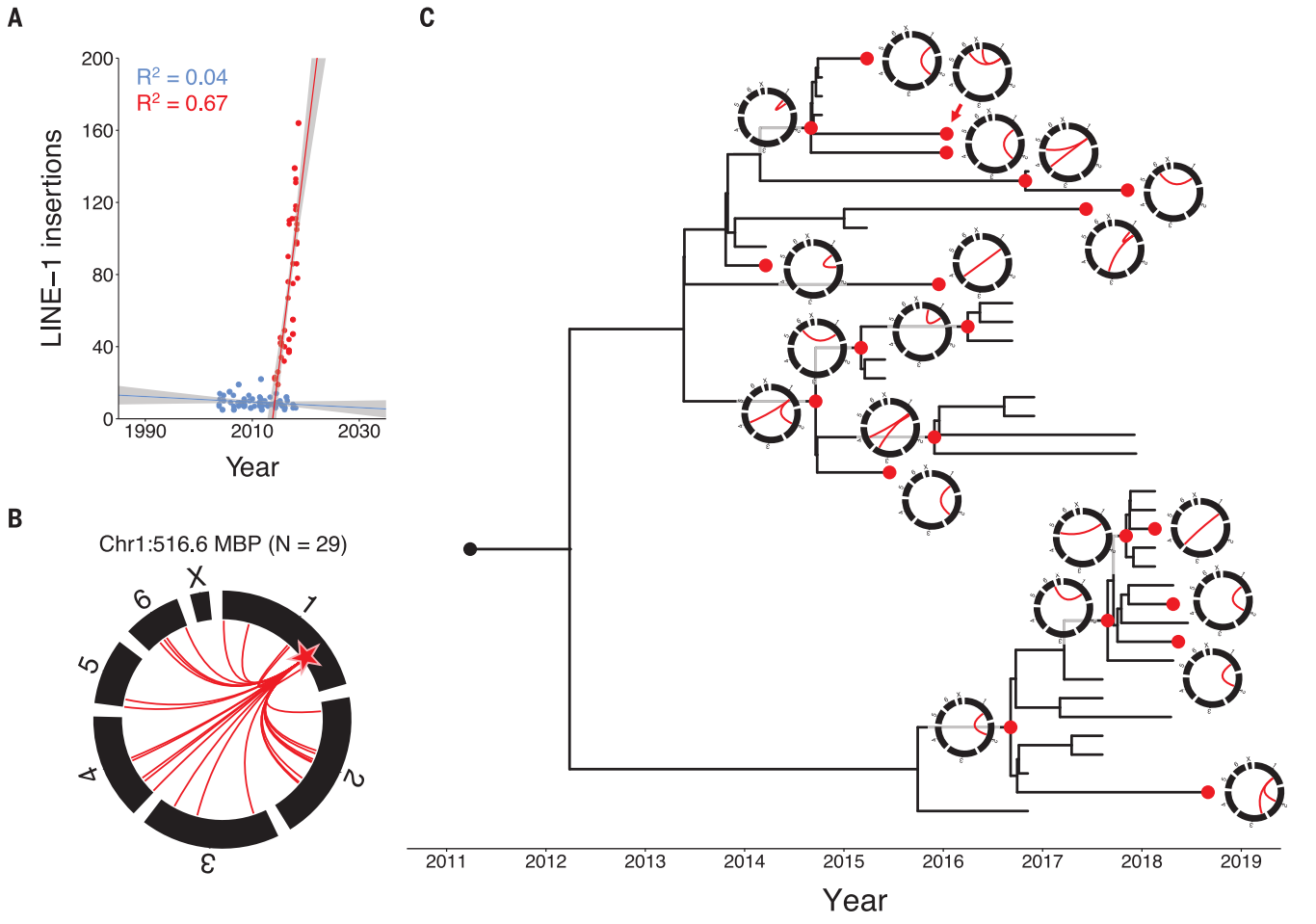


Fig. 4. LINE-1 transposable element activity in DFT1 and DFT2. (**A**) Rate of LINE-1 insertion accumulation in DFT1 (blue) and DFT2 (red). Each point represents a tumor, plotted by sampling date. Lines represent linear regression, gray shading 95% confidence interval. (**B**) DFT2 3′ transduction activity of a LINE-1 source element at chromosome 1:516.6 megabases (Mb) (star). In the circos plot, chromosomes are represented by black bars, and red arcs connect source element to 3′ transduction integration site. (**C**) DFT2 phylogenetic tree as shown in Fig. 1E with circos plots illustrating temporal activity of the LINE-1 source element located at chromosome 1:516.6 Mb. Nodes corresponding to each circos plot are represented in red. Source data available in table S4.

DFT1 tumors (202 mutations per year, 95% confidence interval 166 to 238); however, there are ~1200 fewer mutations genome-wide in the overall clade C2/3 burden than expected. Indeed, clade C2/3 tumors accounted for a significant fraction of the variance in the linear fit for substitutions, attributable to both SBS1 and SBS5, regressed against time (Fig. 3G). These observations suggest that a transient reduction in mutation rate occurred during the chain of transmissions taking place between 1991 and 2003 that transported DFT1 into Tasmania's northwest, perhaps due to a temporary reduction in cell division rate. Such fluctuations in mutation rate may not be uncommon, with detection in this particular case made possible because of the long internal branch and particularly dense sampling of DFT1 clade C2/3.

A DFT1 hypermutator lineage

Although most DFT1 and DFT2 tumors have very similar mutational spectra, a single DFT1 tumor, the unique representative of the early divergent clade E, named 377T1, had a highly distinctive pattern of mutations (Fig. 3H). Signature fitting suggested that, in addition to SBS1, SBS5, ID1, and ID2, this tumor also carried mutations attributable to mutational signatures SBS6 and ID7 (fig. S9). Furthermore, 377T1 carried 6 and 10 times more substitutions and indels, respectively, than expected from other DFT1 tumors that were sampled at a similar time (Fig. 3H). Because SBS6 and ID7, as well as elevated activity of ID1 and ID2, have been linked to deficiencies in DNA mismatch repair (*25*, *26*), these observations suggest that a clonal ancestor of 377T1 lost mismatch repair function. To identify the lesion that disrupted mismatch repair in 377T1, we screened the sequences of genes that encode mismatch repair effectors in DFT1 tumor genomes and discovered a focal deletion specific to 377T1 that removed a single copy of *MLH1* (Fig. 3H). Supporting a role for this gene, the 377T1 mutational spectrum is highly reminiscent of that reported in human cells that lack *MLH1* (*29*). No mutations, however, were detected in the remaining copy of *MLH1*, and we speculate that this may have been transcriptionally silenced, for example, by promoter DNA methylation.

Transposable element activity in DFT1 and DFT2

Transposable elements are frequently active in human cancer (*30*), but it is not known whether these are mobilized in Tasmanian devil cancers. Several families of transposable elements are annotated in mSarHar1.11, including 1948 full-length long interspersed nuclear element 1 (LINE-1) retroelements (table S1). We systematically screened for somatic LINE-1 insertions in DFT1 and DFT2 and found high LINE-1 transposition activity in DFT2, with hundreds of insertions detected. In DFT1, however, no clear evidence of LINE-1 activity was found (table S4). LINE-1 mobilization events were observed throughout the DFT2

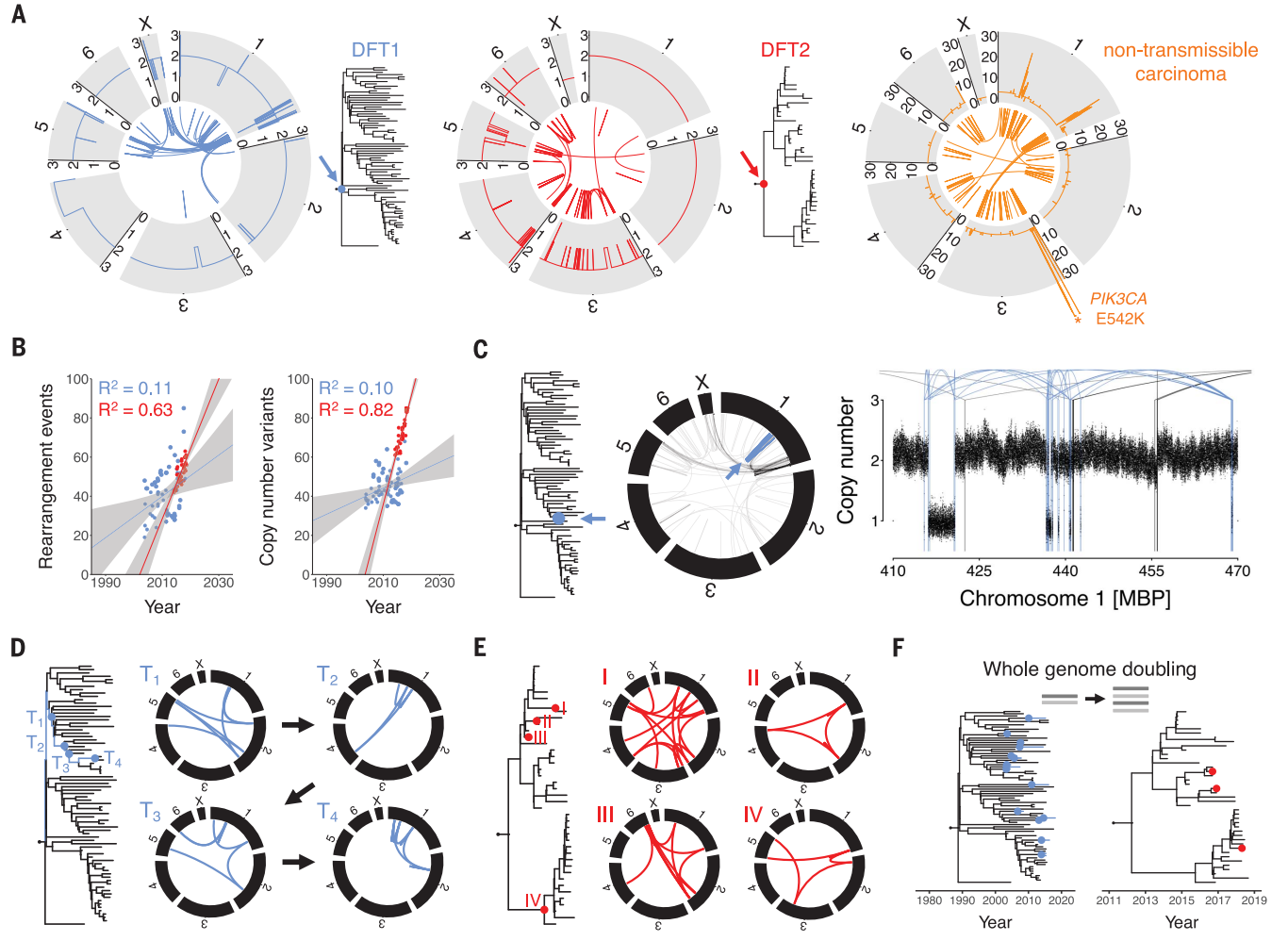


Fig. 5. Genome rearrangement in DFT1 and DFT2. (**A**) Rearrangement and copy number profiles of the DFT1 (left, blue) and DFT2 (center, red) most recent common ancestor tumors (trees, arrows; DFT1 and DFT2 trees as shown in Fig. 1, D and E, respectively). Chromosomes are represented by gray blocks annotated with copy number state. Inner arcs represent rearrangements. Right, rearrangement and copy number profiles of a single Tasmanian devil nontransmissible carcinoma. The location of the highly amplified E542K mutation in *PIK3CA* is labeled (asterisk). (**B**) Rates of accumulation of rearrangement events (left; "events" denotes that clustered rearrangements have been merged) and CNVs (right) in DFT1 (blue) and DFT2 (red). Tumors are represented by points, plotted by sampling date. Lines represent linear regression, gray shading 95% confidence interval. (**C**) Example of a late chromothripsis event in DFT1. A single DFT1 tumor (blue dot, arrow on phylogenetic tree) carries a chromothripsis event on chromosome 1; on the circos plot, rearrangements specific to the affected tumor are drawn in blue, and shared rearrangements that were acquired before this tumor's divergence are drawn in black. Right, copy number plot illustrates rearrangements involving the chromothriptic region (arcs; blue arcs are specific to this tumor, black arcs are shared with other tumors), and copy number is illustrated with binned coverage; each bin represents normalized read coverage in a 1-kb window. (**D** and **E**) Examples of chromoplexy events in DFT1 (left) and DFT2 (right). In both cases, positions of nodes represented by each circos plot are illustrated on the relevant phylogenetic tree, either along a four-step time-resolved (T_1 to T_4) branch trajectory in DFT1 (D) or throughout the DFT2 phylogeny (E). Chromosomes are represented by black blocks and rearrangements by colored arcs. (**F**) Timing of whole-genome doubling events in DFT1 (15 events) and DFT2 (3 events). The estimated date of each whole-genome duplication is illustrated on tree with colored dot. Further information and source data are available in figs. S10 and S11 and tables S5 and S6.

phylogenetic tree and accumulated linearly with time (Fig. 4A, Table 2, and table S4).

Transcriptional readthrough occasionally mobilizes genomic DNA downstream of LINE-1 source elements in a process known as 3′ transduction (*30*). A subset of DFT2 LINE-1 insertions carried 3′ transductions, which identified 35 functional LINE-1 source elements in DFT2 (table S4). Although most DFT2 source elements could be associated with only a single LINE-1 3′ transduction event, one source element located on chromosome 1 spawned at least 29 LINE-1 3′ transductions, with activity continuing throughout the DFT2 phylogenetic tree (Fig. 4, B and C). Overall, these findings reveal that LINE-1 retroelements are transposition competent in Tasmanian devil genomes and that their activity varies substantially between DFT1 and DFT2.

Genome rearrangement in DFT1 and DFT2

The availability of mSarHar1.11 enabled detailed reconstruction of the chromosomal rearrangements that initiated DFT1 and DFT2. The genome catastrophe that marked the origin of DFT1 is focused on the tip of the long arm of chromosome 1 (*10*, *14*, *31*). This region is massively internally rearranged through dozens of inversions interspersed with short deletions and interchromosomal translocations (Fig. 5A and tables S5 and S6). These changes are compatible with a complex chromothripsis event, as previously proposed (*14*). The early rearrangements of DFT2 are less clustered than those of DFT1 (Fig. 5A and tables S5 and S6) (*10*). Chromosome ends are notably involved in

rearrangement in both DFT1 and DFT2, which is consistent with a role for telomere dysfunction in DFT initiation (*10*, *14*, *31*).

The genome of the spontaneous nontransmissible anal sac carcinoma showed extensive rearrangement and copy number alteration (Fig. 5A and tables S5 and S6). This cancer's pattern of stepwise amplification is compatible with the activity of several breakage-fusion-bridge cycles. It is notable that the copy number landscape of this tumor is markedly more complex than those of the respective most recent common ancestors of DFT1 and DFT2, which indicates that, as in humans, there are several routes to carcinogenesis in Tasmanian devils. This is important because it implies that the mutational patterns observed in DFT1 and DFT2 are typical of DFT, not of Tasmanian devil cancer in general.

Rearrangement events and copy number variants (CNVs) both accumulated linearly with time in DFT2 (Fig. 5B and tables S5 and S6). Although slight temporal increases were detected in DFT1, these were only marginally significant, which confirms previous findings that the rate of genomic structural change in DFT1 is barely detectable above background variation among sublineages (*18*). Despite this, it is noteworthy that the group of DFT1 clade C2/3 tumors that carried fewer substitution mutations than expected (Fig. 3G) also showed fewer rearrangement events and CNVs (fig. S10), which suggests that the transient reduction in mutation rate occurring on the westward transmission chain operated across mutation classes.

The spectra of polymorphic (i.e., occurring after each lineage's most recent common ancestor) genomic rearrangements in DFT1 and DFT2 were similar, with small-scale alterations dominating (tables S5 and S6). Several more complex events were also observed in both lineages, however, including occasional chromothripsis (Fig. 5C) and ongoing chromoplexy (Fig. 5, D and E). We investigated the genomic contexts and haplotype specificity of a subset of CNVs that were observed to occur repeatedly either within or between DFT lineages (*18*); one of these was associated with repetitive structural features that likely trigger genome instability (table S6). Copy-neutral variation in minor copy number was rare in DFT1 and undetectable in DFT2, which is consistent with these tumors' overall patterns of copy number stability (*18*).

Whole-genome doubling in DFT1 and DFT2

Among the 78 DFT1 and 41 DFT2 tumors analyzed, 16 DFT1s and 3 DFT2s were identified as likely tetraploid, which defined 15 DFT1 and 3 DFT2 whole-genome duplication events. By counting the number of substitution mutations occurring before and after genome duplication in each tetraploid lineage and applying the previously inferred substitution mutation rates, we estimated the dates upon which genome doubling occurred. This identified whole-genome duplications that predated sampling of tumors by up to 7 years (median 1.8) (Fig. 5F, fig. S11, and table S6). DFT tumors that had undergone genome duplication showed an increased frequency of whole-chromosome or whole–chromosome-arm gain or loss events, compared with diploid tumors (Fisher's exact test $P < 0.01$) (table S6). This may at least in part be due to mitotic spindle defects introduced secondary to centrosome duplication (*32*) or due to a shortage of chromosome replication effectors in the first cell cycle after genome doubling (*33*); alternatively, it is possible that such large-scale aberrations are better tolerated in the tetraploid state.

Signals of selection in DFT1 and DFT2

The mutations that initiated DFT1 remain unknown, although a number of candidates have been proposed (*10*, *11*, *31*). It seems almost certain that the catastrophic event at the origin of DFT1 produced one or more driver mutations. The complex disruption of a single copy of *LZTR1* (*31*) is the most plausible driver candidate associated with this event (Fig. 6, A and B). In DFT2, focal copy number amplification of *PDGFRA* is shared by all DFT2 tumors and remains a strong early driver candidate (Fig. 6A) (*10*). In contrast to DFT1 and DFT2, the nontransmissible carcinoma carries recognizable driver mutations in well-characterized cancer genes (E542K *PIK3CA* mutation amplified to >60 copies; *TP53* truncation; *NOTCH2* mutations) (Fig. 5A and tables S6 and S7). Overall, the paucity of clear early driver mutations in DFT1 and DFT2, as well as the absence of causative cancer genes shared by both lineages, suggests that these cancers arose from a cell type that, perhaps by virtue of its epigenetic or transcriptional state, was predisposed to carcinogenesis and required only minimal genetic perturbation to produce transmissible cancer.

To explore ongoing evolution in DFT1 and DFT2, we first used *dNdScv* (*34*) to analyze evolutionary signal among substitution and indel mutations (Fig. 6C and table S8). This provided no evidence for widespread negative selection acting to remove deleterious mutations from the coding genomes of DFT1 or DFT2. However, a single gene in DFT1, *MGA*, which encodes a transcription factor that opposes *MYC* activity, showed plausible signs of positive selection through repeated truncation (global likelihood ratio test $q < 0.005$) (Fig. 6D). *MGA* has been implicated in cancer, although its driver status is not confirmed (*35*, *36*), and occurs in a haploid state in nearly all DFT1s (Fig. 6A).

Next, we searched for evidence of late drivers involving copy number variation. We created a chromosome map that displays total CNV burden within the sampled DFT1 and DFT2 population and examined this for focal amplification (Fig. 6E). This screen detected the previously described repeated amplification of *PDGFRB* in DFT1 (Fig. 6, A and E) (*10*, *18*) and indicated that further copy number gains of the early *PDGFRA* amplicon in DFT2 have occurred repeatedly in DFT2 clade A (Fig. 6, A and E). This analysis also identified two known recurrent focal amplifications on chromosomes 4 and 5 in DFT1, the latter containing *HMGA2*, and the former carrying 16 genes including *BIRC5* (*18*). In addition, although they are not recurrent, the focal amplification of *RAC1* to four copies in a single DFT1 and focal homozygous deletion of *PTEN* in one DFT2 stand out as potential late driver events (table S6).

DFT2 arose from a male founder devil and thus carries chromosome Y. The skew toward male hosts that is present in the DFT2 population (*7*), as well as a previous observation that chromosome Y had been lost from a single female DFT2 host, prompted speculation that loss of chromosome Y (LoY) may be under positive selection in DFT2 by reducing the immunogenicity of this cancer in female hosts (*10*). We investigated this hypothesis by analyzing the copy number of chromosome Y in our panel of DFT2 tumors. We detected five LoY events throughout the phylogeny of the 41 DFT2 tumors analyzed, one of which occurred in the ancestor of DFT2 clade B and is shared among all tumors of this group (Fig. 6, A and E). Somatic LoY is commonly observed in human normal and cancer cells, and the role of selection in driving this alteration in these contexts is poorly understood (*37*–*39*). Thus, although suggestive, we cannot confirm that DFT2 LoY is under positive selection; indeed, somatic LoY was observed in the analyzed nontransmissible devil anal sac carcinoma (table S6). However, it is noteworthy that a previous study that tracked the karyotype of a $chrY^+$ DFT2 cell line through two hundred passages in vitro made no mention of LoY in this immunologically neutral setting (*40*). If the presence of chromosome Y is indeed an immunological barrier to the colonization of female hosts, then no sex imbalance would be expected among hosts of $chrY^-$ DFT2.

Discussion

The assembly of a highly complete and contiguous reference genome for the Tasmanian devil has enabled comprehensive genomic characterization of this species' two transmissible cancers. DFT1 and DFT2 are independent realizations of a common biological phenomenon. Although the two cancers are overall highly similar in their genome features, especially when compared with a nontransmissible Tasmanian devil cancer, several differences exist; this ecological niche will tolerate different forms.

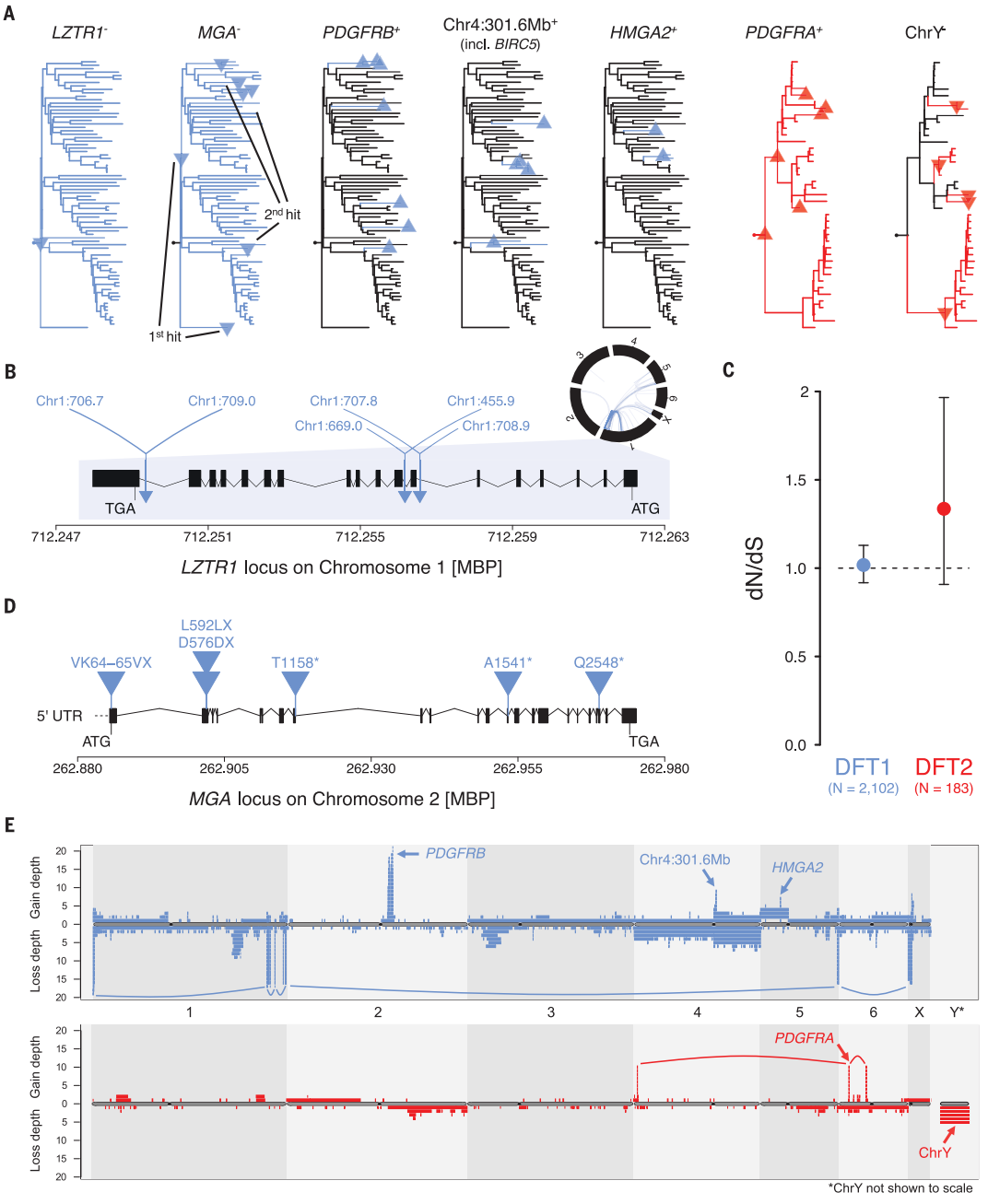


Fig. 6. Signals of selection in DFT1 and DFT2. (**A**) Phylogenetic positions of candidate driver mutations in DFT1 (blue) and DFT2 (red). Upward-pointing triangles and "+" notation represent copy number amplifications; downward-pointing triangles and "–" notation represent copy number losses or gene inactivation events. Multiple gains or losses in the same phylogenetic node are only represented once. DFT1 and DFT2 trees as shown in Fig. 1, D and E, respectively. Chr, chromosome. (**B**) Rearrangement of a single copy of *LZTR1* in DFT1. *LZTR1* (exons represented by black boxes, introns with black connectors) occurs within the densely rearranged region of chromosome 1 that is common to all DFT1s (circos plot; black bars represent chromosomes and blue arcs represent rearrangements common to all DFT1s; table S5). The location of each rearrangement in *LZTR1* is represented by a triangle, with the coordinates of each partner locus labeled. (**C**) Normalized ratio of nonsynonymous-to-synonymous substitutions and indels (dN/dS) in DFT1 and DFT2. Dashed line indicates dN/dS = 1 (neutrality), and bars represent 95% confidence intervals. (**D**) Genomic representation of the *MGA* locus on chromosome 2 in DFT1, exons represented by black boxes, introns with black connectors. Blue triangles represent the six coding mutations identified in this gene, all of which are truncating (tables S7 and S8). 5′ UTR, 5′ untranslated region. (**E**) Map representing CNVs detected within the sampled cohort of 78 DFT1 (upper, blue) and 41 DFT2 (lower, red) tumors. Chromosomes are represented horizontally, with chromosome Y not shown to scale. Each CNV is represented by a colored bar, with copy number gains illustrated above the gray chromosome representation ("gain depth") and copy number losses illustrated below the chromosome representation ("loss depth"). Mitotically inherited CNVs are represented once; thus, each colored bar represents a single CNV occurrence. CNVs that co-occur in the same tumors, and are thus likely to be linked, are connected with colored arcs; in DFT1, the set of linked losses are associated with the unstable small chromosome known as marker 5 (*18*). Arrows label candidate driver genes or genomic coordinates associated with prominent focal amplicons. Data associated with this figure are available in tables S6 to S8. Table S6 shows haplotype phasing of selected recurrent CNVs.

A particularly notable difference between DFT1 and DFT2 is the elevated mutation rate, observable across mutation classes, of DFT2 (Table 2). One explanation for this would be that DFT2 has a faster cell division rate than DFT1 and thus greater opportunity for the accrual of mutations associated with DNA replication. If true, this might influence relative growth rates and generation times of DFT1 and DFT2, with potentially complex epidemiological implications. However, other differences in cell state unrelated to division rate, perhaps, for instance, associated with differentiation state of the two cancers' cells of origin (*9*, *10*, *12*), may underlie this observation. Furthermore, although it is tempting to attribute the elevation in rates across different mutation classes in DFT2 to a common cause, it is possible that these are, in fact, unrelated, particularly as the magnitude of difference varies among mutation classes and signatures (Table 2). In particular, the LINE-1 retrotransposition activity we observed in DFT2, but not in DFT1, may reflect differences in the two lineages' epigenetic states (*41*). More generally, the mutation rates inferred from DFT1 and DFT2 provide evidence that large-scale mutations, including rearrangement events, transposon insertions, and CNVs, can have clock-like properties within individual cancers.

Once arisen, mutations become subject to selection. Positive selection, which acts to increase the frequency of mutations that confer advantageous traits, is usually the dominant force in cancer evolution; negative selection, which operates to remove deleterious mutations, is also detectable in cancer, although weak (*34*). In transmissible cancers, the stochasticity of transmission may decrease the efficiency of selection, and neutral processes, such as genetic drift, are likely to be of particular importance in their evolution (*42*). Nevertheless, and despite the small sample size of our study, plausible signals of positive selection were detectable in DFT1 and DFT2, and it is likely that these are operating to increase the fitness of cells within tumors (e.g., *PDGFRB* and *PDGFRA* amplification in DFT1 and DFT2, respectively, and *MGA* loss of function in DFT1) and to enhance transmission potential (e.g., LoY in DFT2). Genetic variants that increase somatic mutation rate are themselves often causatively involved in cancer through their tendency to predispose cells to acquisition of secondary adaptive mutations. This may be exemplified in the putatively positively selected heterozygous truncating mutation in *MGA* that is observed in mismatch repair–deficient DFT1 clade E.

Predicting the future dynamics and impacts of DFT1 and DFT2 requires knowledge of these diseases' epidemiological parameters. Although estimates of basic reproductive number (R_0) and generation time have been proposed for DFT1 (*15*), considerable uncertainty remains. Phylodynamics methods provide tools for inference of epidemiological metrics from pathogen genomes; however, the small sample size and geographical structuring of our tumor dataset make it unsuitable for such analysis (*43*). Although we cannot predict the evolutionary outcomes of DFT1 and DFT2, one observation that is worthy of comment is the surprisingly long delay between the origin of DFT1 (1982 to 1989) and its detection (1996). During this interval several hundred devils were examined in northeastern Tasmania, the location of DFT1's first observation, but no evidence of DFT was recorded (*4*). This suggests that DFT1 may have remained at low frequency during this time and is compatible with a relatively low R_0 or a longer than expected generation time. This observation, together with that of the superspreading event that occurred shortly after DFT1's origin that involved transmission of a tumor from a single donor to at least six recipients and founded the six DFT1 clades, lends credibility to the hypothesis that R may be overdispersed in DFT and that a large fraction of transmissions may funnel through a small number of infectious tumor donors (*44*). Tumor, host, and seasonal factors may influence individual transmission potential (*45*).

DFT1 and DFT2 have revealed the existence of a biological niche suited for transmissible cancers in Tasmanian devils. There is no evidence that these cancers emerged as a direct consequence of human actions through, for example, the introduction of chemical carcinogens or oncogenic viruses. Thus, it seems most likely that DFTs are a natural part of Tasmanian devil ecology. Although postcolonial human activities may have created conditions that indirectly benefitted DFT emergence or spread, for example, through habitat modification that may have supported increased devil density (*46*), it is very likely that DFTs have occurred in the past and that additional clones will emerge in the future. Notably, many incipient DFTs may die out before detection, particularly if these diseases have superspreading dynamics. Although no specific actions can be taken to prevent the establishment of additional DFTs, it will be important to continue close monitoring of wild and captive devil populations.

Although DFT transmissible cancers might themselves be natural occurrences, these diseases' devastating impact on their host species is exacerbated by anthropogenic threats, including loss of habitat and roadkill (*47*, *48*). Several recent studies have used longitudinal monitoring data to parameterize models that predict future Tasmanian devil population size and have argued against DFT1-induced extinction as a likely outcome (*49*–*51*). However, there is consensus that the species remains under threat, particularly given that its potential for persistence at much reduced density is unknown. It is thus important that adaptive monitoring, research, and management continue to be prioritized to ensure the long-term conservation and resilience of the Tasmanian devil (*47*, *52*–*54*).

Overall, this survey of the genomes of the two Tasmanian devil transmissible cancers has illuminated the evolutionary history of these unusual pathogens. Our analysis suggests that Tasmanian devils host a cell type that is poised for transmissible cancer transformation, with only minimal somatic genetic disruption required for these to be unleashed. Once established, DFT clones continue to acquire mutations at constant rates, and, although most of these are neutral, a small subset drive further adaptation to the niche. The future trajectories of DFT lineages and their Tasmanian devil hosts remain uncertain; however, this study provides a vantage point from which to further explore the evolution and impacts of transmissible cancers in this iconic marsupial species.

REFERENCES AND NOTES

1. A. M. Pearse, K. Swift, *Nature* **439**, 549 (2006).
2. R. Loh *et al.*, *Vet. Pathol.* **43**, 890–895 (2006).
3. R. J. Pye *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 374–379 (2016).
4. C. E. Hawkins *et al.*, *Biol. Conserv.* **131**, 307–324 (2006).
5. B. T. Lazenby *et al.*, *J. Appl. Ecol.* **55**, 1368–1379 (2018).
6. Y. M. Kwon *et al.*, *R. Soc. Open Sci.* **5**, 180870 (2018).
7. S. James *et al.*, *Evol. Appl.* **12**, 1772–1780 (2019).
8. C. E. Hawkins, H. McCallum, N. Mooney, M. Jones, M. Holdsworth, "*Sarcophilus harrisii*" (The IUCN Red List of Threatened Species 2008, e.T40540A10331066, 2008); https://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T40540A10331066.en.
9. E. P. Murchison *et al.*, *Science* **327**, 84–87 (2010).
10. M. R. Stammnitz *et al.*, *Cancer Cell* **33**, 607–619.e15 (2018).
11. L. Kosack *et al.*, *Cancer Cell* **35**, 125–139.e9 (2019).
12. A. L. Patchett *et al.*, *Cell. Mol. Life Sci.* **77**, 1847–1858 (2020).
13. E. P. Murchison *et al.*, *Cell* **148**, 780–791 (2012).
14. J. E. Deakin *et al.*, *PLOS Genet.* **8**, e1002483 (2012).
15. H. McCallum *et al.*, *Ecology* **90**, 3379–3392 (2009).
16. H. V. Siddle *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 5103–5108 (2013).
17. A. Caldwell *et al.*, *eLife* **7**, e35314 (2018).
18. Y. M. Kwon *et al.*, *PLOS Biol.* **18**, e3000926 (2020).
19. W. Miller *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 12348–12353 (2011).
20. A. H. Patton *et al.*, *Mol. Biol. Evol.* **36**, 2906–2921 (2019).
21. M. Jain *et al.*, *Nat. Biotechnol.* **36**, 338–345 (2018).
22. B. L. Aken *et al.*, *Database* **2016**, baw093 (2016).
23. S. Nee, E. C. Holmes, R. M. May, P. H. Harvey, *Phil. Trans. R. Soc. Lond. B* **344**, 77–82 (1994).
24. L. B. Alexandrov *et al.*, *Nature* **500**, 415–421 (2013).
25. L. B. Alexandrov *et al.*, *Nature* **578**, 94–101 (2020).
26. X. Zou *et al.*, *Nat. Cancer* **2**, 643–657 (2021).
27. F. Blokzijl *et al.*, *Nature* **538**, 260–264 (2016).
28. H. Lee-Six *et al.*, *Nature* **574**, 532–537 (2019).
29. J. Drost *et al.*, *Science* **358**, 234–238 (2017).
30. B. Rodriguez-Martin *et al.*, *Nat. Genet.* **52**, 306–319 (2020).
31. R. L. Taylor, Y. Zhang, J. P. Schöning, J. E. Deakin, *Sci. Rep.* **7**, 8761 (2017).
32. N. J. Ganem, Z. Storchova, D. Pellman, *Curr. Opin. Genet. Dev.* **17**, 157–162 (2007).

33. S. Gemble *et al*., . *Nature* **604**, 146–151 (2022).
34. I. Martincorena *et al*., *Cell* **171**, 1029–1041.e21 (2017).
35. F. Martínez-Jiménez *et al*., *Nat. Rev. Cancer* **20**, 555–572 (2020).
36. H. Mathsyaraja *et al*., *eLife* **10**, e64212 (2021).
37. P. Priestley *et al*., *Nature* **575**, 210–216 (2019).
38. T. H. H. Coorens *et al*., *Nature* **597**, 387–392 (2021).
39. D. J. Thompson *et al*., *Nature* **575**, 652–657 (2019).
40. E. D. Ingles, J. E. Deakin, *Genes* **11**, 480 (2020).
41. K. H. Burns, *Nat. Rev. Cancer* **17**, 415–424 (2017).
42. A. Baez-Ortega *et al*., *Science* **365**, eaau9923 (2019).
43. A. H. Patton *et al*., *Science* **370**, eabb9772 (2020).
44. J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, W. M. Getz, *Nature* **438**, 355–359 (2005).
45. M. E. Jones *et al*., in *Wildlife Disease Ecology: Linking Theory to Data and Application*, K. Wilson, A. Fenton, D. Tompkins, Eds. (Cambridge Univ. Press, 2019), pp. 321–341.
46. M. E. Jones, D. Paetkau, E. Geffen, C. Moritz, *Mol. Ecol.* **13**, 2197–2209 (2004).
47. H. McCallum, M. Jones, *PLOS Biol.* **4**, e342 (2006).
48. A. J. Hobday, M. L. Minstrell, *Wildl. Res.* **35**, 712–726 (2008).
49. V. Siska, A. Eriksson, B. Mehlig, A. Manica, A metapopulation model of the spread of the Devil Facial Tumour Disease predicts the long term collapse of its host but not its extinction. arXiv:1806.05449 [q-bio.PE] (2018).
50. C. X. Cunningham *et al*., *Ecol. Lett.* **24**, 958–969 (2021).
51. K. Wells *et al*., *Ecology* **100**, e02613 (2019).
52. A. Patchett, G. Woods, *Science* **365**, 438–440 (2019).
53. C. E. Grueber, E. Peel, B. Wright, C. J. Hogg, K. Belov, *Reprod. Fertil. Dev.* **31**, 1296–1304 (2019).
54. K. A. Farquharson *et al*., *iScience* **25**, 104474 (2022).
55. T. Leitner, J. Albert, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10752–10757 (1999).
56. M. A. Nilsson *et al*., *PLOS Biol.* **8**, e1000436 (2010).
57. M. A. Suchard *et al*., *Virus Evol.* **4**, vey016 (2018).
58. M. R. Stammnitz, Supplementary data sets, *Zenodo* (2022); https://doi.org/10.5281/zenodo.6476544.

ACKNOWLEDGMENTS

We are grateful to B. Lehner, H. Siddle, R. Owen, A. Gastaldello, A. Flies, K. Hughes, S. Mayes, F. Giordano, I. Mickleburgh, A. Scally, B. Aken, the Wellcome Sanger Institute sequencing pipelines team, and past and current members of the Transmissible Cancer Group for technical assistance and helpful discussions. We also thank students, staff, and volunteers involved in Tasmanian devil field work and sample collection. **Funding:** This work was supported by grants from Wellcome (102942/Z/13 and 222551/Z/21), as well as Eric Guiler Tasmanian Devil Research Grants from the University of Tasmania Foundation. F.J.M. and K.B. were funded by Wellcome grants WT108749/Z/15/Z and WT222155/Z/20/Z and an Eric Guiler Tasmanian Devil Research Grant from the University of Tasmania Foundation. Additional field work was supported by grants from the Australian Research Council (LP0561120, LP0989613, and DP110102656 to M.J.; DE170101116 and LP170101105 to R.H.), the US National Science Foundation (DEB-1316549 to M.J. and E.P.M.), and the US National Institutes of Health (1R01GM126563-01 to M.J.). R.H. is supported by the MAVA Foundation. M.R.Sta. was supported by a Gates Cambridge Trust Scholarship. **Author contributions:** E.P.M. designed and directed the project. M.R.Sta. led laboratory efforts, methods development, and computational data analysis. K.G., Y.M.K., E.H., A.B.-O., and H.E. assisted with methods development and tumor genome bioinformatics. M.R.Sta., E.H., F.J.M., K.B., Y.C., W.C., M.A.Q., J.Wo., K.H., and Z.N. generated the mSarHar1.11. reference genome assembly and functional annotation. M.R.Sta., K.G., Y.M.K., S.C., S.F., R.H., M.J., B.L., S.P., R.P., K.S., J.Wa., and E.P.M. undertook fieldwork, tissue sampling, and processing. M.R.Str. provided technical advice. E.P.M. and M.R.Sta. wrote the manuscript and designed the figures. All authors commented on the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Tasmanian devil tumor and normal whole-genome sequence data have been deposited in the European Nucleotide Archive (ENA; www.ebi.ac.uk/ena) under accession no. PRJEB51704. Likewise, the Tasmanian devil reference genome mSarHar1.11 is available under ENA accession no. GCA_902635505.1, with associated raw sequence data under accession no. PRJEB34649 and multitissue transcriptome data for gene annotation under accession no. PRJEB34650. Variant calling and other data supporting our analyses have been deposited at Zenodo (*58*). Custom scripts used for data processing, analysis, and visualization are also available on GitHub (https://github.com/MaximilianStammnitz/TCG2023).

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse. This research was funded in whole or in part by Wellcome (102942/Z/13 and 222551/Z/21), a cOAlition S organization. The author will make the Author Accepted Manuscript (AAM) version available under a CC BY public copyright license.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq6453
Materials and Methods
Figs. S1 to S11
Tables S1 to S8
References (*59*–*117*)
MDAR Reproducibility Checklist

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

Submitted 22 April 2022; accepted 20 March 2023
10.1126/science.abq6453

NEUROEVOLUTION

Syncytial nerve net in a ctenophore adds insights on the evolution of nervous systems

Pawel Burkhardt[1]*, Jeffrey Colgren[1]†, Astrid Medhus[1]†, Leonid Digel[1], Benjamin Naumann[2], Joan J. Soto-Angel[1], Eva-Lena Nordmann[1], Maria Y. Sachkova[1], Maike Kittelmann[3]*

A fundamental breakthrough in neurobiology has been the formulation of the neuron doctrine by Santiago Ramón y Cajal, which stated that the nervous system is composed of discrete cells. Electron microscopy later confirmed the doctrine and allowed the identification of synaptic connections. In this work, we used volume electron microscopy and three-dimensional reconstructions to characterize the nerve net of a ctenophore, a marine invertebrate that belongs to one of the earliest-branching animal lineages. We found that neurons in the subepithelial nerve net have a continuous plasma membrane that forms a syncytium. Our findings suggest fundamental differences of nerve net architectures between ctenophores and cnidarians or bilaterians and offer an alternative perspective on neural network organization and neurotransmission.

For more than a century, the structure and evolutionary origin of the animal nervous system have been at the center of much debate among biologists. Fundamental progress in our structural understanding was put forward by Santiago Ramón y Cajal, who postulated that the nervous system is composed of discrete cells called neurons (*1*). This contrasts with Camillo Golgi's proposition that the nervous system is a syncytial continuum. The discovery of synaptic connections between individual neurons by electron microscopy later confirmed Cajal's theory. However, there is accumulating evidence that ctenophores, gelatinous marine invertebrates that move through the water column by ciliary comb rows, are among the earliest branching extant lineages of the animal kingdom (Fig. 1A) (*2*–*5*). Most ctenophore life cycles include a predatory cydippid stage during which, for some species, the ctenophore is able to reproduce only a few days after hatching (Fig. 1B) (*6*). Ancestral-state reconstruction suggests that the cydippid body plan is a plesiomorphic character of ctenophores (*7*).

The early split of ctenophores from other groups indicates that a nervous system, and maybe even neurons, could have evolved at least twice: once within the ctenophores and once within the lineage of the remaining animals (*8*). Initiated through genomic analyses (*2*, *3*), molecular and physiological features of the ctenophore nervous system were subsequently interpreted to support this scenario (*4*, *5*). In contrast to sponges and placozoans, ctenophores exhibit an elaborate nervous system consisting of a subepithelial nerve net (SNN), mesogleal neurons, a sensory aboral organ, tentacle nerves, and diverse sensory cells in all parts of their body (Fig. 1C and movie S1) (*9*–*14*). Deciphering the development, structure, and function of the ctenophore nervous system is a key element to understand the origin and evolution of animal nervous systems. We have recently shown that a large repertoire of lineage-specific neuropeptides has evolved in the ctenophore *Mnemiopsis leidyi* (*14*). Furthermore, we identified a distinctive feature of SNN neurons: the multiple neurites extending from one soma are interconnected through anastomoses and thus form an extensive continuous network within a single nerve-net neuron (*14*). This characteristic sets them apart from other animal neurons. Additionally, there was little evidence on how these nerve-net neurons connect to each other, to sensory neurons, and to cells within the mesoglea because of the lack of synaptic markers suitable for fluorescent labeling or large-scale electron microscopic data that spans multiple neurons. In this study, we used high-pressure freezing-fixation techniques in combination with serial block face scanning electron microscopy (SBFSEM) to establish the first ultrastructural three-dimensional (3D) network of SNN neurons and other cell types in a ctenophore.

The cydippid SNN is organized in a syncytium

Recent 3D reconstruction of a nerve-net neuron in a cydippid-phase *M. leidyi* has revealed a wide network of anastomosed neurites extending from only one soma (*14*). However, to understand the nature of connections between multiple nerve-net neurons as well as other cell types, we collected a larger continuous SBFSEM dataset of an early cydippid that includes 5 nerve-net neurons, 6 mesogleal neurons, and 22 putative sensory cells. The neurites of all five SNN cells were connected through an anastomosed continuous network (Fig. 2A). Whereas gap junctions could readily be identified within comb plates (fig. S1), we detected neither electrical nor chemical synapses between the cells of the SNN. We confirmed this observation in smaller datasets of the nerve net beneath two comb rows and along the gut in two other cydippid individuals (fig. S2). Additionally, injection of the fluorescent lipophilic dye 1,1'-dioctadecyl-3,3,3′,3′-tetramethylindocarbocyanine perchlorate (DiI) into only one of the cells of two-cell staged embryos led to a fluorescent signal in only one-half of the cydippid body, which was seen in SNN cell bodies throughout the animal consistent with the syncytial nature of the SNN (fig. S3).

Morphologically, neurites within the SNN exhibited no obvious polarity (axon versus dendrite), showing similar diameter, dense-core vesicles throughout their length, and the lack of typical presynaptic triads (Fig. 2, A to C). Moreover, SNN neurites often showed a blebbed or "pearls-on-a-string" morphology (Fig. 2, D to G, and fig. S4). The narrow segments were often just wide enough for microtubules to pass (Fig. 2G and fig. S4), and bulged segments often contained larger clear or electron-dense vesicles and occasionally endoplasmic reticulum (Fig. 2D and fig. S4). A recently developed antibody against the neuropeptide ML02736a (*14*) confirmed the presence of neuropeptides within some of the vesicles of SNN neurons (Fig. 2E and fig. S5). Although SNN neurons seemed to lack synapses between each other, we identified chemical synapses from the SNN to polster cells (fig. S6), which suggest directional signal transmission from the SNN to effector cells.

Mesogleal neurons form direct contacts with the syncytial SNN

We identified and reconstructed six mesogleal neurons exhibiting a starlike morphology with extensive plasma membrane protrusions of variable lengths (Fig. 3A). Their somata were filled with a variety of vesicles and larger vacuoles (Fig. 3B), and the protrusions of these cells did not show the pearls-on-a-string morphology present in neurites of the SNN. Some of the protrusions formed plasma membrane juxtapositions to neurites of the SNN (Fig. 3, A, D, and E). However, we did not find ultrastructural evidence for electrical or chemical synapses (Fig. 3E). In contrast to SNN neurons, we did not observe any electron-dense vesicles in mesogleal neurons (Fig. 3B), but instead small electron-lucent vesicles of a similar size as synaptic vesicles (Fig. 3C), which suggests a different type of information transmission.

Sensory cells form simple circuits involving the syncytial SNN

We identified and reconstructed a total of 22 putative sensory cells from the present

[1]Michael Sars Centre, University of Bergen, 5008 Bergen, Norway. [2]Institut für Biowissenschaften, Allgemeine und Spezielle Zoologie, Universität Rostock, 18055 Rostock, Germany. [3]Department of Biological and Medical Sciences, Oxford Brookes University, Oxford OX3 0BP, UK.
*Corresponding author. Email: pawel.burkhardt@uib.no (P.B.); maike.kittelmann@brookes.ac.uk (M.K.)
†These authors contributed equally to this work.

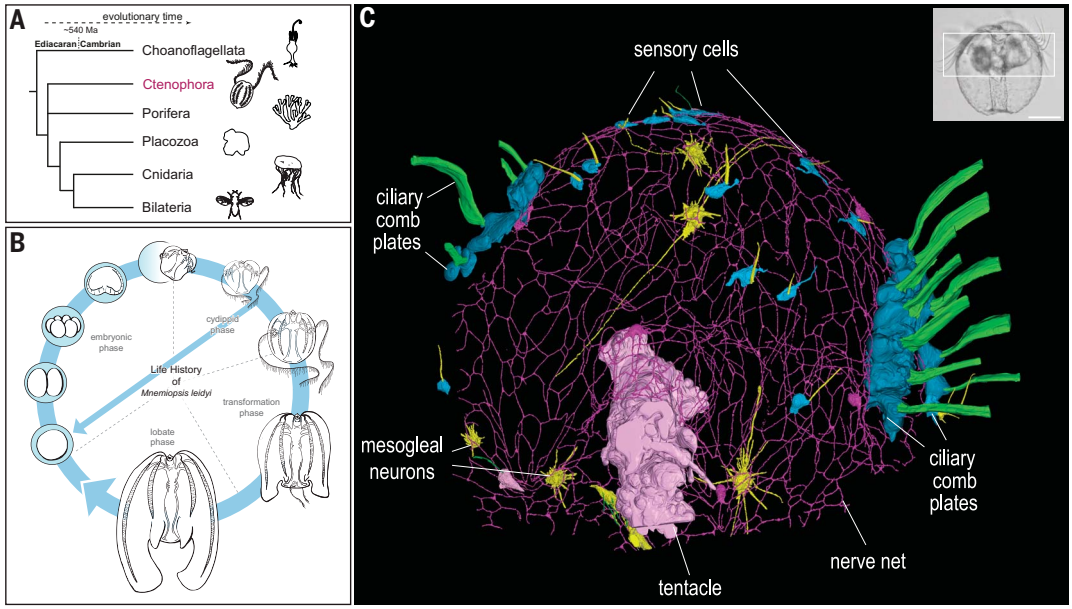


Fig. 1. Ctenophores and their nervous system. (**A**) Ctenophores as one of the earliest branching extant lineages of the animal kingdom. Silhouettes are from PhyloPic (www.phylopic.org). (**B**) The ctenophore *M. leidyi* exhibits complex life cycle stages, including a predatory cydippid phase that hatches from the egg and can reproduce after a few days. (**C**) 3D reconstruction of the nerve net, comb rows, sensory cells, mesogleal neurons, and a tentacle from SBFSEM data of a 1-day-old cydippid. (Inset) Phase contrast image of a 1-day-old cydippid. White box, area reconstructed in (C). Scale bar, 100 μm.

and an earlier dataset (*14*) that fit into five morphological groupings (Fig. 4, fig. S7, and table S1). Some of them resembled known ctenophore sensory cell types (types 1, 4, and 5) (*16*, *17*), whereas others exhibited a morphology that, to the best of our knowledge, has not been described previously (types 2 and 3) (Fig. 4, fig. S7, and table S1). We detected chemical synapses in several but not all putative sensory cells that contact neuronal or other effector cells (Fig. 4 and fig. S7). Type 1 sensory cells exhibited a single long cilium and onion-root basal body (Fig. 4 and fig. S7, A and B). Type 2 sensory cells exhibited a very short single cilium without an onion-root basal body. Long neurites extending from their somata formed chemical synapses to polster cells (Fig. 4B and fig. S7, A and C).

Type 3 sensory cells exhibited multiple cilia without onion-root basal bodies. Many large electron-dense vesicles are localized beneath the cilia (Fig. 4C and fig. S7, A and D). We found one of these cells near the tentacle with a synaptic connection to a mesogleal neuron (Fig. 4C). Type 4 sensory cells exhibited a single long filopodium. Some of them formed synapses to neurites of the SNN (Fig. 4, A and D), and some also received synaptic input from type 1 sensory cells (Fig. 4A). Type 5 sensory cells exhibited multiple long filopodia. They formed plasma membrane contact to polster cells, but we did not detect synaptic contacts from or to this cell type. Last, we used the 3D ultrastructural evidence to identify several discrete and simple neural circuits in early cydippid–phase *M. leidyi*. These circuits included synaptic signal transmission from sensory cells to other cell types including SNN neurons, mesogleal neurons, polster cells, or even other sensory cell types (Fig. 4, A to D).

Discussion

In the debate at the end of the 19th century about the organization of the animal nervous system, Joseph von Gerlach (1871) (*18*) and Camillo Golgi (1885) (*19*) put forward the reticular theory (also known as the syncytial theory). Both proposed the cellular continuity of neurons. This view was challenged by Cajal (1888) (*1*), who proposed an organization from discrete cellular units connected through synapses. Both contestant theories were founded on Golgi's newly invented black staining that enabled scientists to study the detailed morphology of neurons and their neurites (*20*). Golgi and Cajal were honored with the Nobel Prize in Physiology or Medicine in 1906 for their efforts in elucidating the architecture of the nervous system (*20*). However, with the advent of electron microscopy in the 1950s and the discovery of the synaptic cleft, the reticular theory was put to rest in favor of Cajal's hypothesis (*21*, *22*). In our present study, volume electron microscopy revealed the 3D ultrastructural architecture of the SNN in an early cydippid–phase ctenophore, providing evidence for its reticular—or syncytial—organization. Previous work suggested anastomosed nerve cords in adult ctenophores on the basis of chemical staining (*9*) and multiple parallel strands of neurites stained with anti–tyrosylated-α-tubulin (*10*). In this work, we showed that a syncytial nerve net already exists in cydippid-phase *M. leidyi*. This syncytium may be reinforced in adult animals through the anastomosis of additionally formed neurites; however, confirmation of such connectivity will require further detailed, high-resolution analysis of the nerve net throughout development.

Using high-pressure freezing and freeze substitution techniques to preserve fine ultrastructural details with minimal fixation artifacts, we showed that the SNN forms a continuous structure. This is further supported by the unrestricted spread of DiI throughout the nerve net.

Whereas gap junctions could be identified within the comb plates, as previously reported (*15*) in our SBFSEM data as well as in TEM micrographs, we found no evidence of similar structures between neurites of nerve-net neurons that would suggest the presence of electrical synapses. Additionally, a recent characterization of the complete set of *M. leidyi* innexins—responsible for the formation of gap junctions in invertebrates—did not show any mRNA expression in in situ hybridization experiments in nerve-net cell bodies (*23*). However, we did observe synaptic triads and plasma membrane contacts of unknown molecular structure that connect the SNN externally to polster and mesogleal neurons.

Previous characterizations of ctenophore nerve nets have been predominantly based on traditional histochemical staining techniques (*9*, *24*) and more recently on fluorescence microscopy of antibody staining against α-tubulin (*10*, *12*, *13*, *25*). Although both techniques provide valuable insight into the general organization and location of ctenophore neurons, they do not allow the investigation of the ultrastructure and nature of neuronal connections. Data from transmission electron microscopic serial sections (*26*, *27*) may also have overlooked this distinct

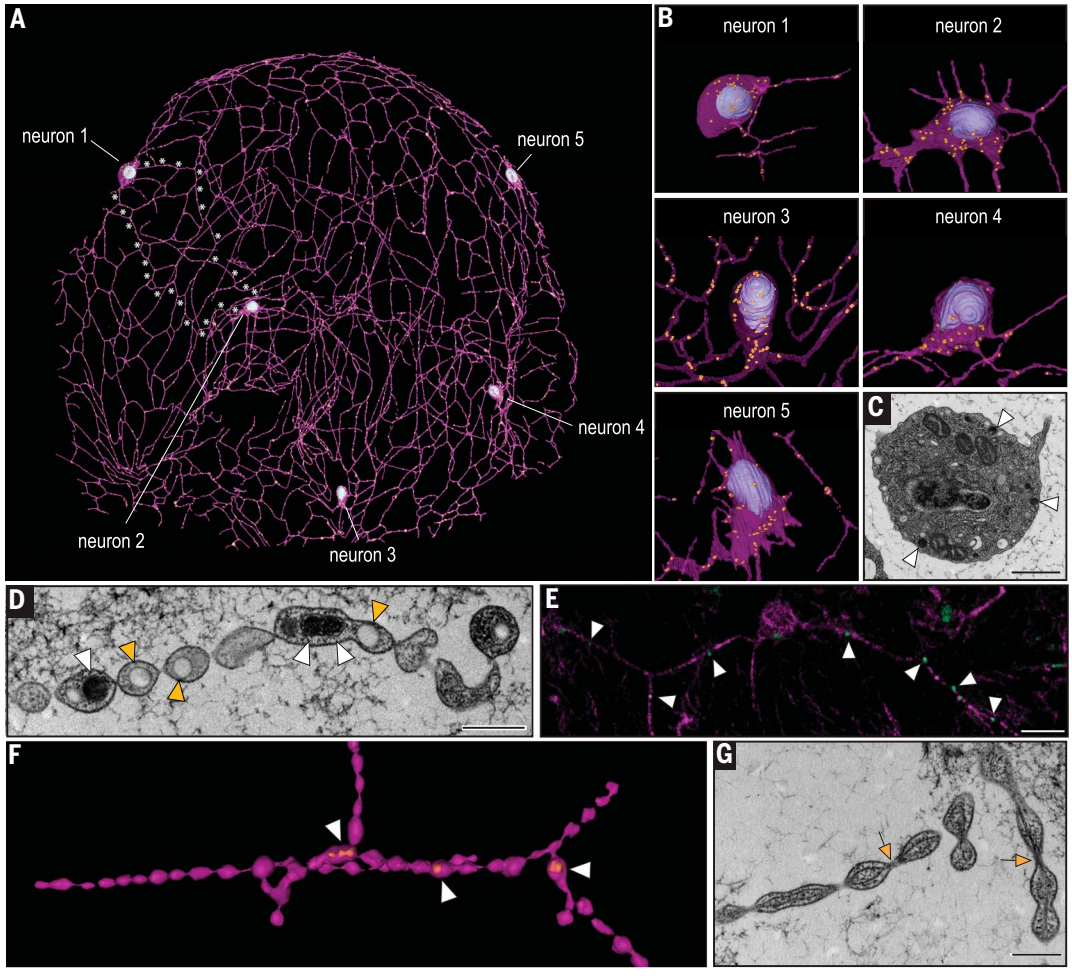


Fig. 2. Connectivity and ultrastructure of the ctenophore SNN. (**A**) 3D reconstruction of five SNN neurons. White asterisks indicate examples of the continuous membrane between the cell bodies of neurons 1 and 2. (**B**) 3D reconstruction of the SNN neuron cell bodies showing the nucleus (blue) and dense-core vesicles (orange). (**C**) TEM cross section of an SNN neuron cell body that shows ultrastructural details, including large, dense-core vesicles (white arrowhead). Scale bar, 1 μm. (**D**) TEM cross section of a SNN neurite with dense-core and clear-core vesicles localized in blebbed areas (white and orange arrowheads, respectively). Scale bar, 500 nm. (**E**) Antibody staining against neuropeptide ML02736a (green) in SNN neurites (magenta) stained for tubulin. (**F**) TEM 3D reconstruction of SNN neurite (violet) and dense-core vesicles (orange), highlighting the blebbed morphology. (**G**) TEM cross section of SNN neurites showing continuous microtubules (orange arrows) passing through narrow segments. Scale bar, 500 nm.

syncytial architecture because of the difficulty to produce continuous section series over such a large volume. Aside from reports on single self-anastomosing neurites in other animals (*28*–*30*), the presence of a complete syncytial nerve net has only been reported for the cnidarian, medusae-like colonial polyp, *Velella* (*31*, *32*). However, to the best of our knowledge, the syncytial organization of this nerve net has not yet been verified on an ultrastructural level. At this time, we have found this feature only in the ctenophore *M. leidyi* nerve net, but further analysis across nerve net–bearing animals may provide exciting insights into early nervous system evolution and modes of neuronal connectivity.

Although neurite fusion and pruning seem to be a common principle during the early neural development in many animals (*33*, *34*), we do not consider the syncytial cydippid SNN to be completely remodeled by such a process later in development. It was suggested that the early cydippid phase is not a larval phase but rather an autonomous life history phase of *M. leidyi* and other ctenophores (*6*). Indeed, cydippid-phase *M. leidyi* are free-swimming pelagic predators, able to reproduce and exhibit complex behaviors as described for their second, reproductive, lobate phase (*35*–*37*).

Our identification of the nonsynaptic architecture of the cydippid-phase SNN raises the question of the mechanism of signal propagation. Genome and single-cell transcriptome analyses revealed that *M. leidyi* SNN neurons express 1 voltage-gated calcium (Ca_v), 35 potassium (K_v), and 2 nonspecific sodium (Na_v) channels (*14*, *38*, *39*). These numbers are similar to those in neurons of other animals, and ctenophore SNN neurons may therefore be able to produce membrane potential or even action potentials (*40*). Moreover, the presence of numerous peptidergic vesicles in the SNN suggests that signal transmission also occurs through neuropeptide release, and the Ca_v channel expressed in these cells might be involved in exocytosis (*14*, *41*). Therefore, we can speculate that the SNN could function as a neuroendocrine system that is able to release transmitters into the mesoglea through vesicle fusion with the plasma membrane at different neurite sites. Such a system would require only a minimum number of chemical synapses and, if acting at short distances, may reach enough effector cells. Studies on the conduction velocity in ctenophores have shown a slower speed of signal propagation compared with that of nerve nets and conducting epithelia of other animals (*42*), indicating that signal propagation could be nonsynaptic.

Additionally, our ultrastructural identification of simple circuits now provides a basis that allows for better understanding of how mechanoreception, swimming, and prey-capture behavior in young cydippid-phase ctenophores could be facilitated. Numerous sensory neurons are connected through chemical synapses to the nerve net, which in turn forms chemical synapses onto effector cells such as the comb rows or ciliated groove cells (*14*). Type 1 ciliated sensory cells and type 4 filopodiated sensory cells, previously described as Tastborsten and Taststifte (*9*), have been postulated to be sensitive to water vibrations and touch (*17*, *43*, *44*). Their abundance throughout the epidermis and direct cell-to-cell contact to the nerve net (many through chemical synapses) underscore the importance of direct transmission of localized vibration and touch information to the SNN. Morphological analysis allows us to speculate that a type 2 sensory

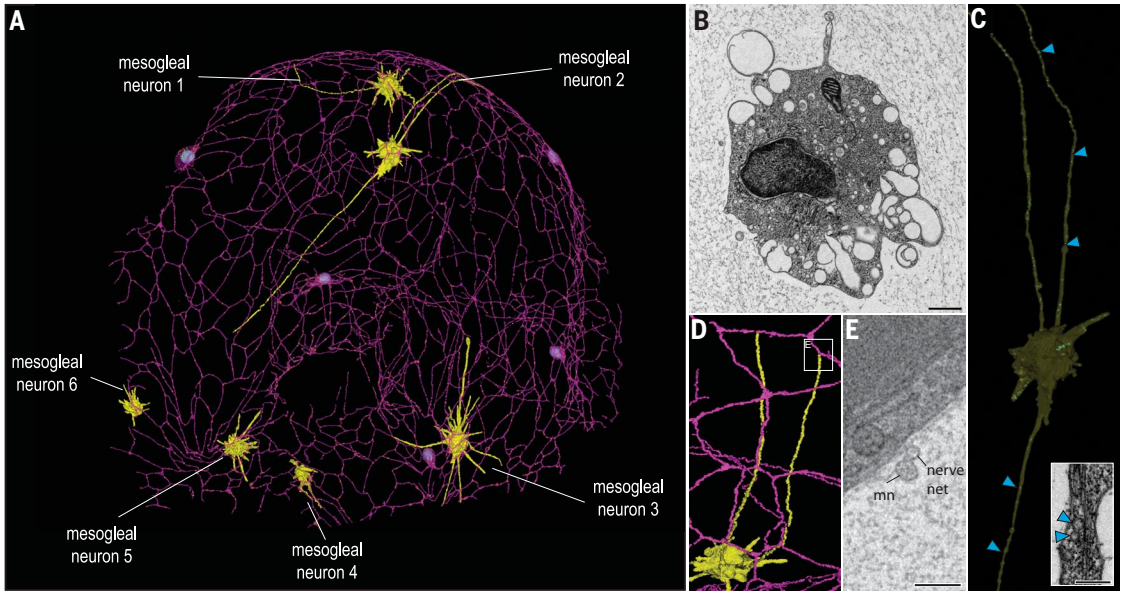


Fig. 3. Close association of mesogleal neurons and the SNN. (**A**) 3D reconstruction of SNN (violet) and mesogleal neurons (yellow) from SBFSEM data. (**B**) TEM cross section of a mesogleal neuron cell body. Different types of clear vesicles and vacuoles but no dense-core vesicles are present. Scale bar, 1 μm. (**C**) 3D reconstructed mesogleal neuron with three long neurites that contain small clear vesicles (blue arrowheads). (Inset) TEM cross section of mesogleal neurites with small clear vesicles. Scale bar, inset, 200 nm. (**D**) 3D reconstruction of mesogleal neuron with contact site (white box) to SNN. (**E**) Corresponding SBFSEM image of contact site between mesogleal neuron and SNN neuron. No chemical or electric synapse structures could be observed. Mn, mesogleal neuron. Scale bar, 500 nm.

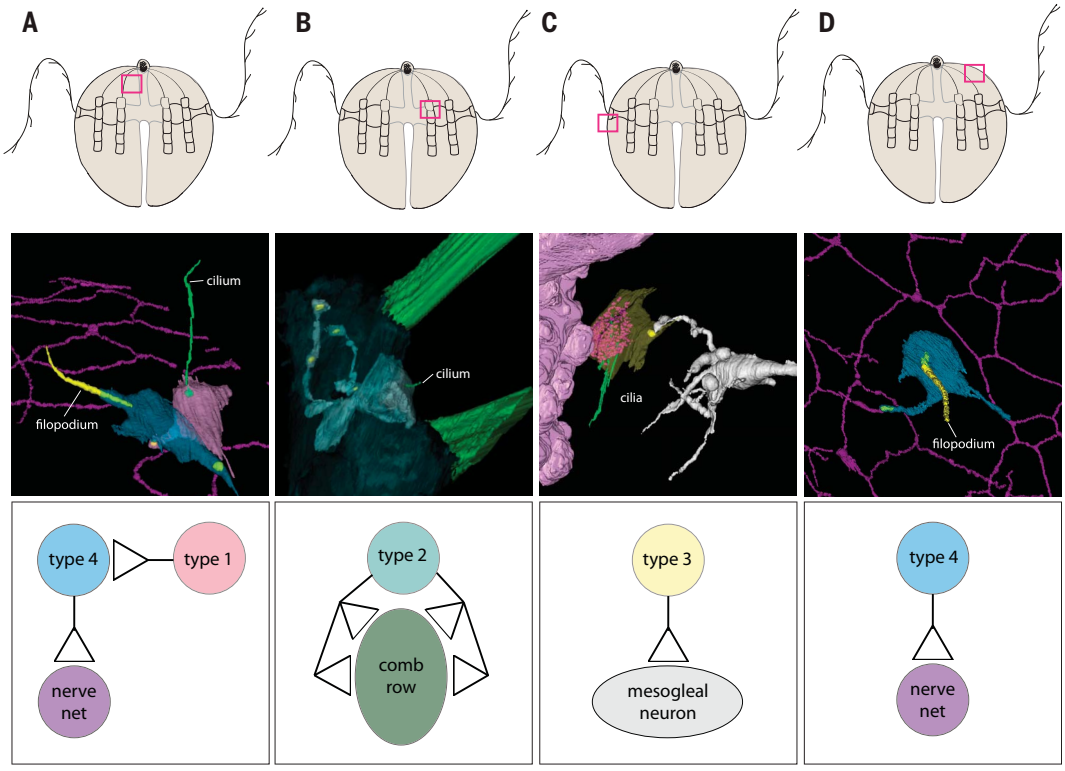


Fig. 4. 3D reconstruction of sensory cells allows for the identification of simple circuits. (Top) Localization of each circuit (pink square). (Middle) 3D reconstructions of sensory and effector cells. Mitochondria are shown in yellow to represent synaptic tripartite complexes in all circuits. (Bottom) Proposed wiring diagram. (**A**) Circuit between type 1 and type 4 sensory cells and SNN. (**B**) Multiple synaptic connections between a type 2 sensory cell with short cilium and comb cells. (**C**) Synaptic connection between a type 3 sensory cell near a tentacle and a mesogleal neuron. (**D**) A type 4 sensory cell with single filopodium synapses onto nerve net.

cell, which wraps around polster cells, may be able to detect water flow and thus alter comb-beat frequency, whereas a type 3 sensory cell, the multiple cilia of which are in close contact with the tentacle, may be triggered by food capture. Functional experiments are needed to fully understand the activity of these circuits and to unravel the different modes of signal transmission used by the different ctenophore neuronal cell types. This study is limited to the analysis of an early developmental stage in which fixation of whole animals with high-pressure freezing is still possible. Comparison with other ctenophore species and investigation of later life history stages of *M. leidyi* are needed to clarify whether a syncytial SNN is a feature restricted to an early ontogenetic phase in only a few species or is a common feature of all ctenophores. This approach will also provide valuable insights into the development of the syncytial SNN on whether neurons divide but remain connected in the cydippid SNN, or whether neurites from different cell bodies reach out and fuse.

Whether neurons of animals have a single origin or possibly originated more than once during evolution is a debated topic. The existing data on the ctenophore nervous system show a specific mosaic of cellular and syncytial components with distinct evolutionary histories. It will be a major future challenge to clearly identify the parts of the mosaic that may have evolved independently and the preexisting parts that were strongly modified. Our study underscores that the resemblance between the nerve net of ctenophores and the nerve nets of cnidarians and bilaterians might only be superficial because it appears that their connectivity is fundamentally different. Our ultrastructural analysis of the ctenophore SNN not only puts ctenophores at the center of nervous system evolution but also provides an opportunity to explore the boundaries of nervous system organization and function.

REFERENCES AND NOTES

1. S. R. Cajal, *Rev. Trim. Histol. Norm. Patol.* **1**, 1–10 (1888).
2. C. W. Dunn *et al*., *Nature* **452**, 745–749 (2008).
3. A. Hejnol *et al*., *Proc. Biol. Sci.* **276**, 4261–4270 (2009).
4. J. F. Ryan *et al*., *Science* **342**, 1242592 (2013).
5. L. L. Moroz *et al*., *Nature* **510**, 109–114 (2014).
6. A. Edgar, J. M. Ponciano, M. Q. Martindale, *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2122052119 (2022).
7. N. V. Whelan *et al*., *Nat. Ecol. Evol.* **1**, 1737–1746 (2017).
8. L. L. Moroz, A. B. Kohn, *Philos. Trans. R. Soc. London B Biol. Sci.* **371**, 20150041 (2016).
9. R. Hertwig, *Ueber den Bau der Ctenophoren* (G. Fischer, 1880).
10. M. Jager *et al*., *J. Exp. Zoolog. B Mol. Dev. Evol.* **316B**, 171–187 (2011).
11. L. S. Babonis *et al*., *Mol. Biol. Evol.* **35**, 2940–2956 (2018).
12. T. P. Norekian, L. L. Moroz, *J. Comp. Neurol.* **528**, 481–501 (2020).
13. A. Courtney *et al*., *J. Comp. Neurol.* **530**, 1438–1458 (2022).
14. M. Y. Sachkova *et al*., *Curr. Biol.* **31**, 5274–5285.e6 (2021).
15. R. A. Satterlie, J. F. Case, *Cell Tissue Res.* **193**, 87–91 (1978).
16. M. L. Hernandez-Nicaise, *Z. Zellforsch. Mikrosk. Anat.* **137**, 223–250 (1973).
17. S. L. Tamm, *Invertebr. Biol.* **133**, 1–46 (2014).
18. J. Von dem Rückenmark, Gerlach in *Handbuch der Lehre der Gewebe des Menschen und der Thiere*, S. Stricker, Ed. (Engelmann, 1871), pp. 665–693.
19. C. Golgi, *Sulla fina anatomia degli organi centrali del sistema nervosa* (Reggio Emilia Tipi di Stefano Calderini e Compagno, 1885).
20. M. Glickstein, *Curr. Biol.* **16**, R147–R151 (2006).
21. G. E. Palade, *Anat. Rec.* **118**, 335–336 (1954).
22. E. G. Gray, *J. Anat.* **93**, 420–433 (1959).
23. J. Ortiz *et al*., *Mol. Biol. Evol.* **40**, msad025 (2023).
24. A. Bethe, *Biol. Zbl* **15**, 140–145 (1895).
25. L. L. Moroz, *J. Exp. Biol.* **218**, 598–611 (2015).
26. M. L. Hernandez-Nicaise, *Nature* **217**, 1075–1076 (1968).
27. M. L. Hernandez-Nicaise, in *Microscopic Anatomy of Invertebrates. Volume 2: Placozoa, Porifera, Cnidaria, and Ctenophora*, F. W. Harrison, E. E. Ruppert, Eds. (Wiley-Liss, 1991), pp. 359–418.
28. M. Oren-Suissa, D. H. Hall, M. Treinin, G. Shemer, B. Podbilewicz, *Science* **328**, 1285–1288 (2010).
29. R. Giordano-Santini, C. Linton, M. A. Hilliard, *Semin. Cell Dev. Biol.* **60**, 146–154 (2016).
30. J. Z. Young, *Philos. Trans. R. Soc. London B Biol. Sci.* **229**, 465–503 (1939).
31. G. O. Mackie, *J. Cell Sci.* **101**, 119–131 (1960).
32. G. O. Mackie, C. L. Singla, S. A. Arkett, *J. Morphol.* **198**, 15–23 (1988).
33. F. Yu, O. Schuldiner, *Curr. Opin. Neurobiol.* **27**, 192–198 (2014).
34. T. E. Faust, G. Gunner, D. P. Schafer, *Nat. Rev. Neurosci.* **22**, 657–673 (2021).
35. L. J. Sullivan, D. J. Gifford, *J. Plankton Res.* **29**, 949–965 (2007).
36. I. S. Baiandina, *Morsk. Biol. Zh.* **5**, 105–108 (2020).
37. I. S. Baiandina, M. P. Kirin, O. V. Krivenko, *J. Sea Res.* **180**, 102152 (2022).
38. A. Sebé-Pedrós *et al*., *Nat. Ecol. Evol.* **2**, 1176–1188 (2018).
39. P. Burkhardt, G. Jékely, *Curr. Opin. Neurobiol.* **71**, 127–138 (2021).
40. J. M. Goaillard, E. Marder, *Annu. Rev. Neurosci.* **44**, 335–357 (2021).
41. Y. Moran, H. H. Zakon, *Genome Biol. Evol.* **6**, 2210–2217 (2014).
42. A. Senatore, H. Raiss, P. Le, *Front. Physiol.* **7**, 481 (2016).
43. G. A. Horridge, *Am. Zool.* **5**, 357–375 (1965).
44. M. L. Hernandez-Nicaise, *Tissue Cell* **6**, 43–47 (1974).

ACKNOWLEDGMENTS

We thank C. Le Goff and A. Jan for the phase contrast image of a 1-day-old ctenophore and R. Gohde for help with the Western blot experiments. High-pressure freezing, SBFSEM, and TEM imaging was done in the Oxford Brookes Centre for Bioimaging. **Funding:** This work was supported by the Michael Sars Centre core budget and the European Research Council Consolidator Grant (101044989, "ORIGINEURO") awarded to P.B. **Author contributions:** Conceptualization: P.B. and M.K.; Methodology: P.B., J.C., M.Y.S., and M.K.; Investigation: P.B., J.C., A.M., J.J.S.-A., E.-L.N., M.Y.S., and M.K.; Funding acquisition: P.B.; Supervision: P.B.; Visualization: P.B., J.C., A.M., L.D., M.Y.S., and M.K.; Writing – original draft: P.B., B.N., and M.K.; Writing – review and editing: P.B., J.C., B.N., M.Y.S., and M.K. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. **License information**: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade5645
Materials and Methods
Figs. S1 to S7
Table S1
References (*45*–*50*)
Movie S1
MDAR Reproducibility Checklist

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

Submitted 24 August 2022; accepted 24 February 2023
10.1126/science.ade5645

CONSERVATION

The value of private properties for the conservation of biodiversity in the Brazilian Cerrado

Paulo De Marco Jr.[1]*, Rodrigo A. de Souza[2], André F. A. Andrade[1], Sara Villén-Pérez[3], Caroline Corrêa Nóbrega[4], Luiza Motta Campello[5], Marcellus Caldas[6]

Areas set aside for conservation within private lands may be key to enhancing biodiversity-friendly landscapes. This conservation strategy should be especially effective in highly threatened regions that are poorly protected by public lands, such as the Brazilian Cerrado. Brazil's Native Vegetation Protection Law has included set-aside areas within private properties, but their relevance to conservation has not been evaluated. We assess whether private lands are contributing to biodiversity in the Cerrado, a global biodiversity conservation priority and major region for food production, where land use conflicts are often at odds with conservation objectives. We determined that private protected areas accommodate up to 14.5% of threatened vertebrate species ranges, which increases to 25% when considering the distribution of remaining native habitat. Moreover, the spatial spread of private protected areas benefits a large number of species. Ecological restoration of private protected lands would improve the benefits of this protection system, especially in the Southeastern Cerrado, where a large economic hub meets a threat hotspot.

Protected areas are the cornerstones for the long-term conservation of biodiversity. They cover about 15% of the terrestrial surface and 7.3% of the ocean surface (*1*), and global analyses show that they are still insufficient to protect biodiversity (*2*). The need for complementary strategies to join or reinforce protection networks is especially urgent to deal with the lack of connectivity due to habitat fragmentation. A promising approach is to make landscapes that are now occupied by economic activity more "biodiversity-friendly" (*3*). Biodiversity-friendly landscapes seek to preserve habitat patches in human-dominated areas to favor the persistence of native species (*3*), including beneficial animals such as pollinators, predators, and fruit dispersers (*4*). Most human-dominated areas are under private ownership, and this represents a large proportion of global land, varying from 44.2% in Brazil (*5*) to 52% in Germany, 75% in the United States (excluding Alaska) (*6*), and nearly 80% in the United Kingdom and Spain (*7*). Thus, improving the biodiversity-friendliness of private landholdings could amplify the benefits of the existing protection system by increasing both the total habitat available and the connectivity among remaining habitats (*8*), ensuring population persistence and richer biodiversity (*9*). However, conservation prioritization efforts have often overlooked the role of private lands while focusing on public protection networks (*10*). Here, we provide an evaluation of the relevance of set-aside areas of private land in one of the most important and vulnerable worldwide arenas for the conflict between food production and biodiversity conservation: the Brazilian Cerrado (*11*). In addition, we present potential scenarios for restoration priorities to optimize the protection of 103 threatened terrestrial vertebrates in the biome.

Biodiversity-friendly landscapes must be designed to increase connectivity among habitat patches and to maintain sufficient habitat to assure the long-term persistence of biodiversity (*8*, *12*). The strategy to implement this approach varies widely across the land-sharing and -sparing continuum and from voluntary to mandatory actions implemented by different countries (*13*). In parts of Australia, for example, there is a well-established model in which some rights are voluntarily relinquished in favor of conservation under a binding legal agreement and in exchange for economic incentives (*14*). Similar approaches are also found in the United States and Canada (*15*). In Latin America, the scheme is similar but shows a larger participation of nongovernmental organizations (NGOs) in land purchase for conservation, especially in Costa Rica, Ecuador, Argentina, and Chile (*16*). Otherwise, mandatory regulations to protect a portion of every rural property may represent a meaningful strategy for conservation in largely human-dominated landscapes. One of the best-established examples of this policy was implemented in the Brazilian Forest C[illegible] almost a century ago (federal decree no. 23.[illegible] 1934 and federal law no. 4.771/1965). It was originally conceived under a utilitarian view that focused on the importance of vegetation to water resources, soil fertility, and wood storage within rural properties (*17*). Even so, the Forest Code has important implications for biodiversity conservation today.

To enforce the sustainable use of natural resources, the Brazilian Forest Code required rural owners to select patches to become legal reserves within their rural properties. The legal reserve area varies between 20% (criteria for most of Brazil) to 80% (for the Amazon) of the property area. By contrast, the location of permanent protection areas is not eligible because they are designed to protect geological stability (e.g., topographic slope higher than 45°) and water resources (e.g., areas around streams, rivers, and springs). The existence of legal reserves and permanent protection areas has suffered long-standing pressure from political and economic sectors, with reiterated attempts to change this legislation during its history. As a consequence, some changes were implemented in the 2012 Native Vegetation Protection Law (federal law no. 12.651/2012) to maintain the general definition for existent categories but allow new deforestation, mainly in the Cerrado biome. The 2012 Forest Code also demands that landowners provide georeferenced information about land uses and protected areas in their rural properties through the Rural Environmental Registry [Cadastro Ambiental Rural (CAR)]. Here, we take the opportunity created by CAR to analyze the spatial distribution of all private protected areas from 684,942 rural properties registered in the Cerrado biome to assess its potential value for the conservation of threatened vertebrate species and to predict the potential benefits of fully restoring set-aside areas. The Cerrado is a wooded grassland, or savanna, covering about 20% of Brazil. It is home to distinctive and threatened species, such as the maned wolf (*Chrysocyon brachyurus*), the giant anteater (*Myrmecophaga tridactyla*), and the highly endangered blue-eyed ground dove (*Columbina cyanopis*). By 2018, natural vegetation loss reached 90 million ha—45% of Cerrado area—mostly in private property (*17*, *18*).

All analyses are based on the overlap between model predictions of species' ranges and the proportion of private protected areas with a 10-km–by–10-km cell from CAR's polygons. We used conservative estimates of species' ranges based on advanced ecological niche modeling techniques that account for dispersal constraints (*19*). Landscape-cell relevance to conservation was estimated by giving higher weight to smaller-ranged species, which are proportionally more affected by habitat

[1]Departamento de Ecologia, Universidade Federal de Goiás, Goiânia, GO 74 690-720, Brazil. [2]Centro Nacional de Informações Ambientais (CENIMA), Instituto Nacional de Meio Ambiente e Recursos Naturais Renováveis (IBAMA), SCEN Ibama, Ed. Sede, Bloco F, Brasília, DF 70818-900, Brazil. [3]Universidad de Alcalá, GloCEE – Global Change Ecology and Evolution Research Group, Departamento de Ciencias de la Vida, 28805, Alcalá de Henares, Madrid, Spain. [4]Aliança da Terra, Av. das Indústrias, 601, Quadra 151 Lote 47 Sala 301, Santa Genoveva, Goiânia, GO 74670-600, Brazil. [5]Universidade Brasília, Instituto de Geociências, Campus Universitário Darcy Ribeiro ICC, Ala Central, Brasília-DF 70910-900, Brazil. [6]Department of Geography and Geospatial Sciences, Kansas State University, Manhattan, KS 66502, USA.
*Corresponding author. Email: pdemarco@ufg.br

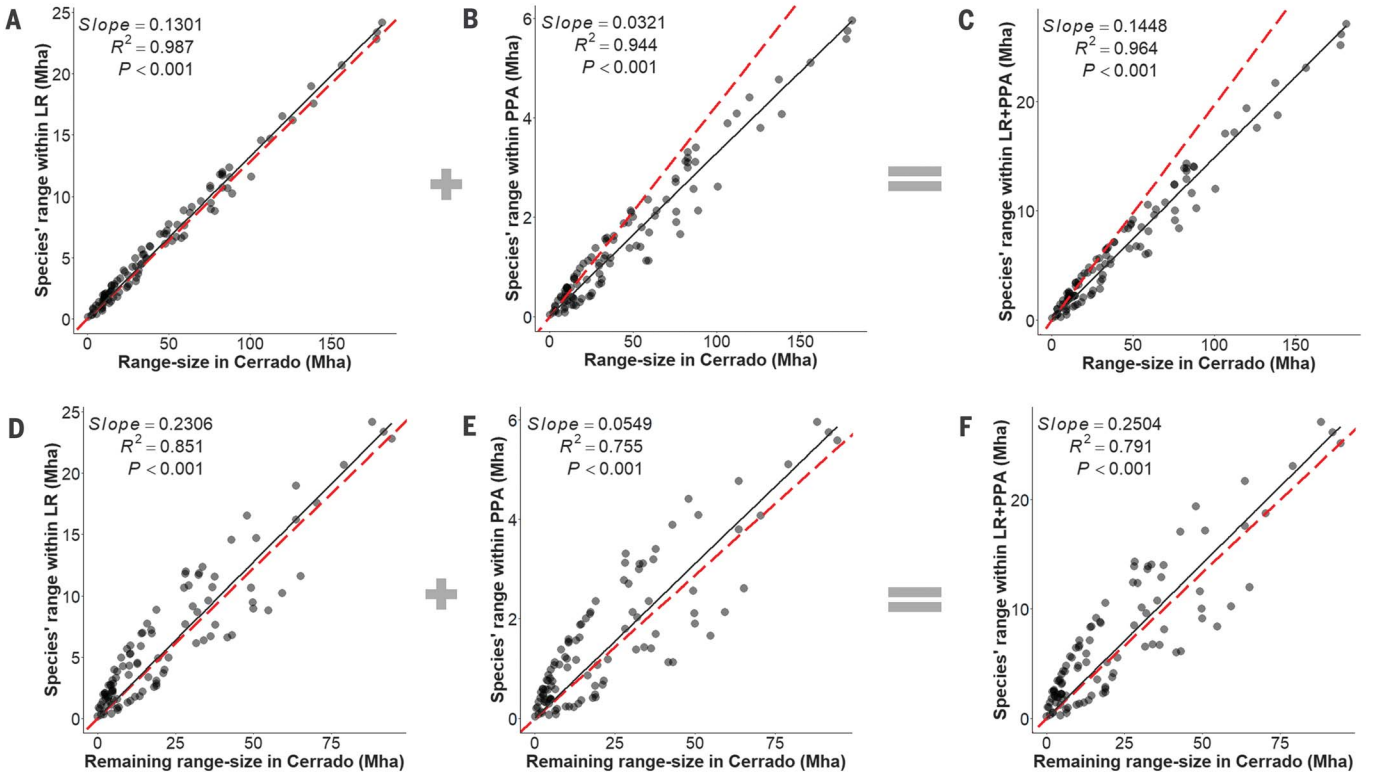


Fig. 1. The proportion of distributional range of threatened vertebrate species that falls within legal reserves and permanent protection areas in relation to its range size in the Cerrado biome. (**A** to **F**) Historical species' ranges that overlap legal reserves (LR) (A) and permanent protection areas (PPA) (B), as estimated by the regression through the origin (black lines), are 13.01 and 3.21%, respectively. The null expectation (red dashed line) is that each species overlaps both categories according to the proportion of those classes across the whole Cerrado area (12.80% for legal reserves and 4.20% for permanent protection areas). The predicted mean proportion of species distribution within total private protected areas (C) is 14.48%. The portion of species' ranges with available remaining habitat that overlaps with legal reserves (D) and permanent protection areas (E) is larger (black lines; 23.06 and 5.49%, respectively) and presents a larger interspecific variation but is close to the null expectation (red dashed line). After discounting habitat loss, the predicted mean proportion of species distributions within total private protected areas (F) is high (25.04%) and close to the null expectation. The estimate of the overlap (slope of black line) is indicated in each plot, together with the R^2 and its statistical significance.

loss. Moreover, we assumed that species can persist in the private protected patches independently of their size, isolation, or type of surrounding matrix because species-specific sensitivity to these variables is unknown for most of the species we evaluated (*20*). We start by assuming that private protected areas are fully restored, though a considerable part of those set-aside areas is, at present, not well preserved and suffers from human interference (*11*). This assumption is relevant because their restoration is mandatory even under the current Forest Code (*21*). Thus, our analysis assesses the conservation value if restoration is properly implemented. Finally, we explore this further by indicating where restoration will bring higher benefits to biodiversity.

We show that an average of 13.01% of the range of threatened species falls within legal reserves [slope of the regression of range within legal reserves and total species' range in Cerrado, forced through the origin; coefficient of determination (R^2) = 0.987; Fig. 1A]. This value is only slightly higher than the null expectation, which is the percentage cover of legal reserves in Cerrado (12.86%). The similarity to the null expectation suggests that these areas are representative of the environmental variation in the Cerrado favoring better representation of species' distribution ranges of threatened vertebrates. This hypothesis was supported both by the frequency distribution of public and private areas in relation to first climatic principal components analysis (PCA) and by the overlap of the entire environmental variation of the Cerrado (figs. S2-1 and S2-2). The mean proportion of the predicted species ranges that fall within permanent protection areas is lower (3.21%; the slope of the regression; R^2 = 0.944; Fig. 1B). This prediction is also slightly lower than the null expectation that species overlap is only determined by the proportion of this category in the whole Cerrado (4.26%). Applying the same analysis for the current public (federal and state levels) protected area system in the Cerrado shows that only 13.78% of species' ranges is protected, slightly lower than the null expectation based on the coverage of public protected network in this region (15.30%; fig. S2-3). Private protected land is more evenly distributed across the Cerrado and thus is better suited to benefit a larger number of species. This is a desirable quality that contrasts with the public protection system, which is biased toward less-favorable lands for agriculture and does not represent the distribution of most threatened species (*10*). Otherwise, private land may represent an even higher proportion of threatened species' ranges if we restrict the analysis to the available remnants of native vegetation in the Cerrado. In that case, the predicted mean proportion of species' ranges within legal reserves is 23.06% and, for permanent protection areas, is 5.49% (R^2 = 0.851 for legal reserves and R^2 = 0.756 for permanent protection areas; Fig. 1, D and E). The general agreement to the null expectation still holds, but there is an increased scattering that suggests higher interspecies variation in their level of protection. This variation supports the nonrandom distribution of habitat loss in the

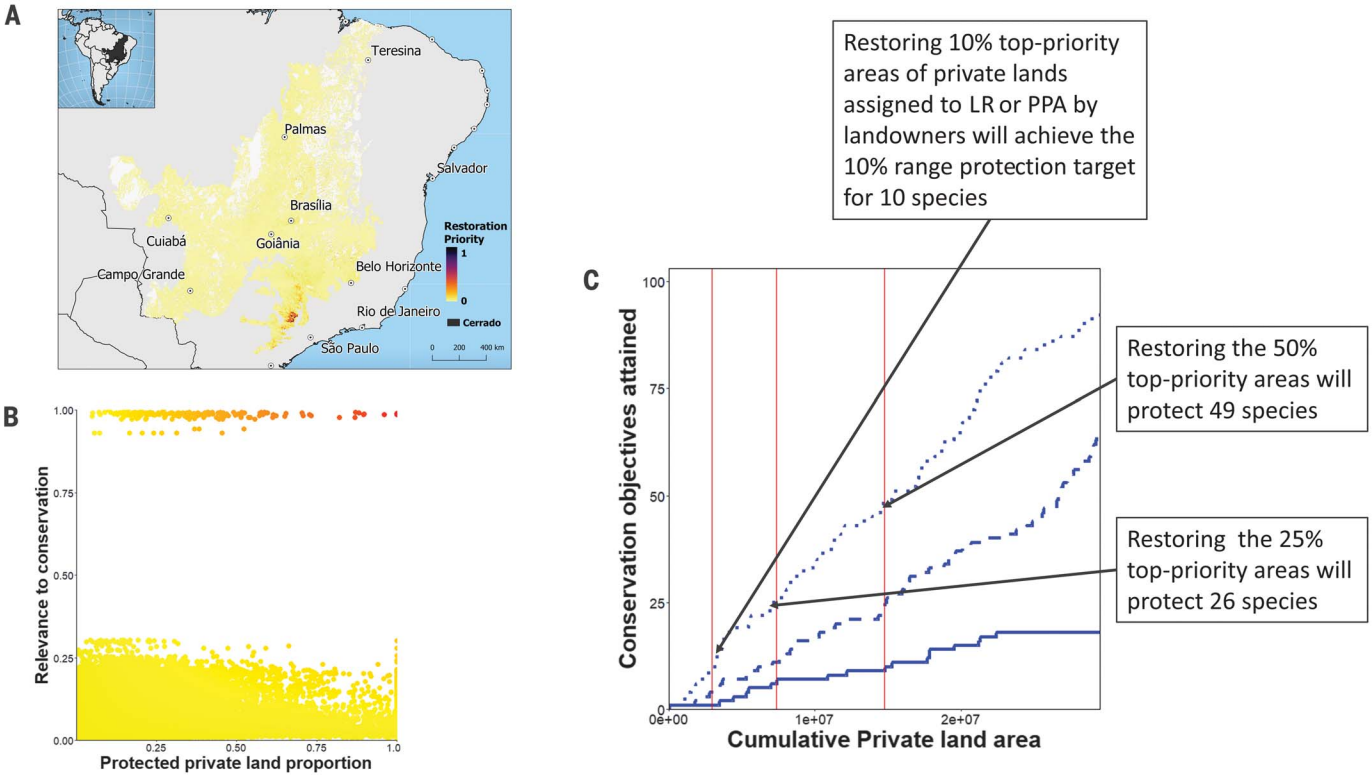


Fig. 2. Priority areas for the restoration of private lands based on their contribution for threatened species conservation. (**A** and **B**) Identification of restoration priorities within private protected areas (A) and their distribution in the Cerrado (B) based on their relevance to threatened species conservation, considering both the biodiversity relevance index and private protected area size. (**C**) Prediction of the conservation milestones that would be reached as an increased number of private protected areas are restored after the prioritization established in (A). Conservation milestones are determined in terms of the number of species that would benefit and the percentage of their range that would be preserved. The dotted, dashed, and solid blue lines represent the conservation targets of 10, 15, and 20% of species range size, respectively.

Cerrado, which causes different levels of exposure among species (*20*).

The combined effect of legal reserves and permanent protection areas may protect up to 14.48% of species' distribution ranges in the Cerrado (R^2 = 0.964; Fig. 1C). This prediction increases to 25.04% after discounting for current habitat loss outside private protected assigned areas (R^2 = 0.791; Fig. 1F). This coverage is consistently higher than that expected from the distribution of total private protected area in the Cerrado (9.7%) and the remaining habitat in these protected areas (19.7%). Our results show that private lands may protect nearly 25% of the remaining climatically suitable habitats for threatened vertebrates in the Cerrado, so that its relevance for conservation is much higher than is now assumed. This also evidences the importance of habitat restoration within set-aside private lands, which is expected to occur under the current legal system. First, an increase in habitat amount due to restoration of private protected areas is expected to increase species' population sizes and reduce their risk of extinction (*22*). In addition, an increase in connectivity among remaining habitats is expected to favor species' dispersion and persistence in the landscape (*3*).

Different private lands are not equally important to species conservation across the Cerrado region. The spatial variance of the relevance to the conservation index (Fig. 2, A and B) shows that a small set of 192 cells, mostly distributed in the most highly affected São Paulo state, has a disproportional importance to conservation. Those areas are part of the distribution of at least 70 small-ranged species and still bear a relatively large amount of protected land to restore within its predicted distribution. The entire set retains nearly 145,000 ha of protected private land, with an estimated cost of restoration not higher than $60 million based on assisted regeneration methods, which is only 0.02% of the exports value of the Brazilian agribusiness sector in 2021 (https://indicadores.agricultura.gov.br/agrostat/index.htm). An analysis of potential scenarios for the prioritization of areas within set-aside private lands shows that after ordering all Cerrado cells according to their relevance for conservation, the cumulative protected private land area points to a positive scenario. Restoration of the 10% top-priority cells will achieve the goal of 10% range protection for 10 threatened species, a 25% restoration will achieve the same goal for 26 species, and a 50% restoration will achieve the same goal for 49 species. More ambitious conservation targets (15 or 20% of species' range protection; Fig. 2C) are attained only for a small number of species or under optimistic restoration scenarios. For instance, restoring 75% of the protected private land would protect 20% of the range of 17 species, which includes many small-ranged species that are well represented in the Cerrado biome. We argue that an explicit policy to assure restoration will return clear conservation benefits based on those scenarios.

Our results support the importance of private lands to the protection of threatened Cerrado species. They indicate that restoring private protected areas is an important conservation goal that deserves special funds and attention. We show that, at least for the conservation of threatened terrestrial vertebrate species, it is possible to devise a prioritization scheme to guide the restoration efforts of those areas. In addition, private protected area restoration would also have direct effects on

essential ecosystem services. Based on recent calculations of carbon storage in Cerrado areas (*23*), we made a conservative estimate that shows that the restoration of private protected areas could capture 12×10^6 tonnes of carbon, which is a substantial contribution toward the 2°C climate target (*24*). Restoring private set-aside areas may also improve pollination services for major crops, such as soybean, and other relevant croplands of fruits and vegetables. Although those services are not always recognized by private owners (*25*), private land protection still carries the possibility of increasing the visibility of its benefits, thereby boosting restoration efforts. The choice to dedicate land and resources for biodiversity conservation is political and influenced by the value that people place on biodiversity (*26*). Conservation in private lands may increase the perception of ecosystem services and promote willing-to-conserve attitudes (*27*), thus reinforcing society's positive view of biodiversity conservation.

REFERENCES AND NOTES

1. UN Environment World Conservation Monitoring Centre (UNEP-WCMC), International Union for Conservation of Nature (IUCN), National Geographic Society (NGS), "Protected Planet Report 2018" (UNEP-WCMC, IUCN, and NGS, 2018).
2. D. Leclère *et al*., *Nature* **585**, 551–556 (2020).
3. F. P. L. L. Melo, V. Arroyo-Rodríguez, L. Fahrig, M. Martínez-Ramos, M. Tabarelli, *Trends Ecol. Evol.* **28**, 462–468 (2013).
4. M. Duru *et al*., *Agron. Sustain. Dev.* **35**, 1259–1281 (2015).
5. G. Sparovek *et al*., *Land Use Policy* **87**, 104062 (2019).
6. Headwaters Economics, "A profile of land use" (Headwaters Economics, 2018).
7. L. A. Powell, *Anim. Biodivers. Conserv.* **35**, 295–306 (2012).
8. V. Arroyo-Rodríguez *et al*., *Ecol. Lett.* **23**, 1404–1420 (2020).
9. J. I. Watling *et al*., *Ecol. Lett.* **23**, 674–681 (2020).
10. U. Oliveira *et al*., *Sci. Rep.* **7**, 9141 (2017).
11. R. Rajão *et al*., *Science* **369**, 246–248 (2020).
12. L. Fahrig, *J. Biogeogr.* **40**, 1649–1663 (2013).
13. S. Kamal, M. Grodzińska-Jurczak, G. Brown, *J. Environ. Plann. Manage.* **58**, 576–597 (2015).
14. C. L. Archibald *et al*., *Environ. Sci. Policy* **115**, 99–107 (2021).
15. J. Owley, A. R. Rissman, *Land Use Policy* **51**, 76–84 (2016).
16. Environmental Law Institute, *Legal Tools and Incentives for Private Lands Conservation in Latin America: Building Models for Success* (Environmental Law Institute, 2003).
17. B. Soares-Filho *et al*., *Science* **344**, 363–364 (2014).
18. Projeto MapBiomas, Coleção 7 da Série Anual de Mapas de Cobertura do Uso do Solo do Brasil (2022); https://mapbiomas-br-site.s3.amazonaws.com/Estatísticas/1_-_TABELA_GERAL_COL7_MAPBIOMAS_BIOMAS_UF_FINAL.xlsx.
19. P. Mendes, S. J. E. Velazco, A. F. A. de Andrade, P. De Marco, *Ecol. Modell.* **431**, 109180 (2020).
20. P. De Marco Jr. *et al*., *Biodivers. Conserv.* **29**, 1637–1658 (2020).
21. K. de Mello *et al*., *Environ. Sci. Policy* **120**, 1–10 (2021).
22. J. J. O'Grady, D. H. Reed, B. W. Brook, R. Frankham, *Biol. Conserv.* **118**, 513–520 (2004).
23. B. Zimbres *et al*., *For. Ecol. Manage.* **499**, 119615 (2021).
24. Y. M. Wei *et al*., *Nat. Clim. Chang.* **11**, 112–118 (2021).
25. F. P. Lima, R. P. Bastos, *Ecosyst. Serv.* **40**, 101029 (2019).
26. A. R. Whittaker, *People Nat.* **2**, 450–467 (2020).
27. F. Pereira Lima, R. Pereira Bastos, *Ecosyst. Serv.* **44**, 101121 (2020).
28. P. De Marco Jr. *et al*., Species distribution data and distribution of private and public protected areas in Cerrado, Brazil. Dryad (2023); https://doi.org/10.5061/dryad.7pvmcvdz5.

ACKNOWLEDGMENTS

This work is dedicated to G. Fonseca, who spent a lifetime dedicated to biodiversity conservation in Brazil and worldwide. We thank E. Bruna and G. Fonseca for their invaluable contributions to an early version of this manuscript. **Funding:** P.D.M. is continuously supported by a CNPq productivity grant (310547/2020-2). S.V.-P. was supported by Comunidad de Madrid and Universidad de Alcalá (2017-T2/AMB-6035; 2022-T1/AMB-237), Sustainability Victoria (CM/JIN/2019), and Ecological Transition and Demographic Challenge (2020/00085/001). M.C. is supported by National Science Foundation NSF-BCS 2117533: Agri-environmental Conservation Incentives in the Extreme Wildfire Context of the U.S. Southern Plains. A.F.A.A. is supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq - 165174/2020-0). **Author contributions:** Conceptualization: P.D.M., R.A.S., C.C.N., M.C.; Methods (geographical data management): R.A.S., L.M.C.; Methods (ecological models): A.F.A.A., R.A.S., S.V.-P., C.C.N.; Data analysis: P.D.M., A.F.A.A.; Writing – first draft: P.D.M., C.C.N., R.A.S.; Writing – review: P.D.M., M.C., S.V.-P. **Competing interests:** The authors declare that they have no competing interests. **Data and availability:** All data (species distribution models and spatialized private and public conservation distribution) are available from the Dryad dataset repository (*28*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq7768
Material and Methods
Figs. S1 and S2
References (*29*–*49*)
MDAR Reproducibility Checklist

Submitted 8 June 2022; accepted 1 March 2023
10.1126/science.abq7768

STRUCTURAL BIOLOGY

Structure of the R2 non-LTR retrotransposon initiating target-primed reverse transcription

Max E. Wilkinson[1,2,3,4,5], Chris J. Frangieh[1,2,3,4,5,6], Rhiannon K. Macrae[1,2,3,4,5], Feng Zhang[1,2,3,4,5]*

Non–long terminal repeat (non-LTR) retrotransposons, or long interspersed nuclear elements (LINEs), are an abundant class of eukaryotic transposons that insert into genomes by target-primed reverse transcription (TPRT). During TPRT, a target DNA sequence is nicked and primes reverse transcription of the retrotransposon RNA. Here, we report the cryo–electron microscopy structure of the *Bombyx mori* R2 non-LTR retrotransposon initiating TPRT at its ribosomal DNA target. The target DNA sequence is unwound at the insertion site and recognized by an upstream motif. An extension of the reverse transcriptase (RT) domain recognizes the retrotransposon RNA and guides the 3′ end into the RT active site to template reverse transcription. We used Cas9 to retarget R2 in vitro to non-native sequences, suggesting future use as a reprogrammable RNA-based gene-insertion tool.

Non–long terminal repeat (non-LTR) retrotransposons are the most abundant class of mobile genetic element (MGE) in the human genome, primarily represented by the LINE-1 and SINE (or Alu) long and short interspersed nuclear elements, respectively (*1*). Despite their prevalence and contribution to genetic diversity and dysregulation through mutagenicity and recombination (*1*–*3*) and their prospective use as gene insertion tools, there is much left to understand about the mobility mechanisms of non-LTR retrotransposons (*4*). Pioneering research on the *Bombyx mori* (silk moth) R2 element (R2Bm), which selectively inserts into the 28*S* ribosomal RNA (rRNA) gene, has contributed substantially to our understanding of this type of MGE (*5*). R2, like all non-LTR retrotransposons, encodes an open reading frame (ORF) with DNA binding, endonuclease, and reverse transcriptase activities (Fig. 1A). The endonuclease domain (restriction-like endonuclease, RLE) nicks the target DNA, and the reverse transcriptase domain uses the exposed 3′ end from the nick to prime reverse transcription of the R2 RNA, resulting in a new genomic copy of the R2 element (Fig. 1B) (*6*, *7*). This process is called target-primed reverse transcription (TPRT) and is characteristic of non-LTR retrotransposons and their group II intron ancestors (*8*, *9*). The nicked strand that primes reverse transcription is referred to as the bottom strand. Complementarity between the bottom strand and the 3′ end of the R2 RNA (3′ homology) is not required to initiate reverse transcription (*10*). Non-LTR retrotransposons are specific for reverse transcribing their own RNA; for R2, this specificity requires an element in the 3′ untranslated region (3′UTR), but the precise motif has not been located (*11*). It is also unclear how R2 specifically recognizes the 28*S* rRNA target gene, or how DNA nicking is coupled to reverse transcription within the same protein. To address these questions, we solved a cryo–electron microscopy (cryo-EM) structure of R2Bm initiating TPRT at the 28*S* rRNA gene using its own 3′UTR. The structure reveals an extensive interface with the target DNA, a small core region of the 3′UTR required for TPRT, and shows that R2Bm can be engineered to reprogram its insertion site.

[1]Howard Hughes Medical Institute, Cambridge, MA 02139, USA. [2]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [3]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [4]Department of Brain and Cognitive Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [5]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [6]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
*Corresponding author. Email: zhang@broadinstitute.org

Reconstitution and cryo-EM structure of an R2 TPRT complex

We overexpressed R2Bm in *Escherichia coli* and purified it to apparent homogeneity (fig. S1). The purified protein was active in vitro, reproducing previously found biochemical activities, including RNA-stimulated nicking of the target DNA bottom strand, site-specific TPRT when supplied with in vitro–transcribed 3′UTR RNA, and low levels of template jumping (Fig. 1C) (*6*, *12*). It is unclear if 3′ homology is required for TPRT in vivo; however, consistent with previous findings, we found that downstream sequences of up to 10 nucleotides (nt) do not inhibit activity in vitro (Fig. 1C) (*10*). Sequencing of TPRT junctions confirmed that homology-mediated TPRT is more likely to initiate reverse transcription at the 3′ end of the 3′UTR rather than skipping bases or inserting untemplated nucleotides (fig. S2) (*10*). To assemble a complex stalled during initiation of TPRT, we incubated R2Bm with target DNA, 3′UTR RNA, and the chain-terminator nucleotide 2′,3′-dideoxythymidine (ddT), which mimics the first nucleotide incorporated in the TPRT reaction (dT) but does not allow further elongation. Purified TPRT complexes contai[illegible] stoichiometric amounts of R2Bm, 3′UTR R[illegible] and target DNA with >99% of the bottom strand nicked (fig. S1). Initial attempts at cryo-EM imaging failed owing to the preferred orientation and flexibility of the complex. To overcome these issues, we used a carbon support on the cryo-EM grid and added 5 nt of downstream 28*S* rRNA sequence to the 3′ end of the 3′UTR RNA to stabilize the complex by forming a primer-template duplex with the target DNA bottom strand. With these modifications, we obtained a cryo-EM reconstruction of the R2 TPRT complex at 3.1-Å resolution (Fig. 1D, figs. S3 and S4, and table S1).

The core of the R2Bm protein is a reverse-transcriptase (RT) domain similar to that of group II intron RTs (*13*), followed by a C-terminal α-helical thumb domain and preceded by a characteristic N-terminal extension domain (NTE0) implicated in template switching (*14*), but the R2Bm RT includes a further N-terminal extension (NTE-1) that binds the 3′UTR RNA (Fig. 1, E and F) (*15*). Preceding the NTE-1 element are two DNA binding domains: the N-terminal C2H2 zinc finger domain (N-ZnF) and a Myb domain. C-terminal to the thumb domain lies an α-helical linker domain that packs against the thumb, followed by a CCHC zinc-finger domain (ZnF) conserved in many LINE ORFs (*4*). The ZnF then links to the C-terminal RLE domain, which cleaves the target DNA. This domain arrangement closely resembles that of Prp8 (*13*, *16*, *17*), the core protein of the spliceosome, underscoring the close relationship between pre-mRNA splicing and retrotransposons.

There are several key interactions between the R2Bm protein, 3′UTR RNA, and target DNA (Fig. 1, E and F). The two strands of the target DNA separate around the ZnF domain, with the bottom strand feeding into the RLE active site where the scissile phosphate remains bound, while the top strand snakes along the opposing surface of the RLE. The RT active site contains a heteroduplex formed by the nicked bottom strand of the target DNA (5′ to the cleavage site) and the 5 nt of 28*S* rRNA homology extension beyond the 3′UTR RNA (Fig. 1G). This target heteroduplex is surrounded by residues important for RT activity (*18*), and the cryo-EM density shows incorporation of the ddT chain terminator nucleotide into the bottom strand (Fig. 1H). The 5′ end of the bottom strand remains base-paired to the top strand as it leaves the RLE, and this downstream DNA region has weak cryo-EM density, suggesting that it is not tightly bound by R2Bm. The 248-nt 3′UTR RNA is mostly not resolved in the cryo-EM density except for a core 40-nt region, which wraps around the NTE-1 α helix of R2Bm and the 3′ end of which is guided into the RT active site via the NTE0 domain.

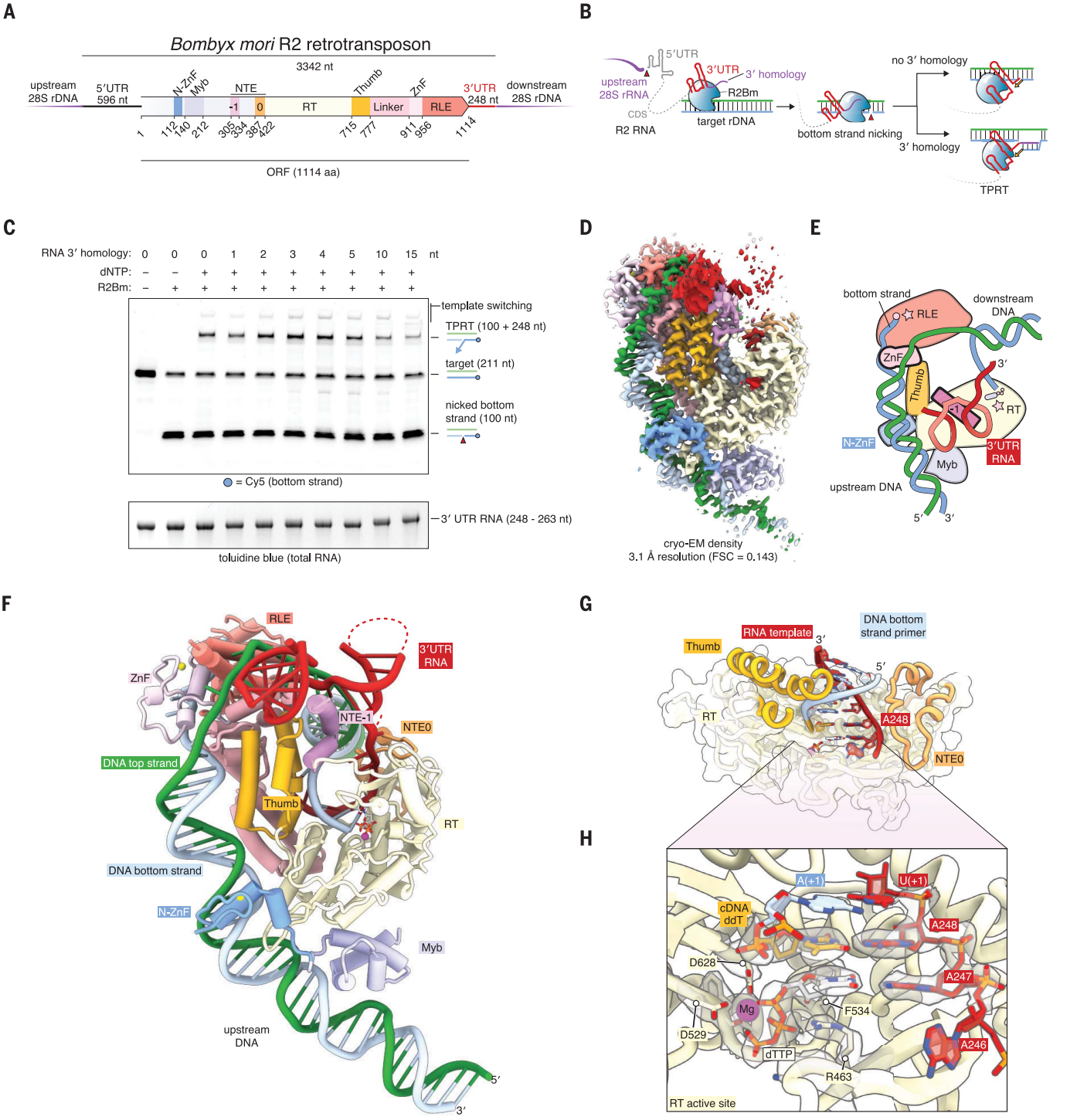


Fig. 1. Cryo-EM structure of the R2Bm retrotransposon. (**A**) Domains of the R2Bm retrotransposon. ZnF, zinc finger; NTE, N-terminal extension; RT, reverse transcriptase; RLE, restriction-like endonuclease. (**B**) Schematic of target-primed reverse transcription (TPRT). (**C**) Denaturing gel of in vitro TPRT reactions on a labeled 211-bp 28*S* DNA target. The same gel was visualized by Cy5 fluorescence and toluidine blue staining. (**D**) Cryo-EM density of the R2Bm TPRT complex. (**E**) Cartoon of the cryo-EM structure. Stars represent active sites. (**F**) Atomic model for the R2Bm TPRT complex. (**G**) Reverse transcriptase domain and template–primer duplex. (**H**) Reverse transcriptase active site. Cryo-EM density is shown as a gray transparent surface.

R2Bm recognizes a sequence motif upstream of the cleavage site

The target 28*S* DNA sequence has extensive interactions with R2Bm (summarized in Fig. 2A). Upstream bases from –38 to –7 and downstream bases from +6 to +21 are respectively paired, whereas the 11 base pairs from –6 to +5 are melted around the RLE domain (bases are numbered relative to the bottom-strand cleavage site). The upstream DNA has a 40° bend and binds along the surface of the RT, linker, and thumb domains in a manner similar to that of the DNA in a recent group IIC intron maturase structure (Fig. 2B and figs. S5 and S6) (*19*). Many of the contacts between

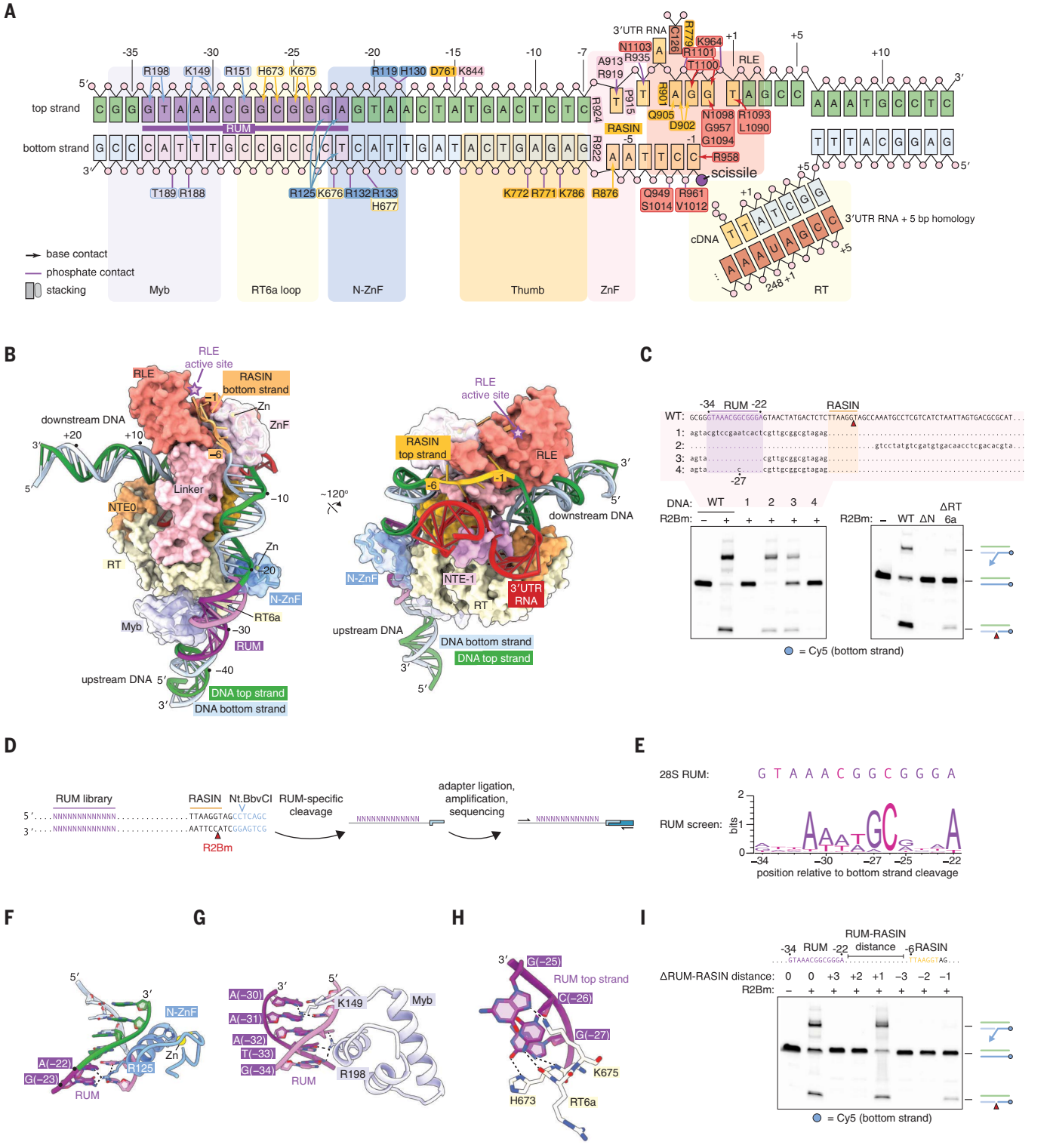


Fig. 2. Target DNA recognition upstream of the R2 cleavage site. (**A**) Schematic of interactions with the target DNA. Bases are numbered relative to the bottom-strand cleavage site. Positions of protein domains are shown by shaded rectangles. (**B**) Structure of R2Bm around the upstream DNA sequences. (**C**) Effect of upstream DNA mutations on target cleavage. The schematic shows the sequences of five DNA sequences tested in top-strand sense; dots represent bases identical to those of wild type. Red triangle, bottom-strand cleavage site. Denaturing gels show in vitro TPRT reactions on labeled 211-bp 28S DNA targets. ΔN, deletion of N-terminal N-ZnF and Myb domains. ΔRT6a, deletion of residues 672 to 677 (DGHRKK) of the RT6a loop. (**D**) Screen for identifying active RUM sequences. Nicking sites of R2Bm and the restriction endonuclease Nt.BbvCI are shown by triangles. (**E**) Sequence logo for sequences enriched in the RUM screen. (**F** to **H**) Details of interactions between the target DNA and the N-ZnF, Myb, and RT6a loop. (**I**) Effect of altering the distance between the RUM and RASIN motifs. Denaturing gel shows in vitro TPRT reactions on labeled 211-bp 28S DNA targets. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; D, Asp; E, Glu; F, Phe; H, His; K, Lys; L, Leu; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; and Y, Tyr.

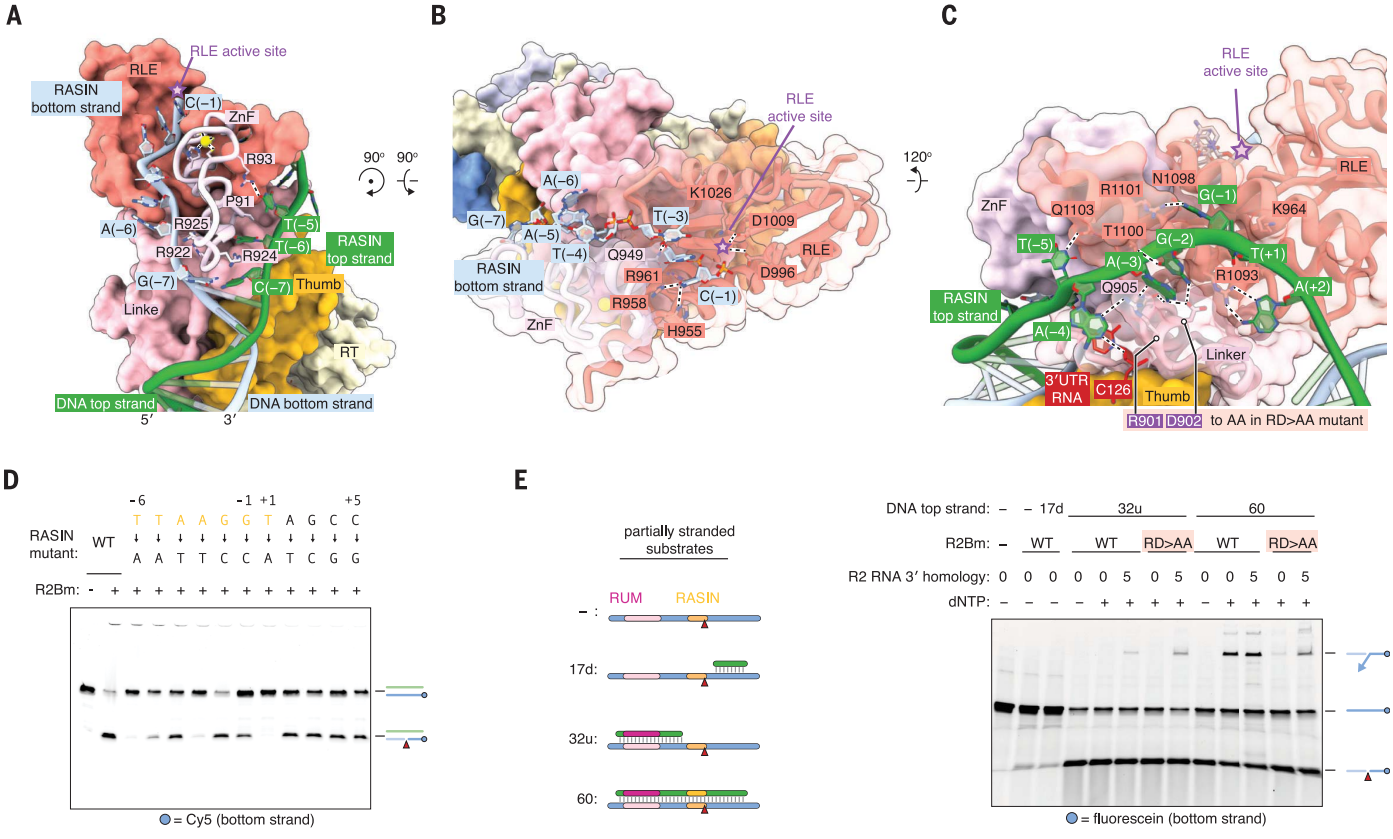


Fig. 3. Target DNA recognition at the R2 cleavage site. (**A**) Interactions of the top and bottom strands of the target DNA with the ZnF domain of R2Bm. Star, RLE active site. (**B**) Interactions of the DNA bottom strand with the RLE domain. (**C**) Interactions of the DNA top strand with the RLE domain. Residues mutated in the RD>AA mutant are highlighted. (**D**) RASIN sequence requirements for bottom-strand cleavage. The labeled 211-bp 28*S* DNA targets were incubated with R2Bm and 3′UTR RNA in the absence of deoxynucleotide triphosphates (dNTPs). The reactions were analyzed with a denaturing gel. Mutations are notated in top-strand sense, but both strands were mutated. (**E**) Denaturing gel showing R2Bm cleavage and TPRT activity on partially stranded substrates. Reactions contained a fluorescein-labeled 76-nt bottom strand. Reactions as indicated also contained 17 nt of downstream top-strand sequence (17d), 32 nt of upstream top strand sequence (32u), or 60 nt of top-strand sequence fully complementary to the bottom strand spanning the upstream and downstream regions. RD>AA, R2Bm R901A D902A.

R2Bm and the DNA are via the phosphate backbone, suggesting that they are not sequence specific. Based on the structure, however, we predicted that two regions are key for sequence-specific DNA recognition by R2Bm: a 13-bp upstream motif from −34 to −22, which is bound by the N-terminal N-ZnF and Myb domains, and the 7 bp from −6 to +1, which are bound by the RLE (Fig. 2A). We term these regions the retrotransposon upstream motif (RUM) and retrotransposon-associated insertion site (RASIN), respectively.

Consistent with the importance of the RUM region for R2 activity, mutating the entire upstream sequence between −38 to −7 eliminated bottom-strand cleavage, whereas mutating the downstream sequences between +6 and +37 preserved wild-type levels of bottom-strand cleavage and TPRT (Fig. 2C) (*20*). Adding just the 13-bp RUM region to the upstream mutant at positions −34 to −22 restored near–wild-type activity, whereas a point mutant RUM (G_{-27} to C) did not rescue activity (Fig. 2C). This region of the target was strongly protected in a previous deoxyribonuclease (DNase) footprinting assay (*21*). To systematically determine the importance of each base within the RUM, we performed an R2 cleavage assay on a DNA target with the upstream region (−38 to −7) mutated and the RUM (−34 to −22) replaced with a 13N library (Fig. 2D). Sequencing of cleaved targets revealed a consensus RUM sequence A_{-31}WWWGCNNNA$_{-22}$, where W is A/T and N is any nucleotide, with minor preferences in other positions (Fig. 2E). This consensus is a close match to the wild-type 28*S* sequence A_{-31}A**CG**GCGGGA$_{-22}$, with the differences shown in bold.

The RUM is recognized by three domains: N-ZnF, Myb, and an R2-specific insertion "6a" in the RT domain between motifs 6 and 7 (Fig. 2B and fig. S7). The N-ZnF has the classical C2H2 fold with a zinc ion coordinated between an α helix and a β hairpin, but unusually the α helix binds in the widened minor groove of the DNA from bases −18 to −23 instead of the typical major groove (Fig. 2F and fig. S6) (*22*). The preference for A at base −22 in the RUM is likely due to N-ZnF Arg^{125}, which hydrogen bonds with the minor-groove–facing side of the A–T base pair (Fig. 2F). The Myb domain forms a typical three-helix bundle, with the third helix bound in the major groove from bases −31 to −34 (*22*) while its linker to N-ZnF engages with base −30 (Fig. 2G). This is reminiscent of other Myb–DNA structures, including telomere-interacting protein Rap1 (*23*). The Myb domain recognizes the A at base −31 through hydrogen bonds with Lys^{149} (Fig. 2G). Although Arg^{198} contacts bases at positions −33 and −34, these contacts appear not to be sequence specific, as the RUM screen showed only weak sequence preferences in this region (Fig. 2, E and G). Deletion of the N-ZnF and Myb domains together (ΔN mutant) completely inhibits target DNA nicking and subsequent TPRT (Fig. 2C) (*20*). The central GC of the RUM is recognized by His^{673} and Lys^{675} of the loop 6a of the RT domain (Fig. 2H). Structural predictions suggest that this loop is specific among non-LTR RT domains to R2 proteins (fig. S7). We found that deletion of the 6a loop inhibits

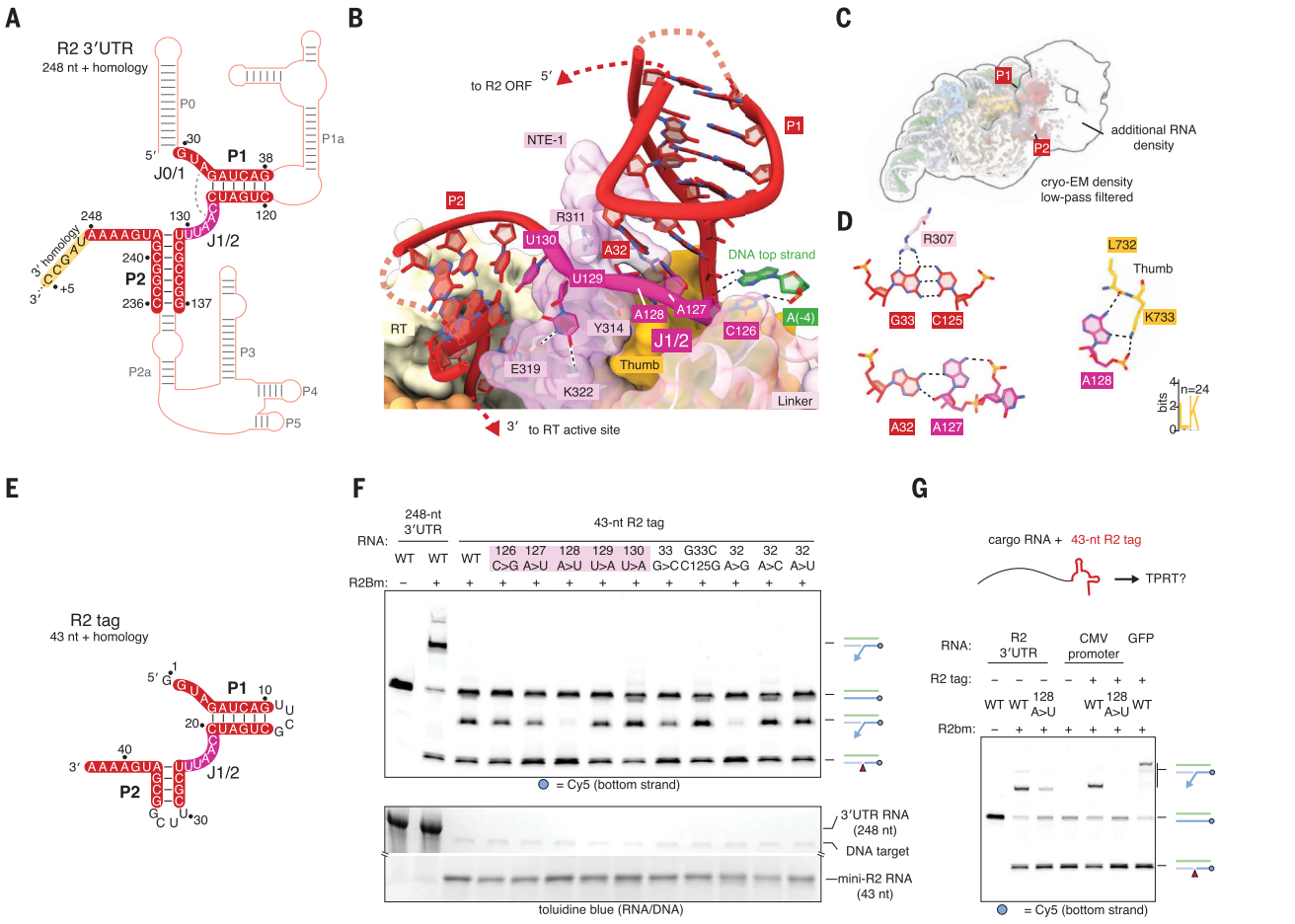


Fig. 4. Interactions of R2Bm with the 3′UTR RNA. (**A**) Secondary-structure diagram of the 3′UTR RNA, based on (*26*). Thicker strokes represent nucleotides visible in the cryo-EM density. Nucleotides are numbered from the first base of the 3′UTR (the base following the stop codon). (**B**) Structure of the 3′UTR RNA core and the R2Bm NTE-1 domain. Dotted lines, hydrogen bonds. (**C**) Low-pass filtered cryo-EM map. (**D**) Interactions between 3′UTR bases. Dotted lines, hydrogen bonds. (**E**) Secondary structure of the R2 tag RNA. Unshaded bases are not in the full-length 3′UTR. (**F**) Denaturing gel of in vitro TPRT reactions on a labeled 211-bp 28S DNA target using various R2 RNAs. Highlighted mutants are in the J1/2 region. The same gel was visualized by Cy5 fluorescence and toluidine blue staining. (**G**) The R2-tag allows TPRT of cargo RNAs. Denaturing gel shows TPRT reactions with equimolar amounts of the indicated RNAs and a labeled 211-bp 28S DNA target. R2 tag (43 nt) was added to the 3′ end of a 239-nt RNA encoding the CMV promoter or a 764-nt RNA encoding GFP.

target DNA nicking (Fig. 2C). Finally, we found that the distance between the RUM and the bottom-strand cleavage site (the RASIN) is important. Increasing the distance by one base was tolerated, but further increase or any decrease to the distance inhibited target cleavage (Fig. 2I).

Target DNA interactions at the cleavage and integration site

The second key region for DNA target recognition by R2Bm is the target site for nicking by the RLE domain and R2 insertion, which we term the RASIN. In our structure, the 11 bp of the RASIN from −6 to +5 are melted around the RLE domain. The ZnF appears to act as the "zip," stacking on the last upstream pair C–G(−7) with Arg^{922} and Arg^{924} and holding unzipped strands apart (Fig. 3A). Strand melting may be enhanced by the 40° bend in target DNA around the RUM (Fig. 1F). Bases −6 to −1 on the bottom strand then follow a cleft between the ZnF and the RLE, which adopts a canonical PD-(D/E)xK-family nuclease fold, but with the characteristic Lys^{1026} on an α helix instead of the usual β strand (Fig. 3B) (*24*). This lysine, along with catalytic residues Asp^{996} and Asp^{1009}, is 4 to 6 Å from the scissile phosphate of C(−1), suggesting that C(−1) may be close to its position during catalysis of bottom-strand cleavage. On the top strand, bases −6 to +2 all make extensive contacts along a cleft between the RLE and linker domains, except for A(−4), which flips out and contacts C126 of the 3′UTR (Fig. 3C). To determine the relative importance of the bases in the RASIN, we mutated each of the 11 bp individually and tested the effect on bottom-strand cleavage. Mutating T(+1) to A abolished cleavage entirely, and mutating T(−6), T(−5), and A(−3) severely decreased activity, whereas other changes were tolerated (Fig. 3D). This suggests the following RASIN motif for cleavage, given in top-strand sense: $T_{-6}TNANNT_{+1}$.

Because only the bottom strand of the RASIN enters the RLE active site, we tested the activity of R2Bm on a single-stranded DNA with the bottom-strand sequence and found that it was cut, albeit weakly (Fig. 3E). Endonuclease activity was strongly stimulated by providing a 60-nt top strand spanning the RASIN and upstream and downstream sequences, but was similarly stimulated by a 32-nt top strand complementary only to the upstream region containing the RUM. A 17-nt top strand complementary to the downstream sequence did not stimulate activity (Fig. 3E). This suggests that the RUM in a double-stranded state is important for recruiting the R2Bm RLE to the RASIN bottom strand and that the top strand of the RASIN, despite its extensive interaction with R2Bm, is dispensable for specific bottom-strand cleavage. However, when we added deoxynucleotides

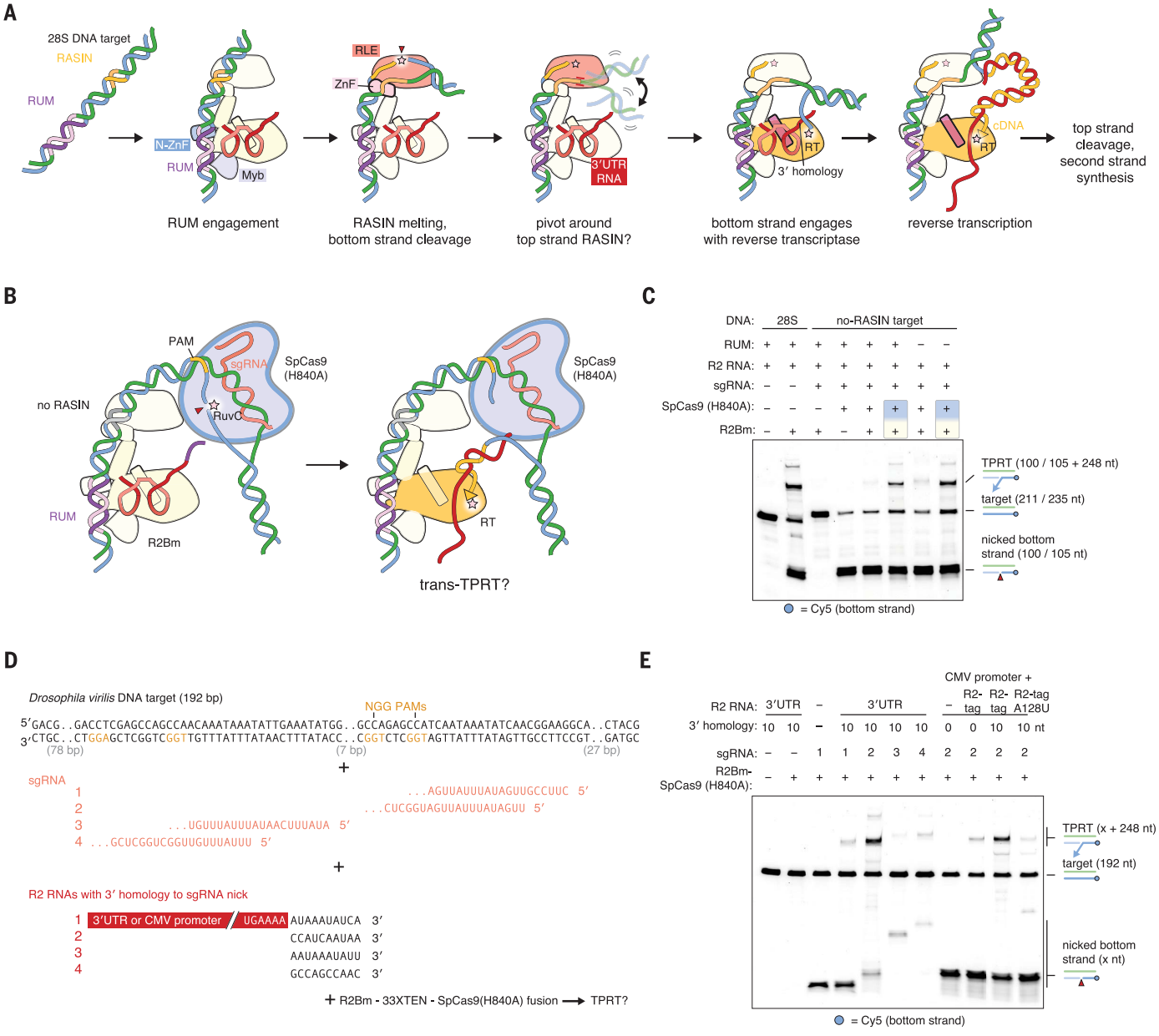


Fig. 5. Mechanism and retargeting of first-strand synthesis by R2Bm. (**A**) Model for the initial stages of target site cleavage and first-strand synthesis. (**B**) Design of R2Bm + Cas9 experiments. (**C**) Complementation of DNA target site mutants by Cas9 cleavage in trans and cis. The denaturing gel shows in vitro TPRT reactions on a labeled 211-bp target corresponding to the wild-type 28*S* target, or two 235-bp targets: one where the RASIN TAAGGTA is replaced by 31 bp of unrelated sequence, and another where the 13-bp RUM is additionally scrambled. R2Bm and SpCas9(H840A) were added in trans, or in cis connected by a 33XTEN linker (fusion indicated by a shaded box). The sgRNA is complementary to the inserted sequence and nicks 40 nt from the last RUM base. The R2 RNA is the 3′UTR with 5 nt of 3′ homology to the nick site. (**D**) Sequences used for retargeting R2Bm to an unrelated locus from the *Drosophila virilis* genome. (**E**) Denaturing gel of in vitro TPRT reactions on the labeled 192-bp *Drosophila virilis* target. sgRNAs are numbered as in (D); all R2 RNAs or R2-tagged RNAs have 10 nt of 3′ homology to the nick site of the sgRNA.

to these reactions, TPRT activity was eliminated in the absence of the top strand from the RASIN downstream but was partially rescued if the 3′UTR RNA contained 3′ homology to the target site (Fig. 3E). The top-strand RASIN bases A(−4), A(−3), and G(−2) are grasped by Arg^{901} and Asp^{902} of the R2Bm linker (Fig. 3C). We mutated these two residues to alanine and tested TPRT activity on a fully double-stranded substrate, and found that TPRT activity was reduced and partially rescued by 3′ homology (Fig. 3E). These results suggest two important factors for initiating TPRT when the 3′UTR RNA lacks 3′ homology: (i) the presence of a top strand downstream of the RASIN, which may help retain the nicked bottom strand, and (ii) contacts between R2Bm and the top-strand RASIN, which help the nicked bottom strand "pivot" into the RT active site.

R2Bm binds a small core region of the 3′UTR

R2Bm can only initiate TPRT on RNAs containing the R2 3′UTR (self-specificity), but the molecular basis for this is not known (*25*). Multiple

models have been proposed for the secondary structure of the R2 3′UTR, and the divergent sequences of R2 RNAs have hindered identification of key bases (*26*, *27*). A model for the R2 3′UTR secondary structure based on chemical probing is shown in Fig. 4A and has at least 11 stems (*26*). In our cryo-EM map, we resolved density for two stems and their flanking single-stranded regions (Fig. 4B). On the basis of nomenclature commonly used for structured RNAs, we name these stems P1 (nucleotides 33 to 38 and 120 to 135) and P2 (nucleotides 131 to 137 and 236 to 242), and term the single-stranded junction between P1 and P2 as J1/2 and the single-stranded region preceding P1 as J0/1. The rest of the 3′UTR may occupy a diffuse cloud of cryo-EM density next to these core regions (Fig. 4C).

P1 and J1/2 are mainly recognized by an α helix from the R2Bm NTE-1 domain, which packs into the major groove of P1 and is wrapped by J1/2 (Fig. 4B). Arg^{307} recognizes the Hoogsteen edge of P1 G33, and the interaction is secured by Arg^{310} and Arg^{311}. Consistently, these residues were previously shown to be essential for RNA binding (*15*), and the first 45 bases of the 3′UTR are essential for TPRT activity (*11*). J1/2 makes numerous sequence-specific contacts (Fig. 4D): A127 forms a sugar-edge pair with the Watson-Crick face of J0/1 A32, A128 hydrogen bonds to Leu^{732} and Lys^{733} of the R2Bm thumb domain and stacks on NTE-1 Tyr^{314}, U129 hydrogen bonds to Glu^{319} and Lys^{322} of NTE-1, and C126 stacks on and hydrogen bonds with A(-4) from the top strand of the DNA target (Fig. 4, B and D).

To test if regions of the R2 3′UTR not clearly visible in the cryo-EM density are required for TPRT activity, we designed a 43-nt minimal 3′UTR—"R2 tag"—that contains only the sequences visible in the cryo-EM density, linked by tetraloops (Fig. 4E). The R2 tag was reverse transcribed as efficiently as the full 248-nt 3′UTR in a TPRT reaction. We tested the importance of the J1/2 linker by making single-base transversions and found that A127U reduced activity and A128U almost completely abolished TPRT activity (Fig. 4F). Mutating G33 to C to disrupt base pairing at the bottom of stem P1 also reduced activity but could be rescued by the compensatory C125G mutation (Fig. 4F). Mutation of J0/1 A32 to G reduced activity, but mutations to C or U were tolerated. Equivalents to P1, P2, J0/1, and J1/2 can be identified in the secondary structures of diverse R2 elements (*26*) (fig. S7). The P1 and P2 stems have different sizes and base compositions, but positions 2 and 3 of J1/2, corresponding to A127 and A128, are conserved as adenosines, consistent with their importance for TPRT.

Because the R2 tag alone is efficiently integrated in a TPRT reaction, we tested if adding the R2 tag to the 3′ end of a "cargo" RNA would allow its integration at the 28*S* target site. We added the R2 tag to the 3′ end of a 239-nt cytomegalovirus (CMV) promoter RNA. This tagged RNA was used as efficiently as wild-type R2 3′UTR in a TPRT reaction, whereas an untagged RNA was not used, nor was an RNA tagged with an R2-tag A128U mutant (Fig. 4G). A larger RNA containing the 720-nt coding sequence for green fluorescent protein (GFP) and a 3′ R2 tag was also reverse transcribed in a TPRT reaction (Fig. 4G).

R2Bm can be retargeted with CRISPR-Cas9

Our structural and biochemical observations suggest a multistep model for initiation of TPRT: The R2Bm N-terminal domains first detect a RUM sequence, followed by cleavage of the bottom strand at the RASIN site, possible pivoting of the nick around the top strand into the RT active site, annealing of any 3′ homology to the nicked bottom strand, and finally initiation of reverse transcription (Fig. 5A). This model implies that R2Bm could prime reverse transcription off an exogenously nicked bottom strand close to the R2Bm binding site (Fig. 5B). To test this, we replaced the RASIN and downstream sequences of the 28*S* DNA target with an unrelated sequence containing an efficient SpCas9 target sequence, but kept the RUM sequence to anchor R2Bm (Fig. 5B). This substrate could not be cleaved by R2Bm but was nicked efficiently by an SpCas9 H840A nickase mutant (Fig. 5C). When SpCas9 and R2Bm were added together with a single-guide RNA (sgRNA) and an R2 3′UTR RNA with 5 nt of 3′ homology to the sgRNA nick site, we detected low amounts of TPRT activity. This activity was enhanced when the R2Bm and SpCas9 proteins were fused with a 33XTEN flexible linker (Fig. 5C). The RUM was not required for Cas9-directed TPRT, as mutating the RUM did not reduce activity (Fig. 5C). This suggests that Cas9 might be able to direct R2Bm to perform TPRT at loci other than the 28*S* target. We mixed the R2Bm-Cas9(H840A) fusion protein with a 192-bp target sequence from *Drosophila virilis*, various sgRNAs, and R2 3′UTRs with 10 nt of 3′ homology to the nick site dictated by the sgRNA (Fig. 5D). We found TPRT activity at all Cas9 nick sites, with one sgRNA (guide 2) giving efficient activity (Fig. 5E). Adding R2Bm and SpCas9(H840A) as separate polypeptides also yielded efficient TPRT with guide 2 but was less robust with other guides (fig. S9). The 239-nt CMV promoter RNA with a 3′ R2 tag and 10 nt of homology to the guide 2 nick site was also reverse transcribed efficiently; this activity required the R2 tag and was reduced in the absence of 3′ homology or with the R2 tag A128U mutation (Fig. 5E). Larger RNAs such as GFP could also be reverse transcribed at the guide 2 nick site (fig. S9). In summary, R2Bm can be retargeted by Cas9 to perform TPRT at unrelated loci, and the R2 tag can direct incorporation of cargo RNAs at these sites.

Discussion

Here we show the structure of a non-LTR retrotransposon during transposition, and we dissect the principles of target DNA and self-RNA recognition. Our structure suggests that R2Bm uses its N-ZnF and Myb domains to locate the endonuclease target sequence, a model that contrasts with the model for other non-LTR retrotransposons in which the endonuclease domain is the only determinant of target site selection (*28*, *29*). We identified two essential target site motifs—the RUM and RASIN—that are recognized by R2Bm, but we note that searching the *B. mori* genome with a RUM-RASIN consensus motif yields many potential off-target sites outside of the ribosomal DNA arrays (fig. S10). We examined the sequence of a previously identified *B. mori* non-28*S* insertion in (*30*) and found that the target site had limited similarity with 28*S* but had a TTAAcG|T RASIN motif ("|" indicates insertion site, lowercase denotes deviation from 28*S*) and a GCTACTTGCGCAT RUM the correct distance upstream of the RASIN (fig. S10). Non-28*S* insertions, however, are rare, so it is likely that other factors are important in regulating R2Bm transposition, including chromatin accessibility, other sequence motifs, or the ability of the target DNA to bend and melt.

Non-LTR retrotransposons form a diverse family, and even within the R2 superclade there are notable differences between elements. R2Bm is a representative of the R2-D clade of elements, which have a single C2H2 N-terminal ZnF domain, but R2-A clade elements have three tandem N-terminal ZnF domains (*31*) that may create a more extensive DNA binding interface with greater stringency in target site selection. More broadly, non-LTR retrotransposons can be divided into two types on the basis of their endonuclease domains: those that, like R2Bm, use a C-terminal restriction enzyme-like (RLE) domain, and those that, like human LINE-1, use an unrelated N-terminal apurinic or apyrimidinic endonuclease (APE) domain (*32*, *33*). Structure prediction using AlphaFold (*34*) suggests that, in these retrotransposons, the position of the APE domain is distinct from that of the RLE domain in R2Bm, suggesting that there may be mechanistic differences in how target cleavage is coupled to reverse transcription (fig. S5) (*35*). Nonetheless, the similarity between the DNA interface on the R2Bm thumb domain and the corresponding interface in the group IIC intron (fig. S5) suggests that this interface might be conserved among most non-LTR retrotransposons (*19*). Indeed, the upstream DNA from R2Bm was easily modeled into an AlphaFold model of human LINE-1 ORF2, including not only the thumb interactions but also strand

separation by the CCHC ZnF domain, which in LINE-1 ORF2 corresponds to the C-terminal cysteine-rich domain (fig. S5).

Overall, the results of this work advance our understanding of transposition by non-LTR retrotransposons and suggest avenues for engineering these transposons for targeted gene insertions.

REFERENCES AND NOTES

1. H. H. Kazazian Jr., J. V. Moran, *N. Engl. J. Med.* **377**, 361–370 (2017).
2. S. J. Priest *et al.*, *Nat. Microbiol.* **7**, 1239–1251 (2022).
3. S. R. Richardson *et al.*, *Microbiol. Spectr.* **3**, MDNA3–0061–2014 (2015).
4. H. Fujiwara, *Microbiol. Spectr.* **3**, MDNA3–0001–2014 (2015).
5. T. H. Eickbush, D. G. Eickbush, *Microbiol. Spectr.* **3**, MDNA3–0011–2014 (2015).
6. D. D. Luan, M. H. Korman, J. L. Jakubczak, T. H. Eickbush, *Cell* **72**, 595–605 (1993).
7. J. Yang, H. S. Malik, T. H. Eickbush, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 7847–7852 (1999).
8. A. M. Lambowitz, S. Zimmerly, *Cold Spring Harb. Perspect. Biol.* **3**, a003616 (2011).
9. S. Zimmerly, H. Guo, P. S. Perlman, A. M. Lambowitz, *Cell* **82**, 545–554 (1995).
10. D. D. Luan, T. H. Eickbush, *Mol. Cell. Biol.* **16**, 4726–4734 (1996).
11. D. D. Luan, T. H. Eickbush, *Mol. Cell. Biol.* **15**, 3882–3891 (1995).
12. A. Bibiłło, T. H. Eickbush, *J. Mol. Biol.* **316**, 459–473 (2002).
13. C. Zhao, A. M. Pyle, *Nat. Struct. Mol. Biol.* **23**, 558–565 (2016).
14. A. M. Lentzsch, J. L. Stamos, J. Yao, R. Russell, A. M. Lambowitz, *J. Biol. Chem.* **297**, 100971 (2021).
15. V. K. Jamburuthugoda, T. H. Eickbush, *Nucleic Acids Res.* **42**, 8405–8415 (2014).
16. W. P. Galej, C. Oubridge, A. J. Newman, K. Nagai, *Nature* **493**, 638–643 (2013).
17. M. M. Mahbub, S. M. Chowdhury, S. M. Christensen, *Mob. DNA* **8**, 16 (2017).
18. S. C. Pimentel, H. E. Upton, K. Collins, *J. Biol. Chem.* **298**, 101624 (2022).
19. K. Chung *et al.*, *Science* **378**, 627–634 (2022).
20. S. M. Christensen, A. Bibillo, T. H. Eickbush, *Nucleic Acids Res.* **33**, 6461–6468 (2005).
21. S. Christensen, T. H. Eickbush, *J. Mol. Biol.* **336**, 1035–1045 (2004).
22. A. Klug, *Annu. Rev. Biochem.* **79**, 213–231 (2010).
23. P. König, R. Giraldo, L. Chapman, D. Rhodes, *Cell* **85**, 125–136 (1996).
24. A. Govindaraju, J. D. Cortez, B. Reveal, S. M. Christensen, *Nucleic Acids Res.* **44**, 3276–3287 (2016).
25. M. Osanai, H. Takahashi, K. K. Kojima, M. Hamada, H. Fujiwara, *Mol. Cell. Biol.* **24**, 7902–7913 (2004).
26. A. M. Ruschak *et al.*, *RNA* **10**, 978–987 (2004).
27. D. H. Mathews, A. R. Banerjee, D. D. Luan, T. H. Eickbush, D. H. Turner, *RNA* **3**, 1–16 (1997).
28. H. Takahashi, H. Fujiwara, *EMBO J.* **21**, 408–417 (2002).
29. Q. Feng, J. V. Moran, H. H. Kazazian Jr., J. D. Boeke, *Cell* **87**, 905–916 (1996).
30. Y. Xiong, W. D. Burke, J. L. Jakubczak, T. H. Eickbush, *Nucleic Acids Res.* **16**, 10561–10573 (1988).
31. A. Luchetti, B. Mantovani, *PLOS ONE* **8**, e57076 (2013).
32. H. S. Malik, W. D. Burke, T. H. Eickbush, *Mol. Biol. Evol.* **16**, 793–805 (1999).
33. I. R. Arkhipova, *Mob. DNA* **8**, 19 (2017).
34. J. Jumper *et al.*, *Nature* **596**, 583–589 (2021).
35. I. Miller *et al.*, *Nucleic Acids Res.* **49**, 11350–11366 (2021).

ACKNOWLEDGMENTS

We thank S. Zhu and M. Walsh for valuable discussions; E. Brignole and C. Borsa for the smooth running of the MIT.nano cryo-EM facility, established in part with financial support from the Arnold and Mabel Beckman Foundation; S. Lövestam for a critical reading of the manuscript; and the entire Zhang lab for support and advice. We thank T. H. Eickbush and colleagues for their inspiring and pioneering work on R2 elements. **Funding:** Helen Hay Whitney Foundation Postdoctoral Fellowship (M.E.W.), Howard Hughes Medical Institute (M.E.W., F.Z.), National Institutes of Health grant 2R01HG009761-05 (F.Z.), Hock E. Tan and K. Lisa Yang Center for Autism Research at MIT (F.Z.), K. Lisa Yang and Hock E. Tan Molecular Therapeutics Center at McGovern (F.Z.), K. Lisa Yang Brain-Body Center at MIT (F.Z.), Broad Institute Programmable Therapeutics Gift Donors (F.Z.), the Pershing Square Foundation, W. Ackman, and N. Oxman (F.Z.), Asness Family Foundation (F.Z.), BT Charitable Foundation (F.Z.), the Phillips family (F.Z.), J. and P. Poitras (F.Z.), D. Cheng (F.Z.), R. Metcalfe (F.Z.). **Author contributions:** M.E.W. and F.Z. conceived the project. M.E.W. designed and performed experiments and solved the cryo-EM structure. C.J.F. generated and analyzed sequencing data. F.Z. supervised the research and experimental design with support from R.K.M. M.E.W. wrote the manuscript with input from all authors. **Competing interests:** F.Z. is a scientific adviser for and cofounder of Editas Medicine, Beam Therapeutics, Pairwise Plants, Arbor Biotechnologies, and Aera Therapeutics. F.Z. is a scientific adviser for Octant. **Data and materials availability:** The cryo-EM map has been deposited in the Electron Microscopy Data Bank with accession code EMD-40033. The coordinates for the atomic model have been deposited in the Protein Data Bank with accession code 8GH6. The raw cryo-EM data have been deposited in EMPIAR with accession code EMPIAR-11458. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse. This article is subject to HHMI's Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the author-accepted manuscript of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adg7883
Materials and Methods
Figs. S1 to S10
Tables S1 to S3
References (*36*–*47*)
MDAR Reproducibility Checklist

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

Submitted 20 January 2023; accepted 21 March 2023
10.1126/science.adg7883

SEXUAL SELECTION

Female preference for rare males is maintained by indirect selection in Trinidadian guppies

Tomos Potter[1]*†, Jeff Arendt[2]†, Ronald D. Bassar[3], Beth Watson[4], Paul Bentzen[4], Joseph Travis[1], David N. Reznick[2]

When females prefer mates with rare phenotypes, sexual selection can maintain rather than deplete genetic variation. However, there is no consensus on why this widespread and frequently observed preference might evolve and persist. We examine the fitness consequences of female preference for rare male color patterns in a natural population of Trinidadian guppies, using a pedigree that spans 10 generations. We demonstrate (i) a rare male reproductive advantage, (ii) that females that mate with rare males gain an indirect fitness advantage through the mating success of their sons, and (iii) the fitness benefit that females accrue through their "sexy sons" evaporates for their grandsons as their phenotype becomes common. Counter to prevailing theory, we show that female preference can be maintained through indirect selection.

Whether female preference for rare males can sustain genetic polymorphisms in nature has long been controversial (*1*). Rarity, as an attractive trait, complicates sexual selection theory for the evolution of female preference because it introduces negative frequency-dependence. Negative frequency-dependent selection occurs when the fitness of a trait decreases as it becomes more common and increases as it becomes rarer. In the absence of negative frequency-dependence, female preference for certain male traits can be explained if the sons of attractive males are also attractive (*2*, *3*) or if paternal attractiveness correlates with enhanced viability in offspring (*4*, *5*). However, when attractiveness is frequency-dependent, as is the case when rare male phenotypes have an advantage, sons can become victims of their father's success. Specifically, the progeny of successful rare males are doomed to become common and thus unattractive. Although there is robust theory describing conditions under which female preference for rare males may evolve (*6*), it is unclear whether the costs of such a preference will outweigh its benefits (*7*).

Numerous laboratory studies have demonstrated a rare-male mating advantage in several taxa [reviewed in (*8*, *9*)]. However, we know of only five such studies in nature (*10*–*14*). As is common in laboratory studies, all studies in nature except one (*12*) reduced the options for female choice to just two male morphs (albeit in some cases noting variation within morphs) (*11*, *14*). These studies were further limited to a single mating season, precluding the detection of long-term fitness consequences for females that mate with rare males. As a result, although these studies have documented rare-male advantage and demonstrated its proximate mechanism through female preference behaviors, the ultimate, evolutionary explanation of why females prefer to mate with rare males remains unknown.

We show an advantage for rare color patterns in males under natural conditions in the highly polymorphic Trinidadian guppy (*Poecilia reticulata*) and quantify the fitness consequences of this advantage over multiple generations. Male color patterns in guppies are reliably transmitted from father to son (*15*–*18*). We identified 27 distinct color patterns in our study population and have confirmed that all males within a patriline share the same pattern (Fig. 1) (*19*). During courtship, males display their pattern to potential mates (*20*). Numerous laboratory studies (*15*, *20*–*30*) and one field manipulation (*31*) have demonstrated female preference for rare or unfamiliar male patterns. Whether those results apply to the much wider level of unmanipulated variation in wild male color patterns is unknown.

We tested these ideas as part of an experimental study of evolution in a natural stream in Trinidad (*19*, *32*, *33*). We used monthly mark-recapture data to determine the presence and movement patterns of individuals within the population over this period. We used a microsatellite-based pedigree to determine the relatedness of individuals and the reproductive success for each individual every month over this period (*19*). Our dataset includes monthly observations of 7173 individuals spanning 10 generations (*34*). We used generalized linear mixed effects models (GLMMs) to test the effects of male pattern rarity and novelty (defined below) on components of fitness in guppies.

Measuring rarity and novelty

We assigned a "rarity" score monthly to each male. We calculated the rarity of a focal pattern (r_i) as a function of the total number of individuals with that pattern (n_i), the total number of individuals of all patterns (N_p), and the degree of polymorphism, i.e., the number of patterns (P):

$$r_i = ln\left(\frac{n_i}{N_p}P\right)$$

Weighting the relative frequency of patterns ($^{n_i}/_{N_P}$) by P allows meaningful comparison across localities with different degrees of polymorphism. Another useful aspect of this approach is that log-transformation results in rare patterns having negative values, common patterns having positive values, and patterns that are neither rare nor common ($^{n_i}/_{N_P} = {}^{1}/_{P}$) having a value of zero. This makes linearized model coefficients directly interpretable. To illustrate our results, we define a "rare" male as one with a pattern half as frequent as expected given the total degree of polymorphism [i.e., $r_i = \ln(0.5)$], and a "common" male as one with a pattern twice as frequent as expected [i.e., $r_i = \ln(2)$]. These illustrative values fall well within the observed distribution of male pattern rarity (fig. S1).

A female's assessment of male rarity will depend upon the males that she regularly encounters, which may be a spatial subset of the total population. The stream habitat is subdivided into discrete pools connected by riffles. Our spatially explicit mark-recapture censuses allowed us to reconstruct patterns of movement and make inferences about population structure. To assess the possibility that female mate preference is shaped by that structure, we calculated rarity at three spatial scales: the local level (e.g., the pool or riffle where the fish was caught that month), the neighborhood, and the whole population. Neighborhoods were defined using network analysis of movement of male guppies between pools (*19*). We identified four distinct multipool neighborhoods characterized by high movement within but low movement between. Males moved around considerably (62% were new arrivals to pools each month, 17% were new arrivals to neighborhoods, fig. S2), whereas females moved around much less (28% in pools, 4% in neighborhoods, fig. S2). When a female assesses how rare a male is, she is likely to see all those males that we collected in the pool with her that month. Although we did not observe this directly, our neighborhood-level analyses include males likely to have passed through the pool in the previous month.

In addition to an advantage to rarity, several studies have demonstrated the advantage of novelty in the form of female preference for unfamiliar males (*15*, *22*, *23*, *26*, *35*), regardless of their color pattern. Female guppies may identify novel males through olfactory cues (*36*). As such, males that are new arrivals in a

[1]Department of Biological Science, Florida State University, Tallahassee, Florida, USA. [2]Department of Evolution, Ecology and Organismal Biology, University of California, Riverside, California, USA. [3]Department of Biological Sciences, Auburn University, Auburn, Alabama, USA. [4]Department of Biology, Dalhousie University, Halifax, Nova Scotia, Canada.
*Corresponding author. Email: tomos.potter@protonmail.com
†These authors contributed equally to this work.

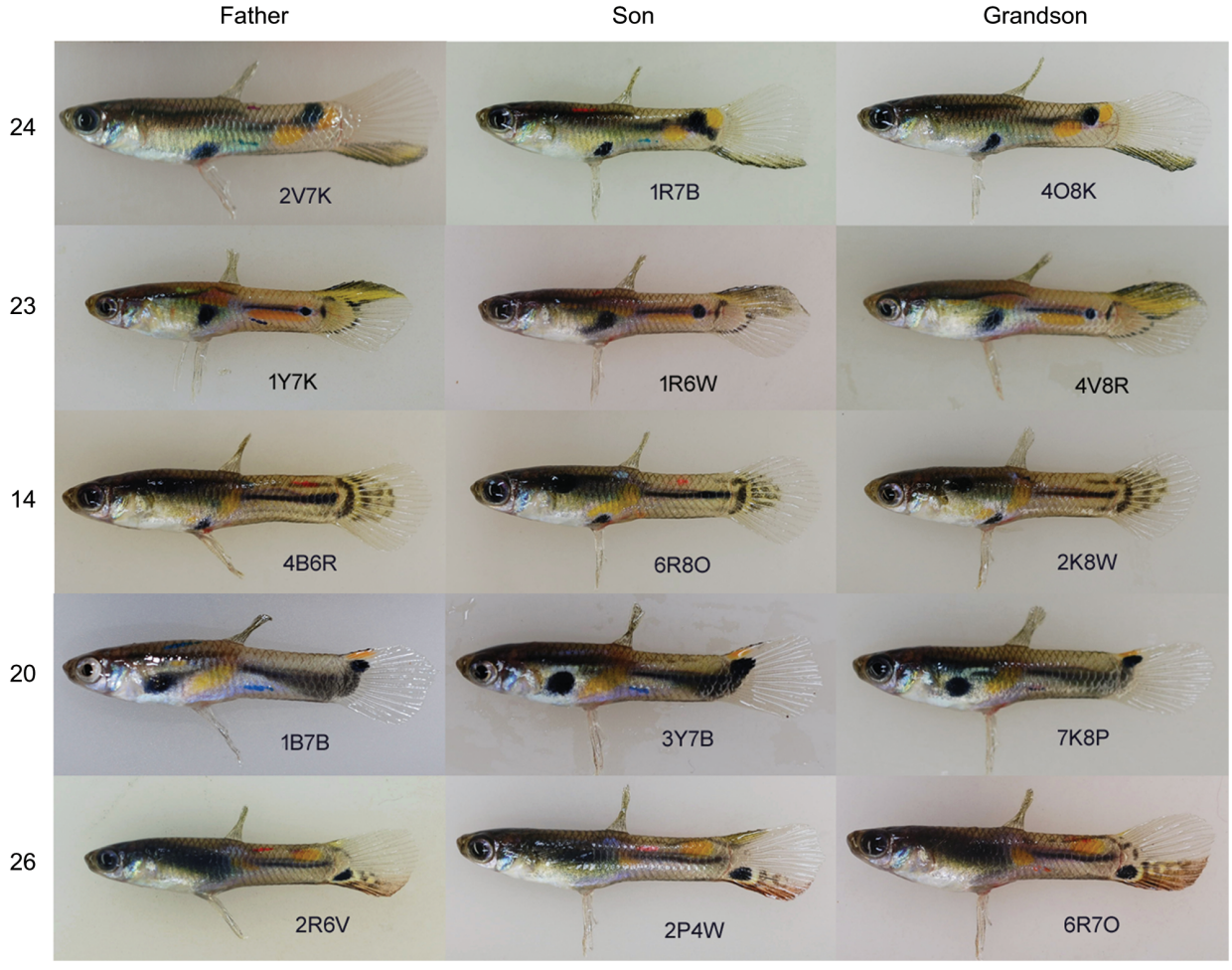


Fig. 1. Male color patterns in guppies are reliably transmitted from father to son. Here, we show three generations (father, son, and grandson) for five example patrilines showing consistency in color pattern within a Y-lineage. Numbers to the left indicate which lineage of 27 are being shown. Some elastomer marks, used to identify individuals, are clearly visible. For example, the son of lineage 14 has a red mark on the dorsal side of the caudal peduncle.

pool or neighborhood may experience a reproductive advantage. To test this, we defined males as "novel" if they were new arrivals to the pool or neighborhood in which they were caught that month. By this definition, males cease to be novel one month after arriving in a locality.

Results

We found evidence for negative frequency-dependent selection operating on male color patterns, resulting in a rare-male advantage (Fig. 2 and table S1). These effects were significant over all three spatial scales over which male rarity was calculated but were strongest and weakest at the neighborhood and local levels, respectively, as determined by Akaike information criterion (AIC) scores (table S1). Each month, males with rarer patterns at the neighborhood level had 36% more mating partners (GLMM, n = 6248, P-value = 2.43×10^{-6}) and ultimately sired 38% more offspring that recruited into the population (GLMM, n = 6248, P-value = 4.75×10^{-6}) (Fig. 2). These results, observed over multiple generations, provide strong evidence that negative frequency-dependent sexual selection is occurring through rare-male advantage.

Novel males (new arrivals to a pool or neighborhood, regardless of their color pattern) also had a large reproductive advantage over residents (Fig. 2 and table S1), with the effect strongest at the local level. Compared with residents of equivalent rarity, new arrivals to pools had 45% more mating partners (GLMM, n = 6248, P-value = 2.77×10^{-4}) and sired 50% more offspring each month (GLMM, n = 6248, P-value = 3.30×10^{-4}).

In line with earlier studies (*15*, *20*–*30*), our results indicate that female guppies prefer to mate with rare color-patterned and/or unfamiliar males. One potential explanation for this preference is inbreeding avoidance: rare or novel males may be less likely to be kin (*37*). However, we found no evidence that male rarity was associated with the relatedness of mating partners (linear model, n = 1259, P-value = 0.273) (*19*). Surprisingly, resident males were more likely to be unrelated to their partners than novel males who were new arrivals to pools [$P(\text{unrelated}|\text{resident} = 0.28)$, $P(\text{unrelated}|\text{novel} = 0.15)$, logistic regression, n = 1580, P-value = 5.74×10^{-8})]. This could occur if males are more likely to remain in pools with unrelated females. Nevertheless, the higher relatedness of novel (and thus attractive) males and the absence of any association of relatedness with rarity contradict the inbreeding-avoidance hypothesis (table S2).

We found no direct benefit for females' preference for rarity or novelty. Mating with rare males did not result in more recruited offspring (table S3 and Fig. 2C, GLMM, n = 2290, P-value = 0.448), nor did mating with new arrivals to neighborhoods (P-value = 0.969) or pools (P-value = 0.397). Our measure of recruited offspring refers to those that survived to be large enough to be individually marked (~2 months old) (*19*), meaning that

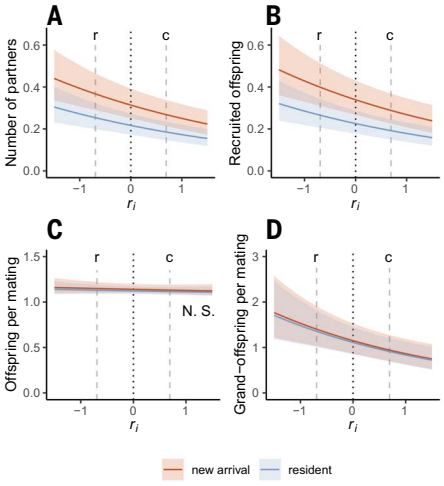


Fig. 2. Rare and novel males have higher reproductive fitness, and females that mate with rare males have more grand-offspring. Effects of male pattern rarity (r_i) and novelty (new arrival or resident) on components of fitness. (**A**) Number of mating partners per month (for males); (**B**) monthly number of offspring recruited into the population (for males); (**C**) number of recruited offspring per mating (for males and females); (**D**) number of grand-offspring that ultimately recruit into the population from a single mating (for males and females). Values for "rare" [$r_i = \ln(0.5)$] and "common" [$r_i = \ln(2)$] males are indicated with dotted lines annotated *r* and *c*, respectively; the dashed line indicates $r_i = 0$. Shaded areas are 95% confidence intervals, N.S. indicates that the slope is not significant ($P > 0.05$). Predictions are based on models where rarity was calculated at the neighborhood level (A) and (B) or the population level (C) and (D).

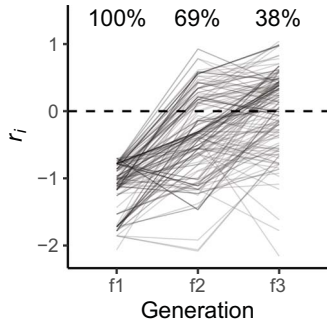


Fig. 3. Rare males have rare sons but common grandsons. In this figure, we track the trajectory of male pattern rarity (r_i, at the population level, *x* axis) over three generations (*y* axis), focusing on rare males [f1, $r_i < \ln(0.5)$], their sons (f2), and grandsons (f3). Lines connect male to sons to grandsons. Percentages describe individuals in each generation where $r_i < 0$, i.e., that are rarer than expected.

variation in offspring viability is captured in this metric. This indicates that preference for rare or novel males is not under direct selection through mechanisms that enhance offspring viability, such as inbreeding avoidance or so-called "good genes" (*4*, *5*).

What then is the ultimate benefit of mating with rare or novel males? Although we detected no indirect fitness benefits for females that mated with novel males [i.e., that were new arrivals to neighborhoods (GLMM, n = 1951, P-value = 0.439) or pools (P-value = 0.573)], matings with rare males (at the population level) ultimately resulted in 48% more grand-offspring recruited into the population than matings with common males (P-value = 3.42×10^{-4}). This is a substantial indirect fitness benefit for those females (table S3 and Fig. 2D).

Females that mated with rare males gained this indirect fitness advantage through the enhanced reproductive success of their sons: a so-called "sexy sons" effect [in the sense of Kokko (*6*)]. The sons of rare males (at the population level) sired more offspring per month (table S4, GLMM, n = 9807, P-value = 0.0039), but there was no such effect in daughters (P-value = 0.575). This occurred because the sons of rare males were also rare (albeit less so than their fathers) and thus still attractive (Fig. 3). This advantage was short-lived, however; after two generations of rare-male advantage the grandsons of rare males became victims of their forefathers' success and were more likely to be common (Fig. 3). Consequently, for males with equivalently rare fathers, having a rare grandfather reduced reproductive success (P = 0.0082).

Discussion

Female guppies that mated with rare males gained no direct fitness advantage in doing so. Their offspring did not have increased viability due to any genetic advantage of their attractive fathers, nor were they less inbred. Instead, females that mated with rare males derived substantial indirect fitness through the attractiveness of their sons. This is at odds with the long-held prediction that such indirect selection cannot maintain female preference (*7*, *38*–*40*). This prediction is not an ineluctable consequence of theory. It is based on assumptions about the genetic variances and covariances of female preference and male traits, which imply that indirect selection must always be overwhelmed by the direct costs of female choice (*7*). We suggest that this may not hold true when the desirable trait is rarity and physical traits are arbitrary. Our study shows that female preference can be maintained by indirect selection when negative frequency-dependence is operating.

Our findings offer a resolution to the "lek paradox" (*39*). To maintain female preferences through indirect selection, there must be a sustained supply of genetic variation in male traits (*39*, *41*). The crux of the lek paradox is that selection on male traits will erode that genetic variation, ultimately resulting in the loss of female preference (*39*). However, when females prefer rare males—regardless of male genotype—negative frequency-dependent selection occurs, ensuring the necessary maintenance of genetic variation.

A notable result is the absence of any detectable fitness benefit, direct or indirect, for females that mated with novel males. Novel males (new arrivals to pools or neighborhoods) had substantially higher reproductive success than residents, regardless of the rarity of their color pattern. Our results illustrate that mating with rare and novel males has distinct fitness consequences for females: Rare males conferred a single-generation reproductive advantage to their sons, driving indirect selection for female preference, whereas novel males conferred no fitness advantages to their partners, either directly or through their offspring. Why then do females prefer novel males?

One possibility is that female preferences for rare and novel males stem from a single, simple mechanism: habituation to familiar males, i.e., females preferring males that are unlike those they have recently encountered (*21*, *22*, *27*). Males with rare color patterns or that are new arrivals to pools (i.e., novel) are both likely to fit this criterion. In this scenario, selection for choosy females is driven by the indirect fitness advantage gained when they mate with rare males. By contrast, preference for novel males emerges as a nonadaptive by-product of the simple behavioral mechanism under selection.

Female choosiness is likely also under frequency-dependent selection (*6*, *42*). Consider what would happen if all females mated with a single male bearing the rarest color pattern: the sexy son benefit would be lost because all male offspring would have the same pattern, making it common and thus unattractive. As a result, selection for choosy females would evaporate. Although we do not know the mean frequency of choosy females in our population, theory suggests that it is likely to be high. Female preference alleles evolve to higher frequencies when the ability to express choice is hindered (*6*). Here, female choice was hindered by the different movement patterns of males and females. The optimum scale on which females should choose rare males is at the level of the population: Males frequently change location so the rarity of sons, upon which the indirect benefits to females depend, is best predicted by the rarity of fathers at the population level (tables S3 and S4). However, females can only assess the rarity of males they encounter. The more limited movement of females meant that they chose males that were

rare at the level of the neighborhood (table S1). This mismatch between the optimum and realized exercises of choice creates the hindrance that could sustain a high frequency of choosy females.

In conclusion, our results challenge the theoretical arguments against the role of sexy sons in sexual selection (*7*, *38*–*40*) by showing that this indirect form of selection can sustain female preference. At the same time, we show that female preference for rare male phenotypes resolves the lek paradox. Both results are a consequence of negative frequency-dependent selection operating on sexual signals and preferences in guppies. Female preference for rare males is well documented in a diversity of organisms (*8*–*14*), but detecting indirect selection in the wild is uncommon because it requires multigenerational studies. The replication of such studies in other organisms will test the generality of our results and determine the broader importance of sexual selection in maintaining, rather than depleting, genetic variation in the wild.

REFERENCES AND NOTES

1. R. C. Lewontin, *The Genetic Basis of Evolutionary Change* (Columbia Univ. Press, 1974), vol. 560.
2. R. A. Fisher, *The Genetical Theory of Natural Selection* (The Clarendon Press, 1930).
3. R. Lande, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 3721–3725 (1981).
4. M. Andersson, *Evolution* **40**, 804–816 (1986).
5. A. Grafen, *J. Theor. Biol.* **144**, 473–516 (1990).
6. H. Kokko, M. D. Jennions, A. Houde, *Proc. Biol. Sci.* **274**, 1317–1324 (2007).
7. E. Cameron, T. Day, L. Rowes, *J. Evol. Biol.* **16**, 1055–1060 (2003).
8. P. Knoppien, *Biol. Rev.* **60**, 81–117 (1985).
9. H. Ajuria Ibarra, T. Reader, *J. Zool.* **290**, 81–95 (2013).
10. D. A. S. Smith, *Science* **187**, 664–665 (1975).
11. J. Muggleton, *Heredity* **42**, 57–65 (1979).
12. V. M. Salceda, W. W. Anderson, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 9870–9874 (1988).
13. B. S. Masters, B. G. Hicks, L. S. Johnson, L. A. Erb, *Proc. Biol. Sci.* **270**, 1393–1397 (2003).
14. D. Estévez *et al*., *Front. Mar. Sci.* **7**, 614237 (2020).
15. J. A. Farr, *Evolution* **31**, 162–168 (1977).
16. A. E. Houde, *Anim. Behav.* **36**, 888–896 (1988).
17. A. E. Houde, *Heredity* **69**, 229–235 (1992).
18. J. R. Paris *et al*., *Nat. Commun.* **13**, 1233 (2022).
19. Materials and methods are available as supplementary materials.
20. J. A. Farr, *Behaviour* **74**, 38–90 (1980).
21. K. A. Hughes, L. Du, F. H. Rodd, D. N. Reznick, *Anim. Behav.* **58**, 907–916 (1999).
22. A. L. Eakley, A. E. Houde, *Proc. R. Soc. B.* **271** (2004).
23. S. R. K. Zajitschek, J. P. Evans, R. Brooks, *Behav. Ecol.* **17**, 911–916 (2006).
24. S. R. K. Zajitschek, R. C. Brooks, *Am. Nat.* **172**, 843–854 (2008).
25. K. J. Hampton, K. A. Hughes, A. E. Houde, *Ethology* **115**, 475–481 (2009).
26. M. M. Mariette, S. R. K. Zajitschek, C. M. Garcia, R. C. Brooks, *J. Evol. Biol.* **23**, 1772–1782 (2010).
27. M. J. Daniel, L. Koffinas, K. A. Hughes, *Proc. R. Soc. B.* **286**, 20190435 (2019).
28. T. Lucon-Xiccato, A. Bisazza, A. Pilastro, *Anim. Behav.* **155**, 217–224 (2019).
29. J. J. Valvo, F. H. Rodd, K. A. Hughes, *Behav. Ecol.* **30**, 1672–1681 (2019).
30. M. J. Daniel, L. Koffinas, K. A. Hughes, *Am. Nat.* **196**, 414–428 (2020).
31. K. A. Hughes, A. E. Houde, A. C. Price, F. H. Rodd, *Nature* **503**, 108–110 (2013).
32. J. Travis *et al*., in *Advances in Ecological Research, Vol 50: Eco-Evolutionary Dynamics*, J. MoyaLarano, J. Rowntree, G. Woodward, Eds. (Elsevier, 2014), vol. 50, pp. 1–40.
33. D. N. Reznick *et al*., *Am. Nat.* **194**, 671–692 (2019).
34. T. Potter *et al*., Data and code for the manuscript: "Female preference for rare males is maintained by indirect selection in Trinidadian guppies"(2023), doi: 10.5281/ZENODO.7644306
35. R. E. Graber, M. Senagolage, E. Ross, A. E. Houde, K. A. Hughes, *Ethology* **121**, 17–25 (2015).
36. A. J. Shohet, P. J. Watt, *Behav. Ecol. Sociobiol.* **55**, 363–369 (2004).
37. A. M. Johnson *et al*., *Ethology* **116**, 448–457 (2010).
38. M. Kirkpatrick, N. H. Barton, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1282–1286 (1997).
39. M. Kirkpatrick, M. J. Ryan, *Nature* **350**, 33–38 (1991).
40. J. S. Kotiaho, M. Puurtinen, *Funct. Ecol.* **21**, 638–644 (2007).
41. H. Kokko, I. Booksmythe, M. D. Jennions, *J. Evol. Biol.* **28**, 259–266 (2015).
42. H. Kokko, M. D. Jennions, R. Brooks, *Annu. Rev. Ecol. Evol. Syst.* **37**, 43–66 (2006).

ACKNOWLEDGMENTS

We thank the many field managers and interns, past and present, for their contributions to the Guppy Project and J. Ramlal and M. Ramlal and their family for hosting us in Trinidad. **Funding:** We gratefully acknowledge the continuing support of the National Science Foundation [DEB-0623632EF, DEB-0808039, and DEB-1258231 (to D.N.R.), DEB-1556884 (to J.T.), and DEB-2100163 (to R.D.B.)]. **Ethics:** All animals were collected and handled with the approval of the University of California Riverside IACUC AUP (a-20080008). **Author contributions:** Conceptualization: J.A., D.N.R., T.P., J.T.; Methodology: all authors; Investigation: T.P., J.A., J.T., D.N.R.; Visualization: T.P.; Funding acquisition: D.N.R., J.T., R.D.B.; Project administration: D.N.R., J.T., R.D.B.; Supervision: D.N.R., J.T.; Writing – original draft: T.P.; Writing – review and editing: all authors. **Competing interests:** Authors declare that they have no competing interests. **Data and materials availability:** All data and code required to reproduce the analyses presented here are hosted at Zenodo (*34*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade5671
Materials and Methods
Figs. S1 to S4
Tables S1 to S5
References (*43*–*55*)
MDAR Reproducibility Checklist

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

Submitted 25 August 2022; accepted 22 March 2023
10.1126/science.ade5671

WORKING LIFE

By **E. Celeste Welch**

Serious about research—and teaching

I arrived at the "future faculty" workshop hoping to glean tips on how to apply for tenure-track jobs. But when the time came to discuss my application materials, I was taken aback by the advice. "Cel," the faculty mentor said hesitantly, "I'm going to give you some harsh advice that I think you need to hear. Try to tone down your service and teaching—it doesn't make you look serious about research." I felt my cheeks burning in embarrassment as some of the other attendees nodded their heads in agreement. And I left feeling bewildered. I'd spent years building up my teaching and mentoring skills and devoting myself to serving as a positive role model for students. Shouldn't that hold value in a faculty job search? Why would I be penalized for it?

During graduate school, I had become passionate about making engineering accessible to everyone. I did not want others to feel like an imposter, as I did. As an undergraduate, advisers had told me I wasn't cut out for a career in engineering. Convinced an academic position was off the table, I planned to pursue a position in industry after graduating. I submitted applications to graduate school on a whim. I was floored when I was accepted.

When I became a teaching assistant, I started out badly. I mirrored the methods that had been used on me for all my life—relying on traditionally structured lectures and high-stress assignments—and was frustrated when students did poorly. I expected a lot from them but was not sure how to give them the resources they needed to succeed.

"I hope to land in a place that values my commitment to teach, serve, and care for my students."

In my third year, I decided to pursue a teaching certification. I learned that the methods I'd been using had long been debunked as ineffective, as they exacerbated performance gaps in students from marginalized backgrounds. I remembered that I excelled in courses with interactive lectures, small group work, low-stakes assignments, and caring professors who led with empathy. Such approaches, I realized, could help other marginalized students succeed.

I modified my teaching based on what I learned. Later that year, after I served as a precollege instructor at another university, one student wrote, "I wasn't sure if I could do engineering, but Dr. Welch made everything so simple and clear. I realized it's not that scary after all." This feedback filled me with immense satisfaction.

I hadn't forgotten about my research; that was always my main focus. But I enjoy multitasking, and teaching also benefited my research. In the lab, I adjusted my mentoring style to match teaching approaches that worked well—leading with empathy and providing individualized support—and my team exploded with productivity. Seeing evidence of my ability to lead a research team convinced me that I could excel in a tenure-track faculty position.

When I asked my mentors how to ensure I'd be competitive on the academic job market, they always said the same thing: be excellent in research, teaching, and service. But in the same breath they would remark on my teaching and service work, hinting that those who pursue such activities are best suited to serving as teaching faculty members or diversity, equity, and inclusion officers.

Then, at the future faculty workshop, it happened again. The advice rattled around my head long after. At first, I responded by reducing some of my nonresearch commitments. But then a mentee would need advice and I could feel the teaching skills I had spent time cultivating coming out. I vowed to not push those activities to the side to placate nearsighted views about how researchers on the academic track should allocate their time.

As I apply to faculty positions, I hope to land in a place that values my commitment to teach, serve, and care for my students. I'm not hiding my track record—I'm embracing it. Because I've seen first-hand the difference professors can make when they invest time and energy in teaching and mentoring. ■

E. Celeste "Cel" Welch is a Ph.D. candidate at Brown University and instructor at Columbia University.