

What happens when ordinary
people shape climate policy p. 518

Lockdown quiet clears the way
for better bird song pp. 523 & 575

Forgotten research on
animal cognition p. 530

Science

\$15
30 OCTOBER 2020
sciencemag.org

AAAS

SPRING OR FALL?

Genetic determination of Chinook salmon
migration timing pp. 526 & 609

CONTENTS

30 OCTOBER 2020 • VOLUME 370 • ISSUE 6516



522 & 557

Dogs, including this endangered dingo, were the first domesticated animal. Researchers have examined ancient dog genomes to identify the complex way in which man's best friend evolved.

NEWS

IN BRIEF

508 News at a glance

IN DEPTH

511 U.S. funds new fleet of probes to explore ocean biogeochemistry

Drifting Argo floats will track oxygen loss from global warming and the oceans' ability to soak up carbon *By P. Voosen*

512 Denisovan DNA found in cave on Tibetan Plateau

"Molecular caving" shows archaic humans visited high-altitude cave over long period *By A. Gibbons*

REPORTS pp. 579 & 584

513 Herd immunity? India still has a long way to go, scientists say

Cases are dropping, but critics question optimistic forecast *By V. Chandrasekhar*

514 How an immunologist pivoted to tackle COVID-19

For Akiko Iwasaki, the pandemic has brought new research priorities—and new urgency *By J. Couzin-Frankel*

515 Edit reviews without permission? Some journal editors say it's OK

Deleting offensive language is common, but some editors alter reviewers' overall recommendation *By C. O'Grady*

516 Efforts to control monkey brains get a boost

With open data project, primate optogenetics confronts stumbling blocks *By K. Servick*

FEATURES

518 Power to the people

Nations are turning to citizen assemblies to weigh up climate policies *By C. O'Grady*

PODCAST

INSIGHTS

PERSPECTIVES

522 Of dogs and men

Ancient genomes reveal the common history of human and dog *By P. Pavlidis and M. Somel*
RESEARCH ARTICLE p. 557

523 The quiet spring of 2020

Anthropogenic noise has no legacy effects on bird song *By W. Halfwerk*
REPORT p. 575

525 Speeding protons with metal vacancies

Membranes made from nanosheets derived from CdPS₃ have high proton conductivity *By F. Wang and J. He*
REPORT p. 596

526 Should I stay or should I go?

Pinpointing genetic alleles that influence the timing of salmon migration might aid restoration efforts *By G. McKinney*
REPORT p. 609

527 Will SARS-CoV-2 become endemic?

Reinfection, seasonality, and viral competition will shape endemic transmission patterns *By J. Shaman and M. Galanti*

530 Charles H. Turner, pioneer in animal cognition

An African American scientist's early discoveries are forgotten for all the wrong reasons *By H. S. Galpayage Dona and L. Chittka*
PODCAST

532 Paolo Sassone-Corsi (1956–2020)

Groundbreaking molecular biologist and epigeneticist *By E. Verdin*

POLICY FORUM

533 Political sectarianism in America

A poisonous cocktail of othering, aversion, and moralization poses a threat to democracy *By E. J. Finkel et al.*

BOOKS ET AL.

537 The vibrant lives of Neanderthals

A nuanced portrait of our evolutionary cousins encourages empathy and understanding *By E. Pomeroy*

538 A profound plan to save the seas

Rejecting piecemeal strategies, a conservationist encourages total ocean protection *By M. E. Hannibal*

LETTERS

539 Ban unsustainable mink production

By C. Xia et al.



Excavation in Baishiya Karst Cave identified Denisovan mitochondrial DNA dating to 100,000 years ago.

539 Support U.S. research during COVID-19

By M. M. Crow et al.

540 Hainan peacock-pheasant needs focused protection

By S. Guan et al.

RESEARCH

IN BRIEF

543 From *Science* and other journals

REVIEW

546 Cancer

Physical traits of cancer H. T. Nia et al.

REVIEW SUMMARY; FOR FULL TEXT:
DX.DOI.ORG/10.1126/SCIENCE.AAZ0868

RESEARCH ARTICLES

547 Structural biology

The coupling mechanism of mammalian respiratory complex I

D. Kampjut and L. A. Sazanov

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT:
DX.DOI.ORG/10.1126/SCIENCE.ABC4209

548 Hybrid perovskites

Atomic-scale microstructure of metal halide perovskite M. U. Rothmann et al.

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT:
DX.DOI.ORG/10.1126/SCIENCE.ABB5940

549 Microbiota

Multi-omics analyses of radiation survivors identify radioprotective microbes and metabolites H. Guo et al.

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT:
DX.DOI.ORG/10.1126/SCIENCE.AAY9097

550 Plant science

Receptor kinase module targets PIN-dependent auxin transport during canalization J. Hajnó et al.

557 Ancient dog genomics

Origins and genetic legacy of prehistoric dogs A. Bergström et al.

PERSPECTIVE p. 522

564 Coronavirus

The emergence of SARS-CoV-2 in Europe and North America M. Worobey et al.

REPORT p. 571

REPORTS

571 Coronavirus

Cryptic transmission of SARS-CoV-2 in Washington state T. Bedford et al.

RESEARCH ARTICLE p. 564

575 Pandemic pause

Singing in a silent spring: Birds respond to a half-century soundscape reversion during the COVID-19 shutdown E. P. Derryberry et al.

PERSPECTIVE p. 523

579 Human evolution

Denisovan ancestry and population history of early East Asians D. Massilani et al.

NEWS STORY p. 512; REPORT p. 584

584 Paleogenomics

Denisovan DNA in Late Pleistocene sediments from Baishiya Karst Cave on the Tibetan Plateau D. Zhang et al.

NEWS STORY p. 512; REPORT p. 579

587 Magnetism

Metal-organic magnets with large coercivity and ordering temperatures up to 242°C P. Perlepe et al.

592 Quantum systems

Parallel single-shot measurement and coherent control of solid-state spins below the diffraction limit S. Chen et al.

596 Proton membranes

CdPS₃ nanosheets-based membrane with high proton conductivity enabled by Cd vacancies X. Qian et al.

PERSPECTIVE p. 525

600 Topological systems

Generation of helical topological exciton-polaritons W. Liu et al.

605 Geophysics

Surface deformation associated with fractures near the 2019 Ridgecrest earthquake sequence X. Xu et al.

609 Migration

A complex phenotype in salmon controlled by a simple change in migratory timing N. F. Thompson et al.

PERSPECTIVE p. 526

DEPARTMENTS

502 Editorial

Untapped computer scientists
By Henry T. Frierson et al.

503 Editorial

It's just louder this time By H. Holden Thorp

630 Working Life

The things I wish I could say By Anonymous

ON THE COVER

Chinook salmon swimming in the Rogue River, Oregon. Genome sequencing and analysis of tribal fishery catch in the Klamath River, California, show that salmon ecotypes in diverse populations result from a simple Mendelian polymorphism that affects seasonality of spawning migration. Physiological and

reproductive differences long thought to distinguish ecotypes appear as a result of environmental effects from summer residence in different ecosystems. See pages 526 and 609. Photo: Mark Conlin/Alamy Stock Photo



AAAS News & Notes 541
Science Careers 614

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2020 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership, including subscription (12 months): \$165 (\$74 allocated to subscription). Domestic institutional subscription (61 issues): \$2148; Foreign postage extra: \$98. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request. GST #125488122. Publications Mail Agreement Number 1069624. Printed in the U.S.A.
Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. Postmaster: Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090-6178. Single-copy sales: \$15 each plus shipping and handling available from backissues.sciencemag.org; bulk rate on request. Authorization to reproduce material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for Science is 0036-8075. Science is indexed in the Reader's Guide to Periodical Literature and in several specialized indexes.

Untapped computer scientists

Henry T. Frierson is the associate vice president and dean of the Graduate School, University of Florida, Gainesville, FL, USA. hfrierson@ufl.edu

Juan E. Gilbert is the the Banks Family Preeminence Endowed Professor and Department Chair, Department of Computer Information Science Engineering, Wertheimer College of Engineering, University of Florida, Gainesville, FL, USA. juan@ufl.edu

Jeremy A. M. Waisome is a lecturer in the Department of Engineering Education, Herbert Wertheim College of Engineering, The University of Florida, Gainesville, FL, USA. jwaisome@eng.ufl.edu

Underrepresented U.S. women of color (UWOC)—Black Americans, Latina Americans, Native Americans, and Indigenous Pacific Islanders—have long been underrepresented among Ph.D. recipients in computer science. The number of computer science Ph.D. degrees annually awarded to UWOC is miniscule. Moreover, even when compared with white women, the discrepancies are overwhelming. The Computer Research Association's Taulbee Survey revealed that in 2019, only seven UWOC (U.S. citizens) received Ph.D. degrees in computer science—a mere 0.43% of all computer science doctorates attained that year and only 2.1% of the doctorates awarded to all women.

A sincere will to increase the number of UWOC in computer science graduate programs may be effective in getting more undergraduate UWOC to enroll. A start would be to encourage undergraduate UWOC to engage in computer science research and then to pursue a computer science graduate degree. Another tactic is to establish communication with other computer science departments to identify UWOC who are Ph.D. prospects and recruit them to graduate programs. Once enrolled, they must be provided financial support, effective mentoring, and a supportive environment. To reduce the sense of isolation, faculty leaders should avoid putting UWOC in situations where they are the only underrepresented persons of color. In addition, faculty and students should ensure that their UWOC colleagues engage with and build relational networks within the computer science academic and professional communities. This will require institutions and departments to understand the existing systems in which UWOC are minoritized and marginalized, and a commitment to dismantling them.

Many computer science departments express the desire to increase diversity by enrolling more U.S. citizens of color in their doctoral programs. Although strides have been made at universities such as Clemson, Auburn, and the University of Florida (UF), more favorable outcomes are welcomed. Computer science departments at Auburn, Clemson, and UF share similar features, such as established connections with historically black colleges and universities and the recruitment of faculty of color. For example, from 2014 to 2020, the UF Department of Computer and Information Science and Engi-

neering (CISE) enrolled 19 UWOC; thus far, eight have graduated with a Ph.D. degree. Of these eight women, six were Black Americans, one was Latina, and one was Native American. Among the remaining 11, three are currently doctoral candidates. Only two have left CISE without a Ph.D. degree. Compared to the national 0.43% UWOC (U.S. citizens) who graduated with a computer science Ph.D., the graduation trend at CISE is encouraging and is an example of what can be achieved.

To promote retention and graduation of UWOC, CISE places students with faculty who foster research cultures and provide holistic advising that enhance success. Further, students receive funding to attend conferences on

research and community building, such as the Grace Hopper Celebration of Women in Computing Conference, the Association for Computing Machinery Conference, and the Richard Tapia Celebration of Diversity in Computing Conference. Students and faculty attend these conferences to network and broaden the participation of underrepresented students in computing. CISE Ph.D. students are encouraged to guide undergraduate students in research and to connect with other graduate students and faculty across departments. They also engage in organizations that provide professional development opportunities, such as the Black Graduate Student Association and the Society of Hispanic Professional Engineers. En-

rolled students receive sufficient funding to relieve them from financial pressures and are placed in established student communities that provide ongoing peer support. Students thrive in CISE's culture of diversity. The department has worked to strategically infuse culturally relevant pedagogy into the curriculum. The department faculty currently has four Black American women, two Black American men, and one Hispanic American man. These faculty members provide underrepresented students of color with role models who come from similar cultures, which encourages success and mitigates feelings of isolation.

The UF CISE department has recognized an existing pool of potential computer scientists among UWOC and has shown the willingness to draw from a source proven to be fruitful. With sincere will, other institutions are encouraged to do the same.

—Henry T. Frierson, Juan E. Gilbert,
Jeremy A. M. Waisome

**“Although
strides
have been
made...more
favorable
outcomes are
welcomed.”**

It's just louder this time

As if there were any doubt that U.S. President Donald Trump has no respect for scientists, he now refers to public health scholars as “Fauci and all these idiots.” That’s how he’s describing experts in virology, immunology, epidemiology, and infectious disease. Never mind that after recovering from coronavirus disease 2019 (COVID-19), Trump suddenly became excited about future vaccines and “Regeneron,” which is what he calls monoclonal antibodies in general. (Regeneron Pharmaceuticals, Inc. is probably thrilled to have achieved the product-brand status of Xerox and Kleenex, but Eli Lilly also has developed promising monoclonals, and more are in clinical trials.) Apparently, no one told the president that scientists from these same fields—many of whom live in “Democrat-run cities” or college towns and are immigrants who wouldn’t be here under his policies—created these drugs and carried out the decades of science that made them possible. This paradox of loving the drug but hating the science is nothing new. It’s just louder this time.

Republican presidents were not always rhetorically hostile to science. As described earlier this year on this page, on the 50th anniversary of Earth Day, James Morton Turner and Andrew Isenberg carefully traced how the United States got to this point. In the 1970s, President Richard Nixon worked hard to pass important pieces of public health and environmental legislation that were approved with large bipartisan majorities in Congress. Then, when Ronald Reagan arrived as a candidate in the 1980 election, he advocated teaching creationism in public schools and mocked environmental science and regulation. In his brand of conservatism, the free market and American exceptionalism could not coexist with a shared responsibility for caring for the planet or its inhabitants.

Vice President Mike Pence is carrying on Reagan’s tradition. In a widely viewed speech on the House floor when he was a member of Congress, Pence extolled “intelligent design.” He cited a then-recent study of new fos-

sils, which enhanced our understanding of how human life unfolded on Earth, as evidence that evolution was invalid because scientists were always changing theories when new data were obtained. He was criticizing scientists for doing science, as my colleague Jon Cohen recently tweeted. If Pence thinks we can’t change our understanding with new data, then we’d have to go back to breathing phlogiston and being orbited by the Sun.

The paradox has played out for years. Many Republicans in Congress have been strong advocates for science funding, especially for the National Institutes of Health, although some simultaneously espouse antiscience views and embrace creationism. Biology is the study of evolution, and biomedicine is applied evolution. Why would creationists spend money to study and apply

this heresy? Because they want their new medicines. They want to tell their constituents that they are fighting diseases that are harming their families. Arguing for science funding by promising new cures has been a winning political strategy for the 75 years that the United States has had federally funded science.

A recent survey from the Pew Research Center found that only 20% of the political right has “a lot” of confidence in scientists. Yet when folks at this end of the political spectrum get sick, they want the

best treatments that secular academic medicine can provide. The consequences of this are profound and especially apparent in the COVID-19 crisis. The same politicians who are criticizing public health guidance are praising vaccines and antibodies without acknowledging that they come from the same principles and researchers as masks and social distancing.

When the presidential election is over, science will face an important choice. Should the scientific community try to get the missing 80% of the ideological right to understand its people and its methods? Or should science write it off as a lost cause and continue to take the funding while providing the outstanding new medicines?

—H. Holden Thorp



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

**“This paradox
of loving the
drug but hating
the science is
nothing new.”**

NEWS

IN BRIEF

Edited by Jeffrey Brainard

PUBLIC HEALTH

Europe starts shutdowns amid COVID-19's second wave

A second wave of coronavirus infections is prompting renewed restrictions across Europe, with countries from Ireland to Italy imposing nightly curfews and limiting the number of people allowed to gather. The hardest hit countries, Belgium and the Czech Republic, are reporting roughly 800 new cases per 100,000 residents per week. In Prague, the Czech army last week set up a 500-bed field hospital to treat COVID-19 patients, and experts said the Belgian health system was nearing collapse as large numbers of health care workers were infected. During the pandemic's

first wave in the spring, Belgium had one of Europe's highest rates of cases and deaths, but the Czech Republic was largely spared. France, Italy, Spain, and the United Kingdom, also hard-hit then, are now reporting tens of thousands of cases per day. Authorities in many countries are attempting to use targeted shutdowns to slow the spread of the virus in hot spots, leading to empty streets in front of Rome's Colosseum (above) and other locations. Ireland has imposed the strictest restrictions, with most nonessential shops closed nationwide through 1 December.

Extra pandemic deaths hit young

COVID-19 | The United States recorded some 299,000 more deaths than expected between late January and early October, with big increases among people 25 to 44 years old and nonwhite people, especially Latinos, the U.S. Centers for Disease Control and Prevention (CDC) reported last week. Roughly two-thirds of the excess deaths were directly attributable to COVID-19, the authors estimated; others may have been indirectly related to the disease,

resulting for example from lack of access to care for other conditions. The findings in CDC's *Morbidity and Mortality Weekly Report* are a comprehensive analysis of U.S. excess deaths, defined as those from all causes beyond what would be expected from historical trends, in this case average numbers from 2015 to 2019. The authors reported that deaths among 25- to 44-year-olds increased 27%, more than in any other age group. Deaths rose 54% for Latinos; from 29% to 37% for other non-white groups; and 12% for white people.

FDA allows first COVID-19 drug

THERAPEUTICS | The U.S. Food and Drug Administration (FDA) last week approved the first drug to treat COVID-19, Gilead's antiviral remdesivir, despite a lack of evidence of efficacy from the world's largest trial of it. The drug is now licensed to treat people as young as 12 years old who are hospitalized with COVID-19. Since the pandemic's early days, researchers have viewed remdesivir as one of the most promising drugs that might be repurposed

against COVID-19. In May, FDA granted an emergency use authorization for remdesivir after a U.S. trial found it reduced the median time to recovery in hospitalized COVID-19 patients from 15 days to 10. The drug has not been shown to reduce deaths among such patients, however. And recently published data from the World Health Organization's large Solidarity trial showed no clear mortality benefit or reduction of time to discharge from the hospital. FDA did not mention the trial in its decision to grant full approval.

Japan aims for carbon neutrality

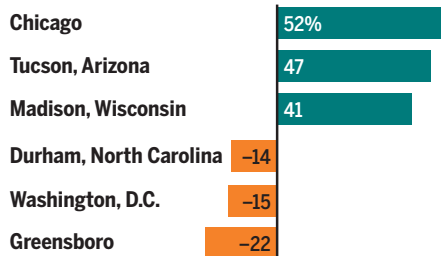
CLIMATE SCIENCE | Japan this week joined the growing list of countries that have set target dates to achieve carbon neutrality, pledging to reduce greenhouse gas emissions to zero by 2050. Prime Minister Yoshihide Suga told the nation's parliament that Japan plans to meet that target by bolstering conservation, increasing renewable and nuclear energy generation, and making a "drastic change" in the use of coal. Environmental organizations welcomed the announcement but are urging Japan to dramatically reduce emissions sooner, by 2030, to help keep global warming to within the 2015 Paris accord's preferred target of 1.5°C. They suggest the country ramp up renewable capacity, drop nuclear power, and phase out coal. Japan, which has the world's third largest economy, has been building new coal plants to replace electricity from nuclear reactors idled since the 2011 Fukushima Daiichi nuclear disaster.

Judges nix altering census count

DEMOGRAPHICS | A panel of U.S. judges last week blocked President Donald Trump's plan to leave out undocumented residents in deciding how many seats each state gets in the 435-member House of Representatives. The three judges ruled that subtracting undocumented residents from the final tally for the 2020 census, which the Census Bureau hopes to deliver by the end of the year, would violate the U.S. Constitution's mandate to count every resident, as well as federal law on how to carry out apportionment. The judges didn't prohibit Census officials from following a presidential order to calculate the number of undocumented residents in each state, which demographers say is a nearly impossible task. But the information "won't have any legal weight and Congress can ignore it," says Justin Levitt, a constitutional law scholar. The Department of Justice has appealed the ruling, by a court in Northern California, and on 30 November the Supreme Court will hear arguments in a similar case from New York.

Promises not always kept

Greensboro, North Carolina, and a few other cities are beating their goals for lowering carbon emissions, but many are not. Percentages show the difference between the actual and target level in the year most recently inventoried.



U.S. cities labor to cut emissions

CLIMATE SCIENCE | Two-thirds of the 100 largest U.S. cities that have adopted targets for reducing carbon emissions are falling short of meeting them, according to a study released last week. Setting those targets is still a worthwhile exercise given the lack of consensus at the federal level for how to combat climate change, say the

authors at the Brookings Institution. Some 45 cities have adopted targets; their levels are generally in line with the goals of the 2015 Paris climate accord and, if achieved, would be equal to removing 79 million cars from U.S. roads. But these "bottom-up" approaches have been buffeted by stiff economic and political headwinds, the report notes. Many cities lack plans for achieving the targets, and what plans do exist don't account for how population growth and economic development could boost greenhouse gas emissions.

Shortened treatment for TB

INFECTIOUS DISEASES | A novel combination of existing antibiotics can cure tuberculosis (TB) in 4 months instead of six, a study has found. TB specialists say the shorter regimen promises to be one of the most important changes to standard treatment since the 1960s. The cocktail includes four antibiotics, including one, moxifloxacin, previously reserved for multidrug-resistant TB, and a high dose of rifapentine, one of the most commonly



The OSIRIS-REx probe's sampling arm touched its target on an asteroid 321 million kilometers from Earth.

PLANETARY SCIENCE

NASA grabs its first asteroid sample

NASA's OSIRIS-REx probe last week took a bite of the asteroid Bennu and came away stuffed. The \$800 million spacecraft captured a sample of asteroid grit and will ferry it back to Earth—the agency's first ever collection of this kind. When OSIRIS-REx touched its target for 6 seconds with its robotic collector arm, Bennu's surface proved surprisingly soft, and the collector penetrated as deep as 50 centimeters. Photos indicate the mission collected far more than the 60-gram minimum required by NASA: The collector was overstuffed to the point that a mylar flap wouldn't close, allowing dust to slowly leak out. To prevent further losses, NASA planned to command the spacecraft to seal the sample in its return capsule this week. Next year, when Bennu's orbit is aligned favorably with Earth's, the spacecraft will depart, ultimately sending its cargo to a parachute landing in the Utah desert in 2023.

used drugs against TB. The combination proved safe and well tolerated in a randomized, controlled study involving 2500 patients. Presented last week at the 51st Union World Conference on Lung Health, the study took place in 13 countries. It was run by the Tuberculosis Trials Consortium in collaboration with the AIDS Clinical Trials Group.

Mexico cuts off science dues

SOCIETIES | Mexican scientists say a government funding cut may force them to leave a slew of international science organizations, which they say would isolate the country and deprive it of scientific opportunities. The National Council of Science and Technology has informed national science organizations that it will no longer fund their membership dues for at least nine international groups, *Science*

has learned. They include the Global Biodiversity Information Facility, the Ibero-American Federation of Physical Societies, the International Science Council, the International Union of Pure and Applied Physics, and the Third World Academy of Sciences. Finding the money elsewhere may be a tall order as some membership dues exceed \$100,000, the organizations say. The government, which has been cutting the societies' operating budgets since 2019, plans to use the funds to help fight the COVID-19 pandemic. The move is "very myopic," says María de la Luz Jiménez de Teresa de Oteyza, a former president of the Mexican Mathematical Society.

Berkeley renounces eugenics fund

HUMAN RIGHTS | The University of California (UC), Berkeley, said this week it would repurpose and rename a 60-year-old

fund established by a donor to fund eugenics research. UC Berkeley apologized for its School of Public Health's Genealogical Eugenics Institute Fund, which in recent years had provided \$70,000 annually for faculty and student research expenses unrelated to eugenics. UC Berkeley froze the \$2 million fund in 2018 after faculty members called its origins unethical; it was set up "for the primary purpose of improvement of the human race through research and education in the field of eugenics." The university is considering new uses for the money, including financial aid for members of underrepresented minority groups and public education about the damage wrought by the eugenics movement. For decades during the 20th century, people who were Black, poor, or had disabilities were sterilized involuntarily at disproportionately high rates in California and other states under eugenics laws.



INVASIVE SPECIES

U.S. murder hornets get rounded up

For the first time, entomologists last week found a nest of the Asian giant hornet—dubbed the murder hornet—in the United States. The largest hornet in the world (*Vespa mandarinia*), up to 5 centimeters long, rarely attacks people. But the hornets do capture honey bee hives, decapitating all the bees in hours and claiming the hive as their own. Native to Japan and now found across Asia, the species was spotted last year in Washington state and Canada's British Columbia province, most likely arriving as stowaways in shipping containers. Researchers trapped several hornets, fitted them with radio tracking devices, and followed one

to the nest in a tree trunk near Blaine, Washington. On 24 October, workers in protective suits wrapped the trunk in plastic and vacuumed about 100 hornets from the nest. The tree will be removed and searched for any remaining insects and signs that queens escaped capture. Researchers reported in the 6 October issue of the *Proceedings of the National Academy of Sciences* that the species could spread throughout coastal portions of Washington, Oregon, and British Columbia within 20 years, threatening economic losses for growers who depend on bees for pollination. Researchers continue to search for other nests.

PHOTO: ELAINE THOMPSON/AP PHOTO



Since 2014, researchers have deployed more than 150 biogeochemical Argo floats in the Southern Ocean.

paigms are expensive, and often limited to common ocean trade routes, says Ken Johnson, a chemical oceanographer at the Monterey Bay Aquarium Research Institute and a leader of the new program. The floats, he says, are “not as good as a ship, but they’re there all the time.”

In addition to standard Argo measurements of temperature and salinity, the new floats will have sensors measuring oxygen, sunlight, particles, chlorophyll (a gauge of phytoplankton abundance), nitrate (a key nutrient), and pH (acidity). Researchers will be watching that last reading closely, because acidity reflects both the ocean’s uptake of CO₂ and its pernicious effect. When the gas dissolves in seawater, it forms carbonic acid that eventually splits into bicarbonate and hydrogen ions, the latter increasing the water’s acidity. Ecologists are concerned that acidification, already 30% worse in surface waters than preindustrial times, will make it more difficult for some phytoplankton, corals, bivalves, and many other species to assemble their shells of calcium carbonate.

Researchers have been testing more than 150 prototype BGC floats in the Southern Ocean since 2014, and the findings are tantalizing. This ocean, encircling Antarctica, is home to two seasonally contrasting CO₂ fluxes. In the summer, algae draw down CO₂. But in the winter, ancient carbon, stashed away for centuries in deep currents traveling south from the Atlantic and Pacific oceans, wells up, some escaping into the atmosphere. Scientists have long assumed that the carbon absorption dominates, with the Southern Ocean accounting for a significant share of the oceans’ global CO₂ storage. But because almost no ships ventured to the stormy, cold Southern Ocean, they couldn’t check whether the winter release really was smaller.

In the past few years, Alison Gray, an oceanographer at the University of Washington, Seattle, and her colleagues have used the trial BGC floats to show that the winter exhalation of CO₂ comes close to canceling out the summer’s gains. She is now seeking to understand why the outgassing seems to peak on the Pacific side of the ocean. Surface winds or currents smashing into undersea ridges might sweep up more of the deep waters there. Another explanation is biological: The region’s phytoplankton might be less capable than their Atlantic counterparts at staunching the CO₂ outflow.

The BGC floats could also shed light on a less-known ocean trend: a slow drop in oxygen. Since the mid-20th century, it has declined by some 2%. Much of the loss takes

IN DEPTH

OCEANOGRAPHY

U.S. funds new fleet of probes to explore ocean biogeochemistry

Drifting Argo floats will track oxygen loss from global warming and the oceans’ ability to soak up carbon

By **Paul Voosen**

A single drop of seawater holds millions of phytoplankton, a mix of algae, bacteria, and protocellular creatures. Across the world’s oceans these photosynthesizing microbes pump out more than half of the planet’s oxygen, while slowing climate change by capturing an estimated 25% of the carbon dioxide (CO₂) released from humanity’s burning of fossil fuels. But the scale of this vital chemistry is mostly a guess, and there’s little sense of how it will change as temperatures rise. “What’s happening out there? We have no idea really,” says Susan Wijffels, a physical oceanographer at the Woods Hole Oceanographic Institution.

Soon, 500 drifting ocean floats studded with biogeochemical sensors will deliver

answers. This week, the National Science Foundation (NSF) announced it will spend \$53 million to fund the new floats, marking the first major expansion of the Argo array, a set of 4000 floats that for 15 years has tracked rising ocean temperatures. “This is going to be revolutionary,” says Wijffels, a leader of the original Argo program.

The biogeochemical (BGC) Argo floats, in development for nearly as long as Argo itself, will operate much like their forerunners. After being tossed off a ship, each of the skinny, 1-meter-tall floats drifts with deep ocean currents 1000 meters down. Every 10 days or so, it uses an oil-filled bladder to change its density, dropping to 2000 meters and then slowly rising to the surface, where it beams the resulting profile home. Although instruments lowered from ships can make deeper and more precise measurements, ship cam-

place in anoxic dead zones like one in the Gulf of Mexico, where nutrient runoffs from overfertilized lands lead to algal blooms, and, eventually, bacterial surges that use up oxygen in a frenzy of aerobic decomposition, suffocating fish. But researchers have also found that vast tracts of open ocean are losing oxygen because of warming, which limits water's ability to hold oxygen, and decreased downward mixing. BGC-Argo could reveal the true extent of these minimum zones. "If you eat seafood or care about sea life, you should care about deoxygenation," Gray says.

The floats aren't only useful for large-scale trends. For instance, the trial floats discovered massive phytoplankton blooms in the Southern Ocean, far from typical nutrient sources like melting sea ice, suggesting a seafloor hydrothermal vent was providing the nutrients. The floats should also be able to detect short-lived phenomena such as the oxygen loss after a hurricane's upheaval, or the injection of nutrients for species in the North Atlantic twilight zone, up to 1000 meters down, that come from late-winter plunges of surface water. "The ability to observe these biogeochemical properties in three dimensions is going to be huge," says Katja Fennel, an oceanographer at Dalhousie University, who is pushing for Canada to add 40 floats to BGC-Argo.

Researchers plan to begin deploying the new floats next year in the equatorial Pacific, where El Niño and La Niña drive large temperature swings every year or two. The floats could show how the swings affect the ability of phytoplankton to soak up carbon, offering clues to how a warming climate will change the ocean's overall carbon uptake.

After the first 5 years of NSF financing, Johnson and his colleagues would likely need support from the National Oceanic and Atmospheric Administration, which pays for much of the main U.S. Argo fleet. Expanding to 1000 floats, as the team hopes, would require funding from France, Australia, Canada, China, and other countries. At some point, they hope the Argo fleet will be bolstered by 1200 deep Argo floats, which can sink to 6000 meters without crushing (*Science*, 8 September 2017, p. 956). But money for upgrading the fleet, which would remain at 4000 total floats, has been elusive—even though Argo is a relative bargain compared with ships.

If anything, the pandemic has underscored that point, Johnson says. "We're all sitting at home now," he says. Yet out in the ocean, the robotic fleet calls home week after week, capturing otherwise invisible changes. "Ten years ago, the ship was the observing platform," Johnson says. "Now, it's the tender." ■



Archaeologists wearing anticontamination gear sampled Baishiya Karst Cave on a winter night.

ARCHAEOLOGY

Denisovan DNA found in cave on Tibetan Plateau

"Molecular caving" shows archaic humans visited high-altitude cave over long period

By Ann Gibbons

For today's Buddhist monks, Baishiya Karst Cave, 3200 meters high on the Tibetan Plateau, is holy. For ancient Denisovans, extinct hominins known only from DNA, teeth, and bits of bone found in another cave 2800 kilometers away in Siberia, it was a home. Last year, researchers proposed that a jawbone found long ago in the Tibetan cave was Denisovan, based on its ancient proteins. But archaeologist Dongju Zhang of Lanzhou University and her team wanted more definitive evidence, including DNA, the molecular gold standard. So in December 2018, they began to dig, after promising the monks they would excavate only at night and in winter to avoid disturbing worshippers.

After working from dusk to dawn while temperatures outside plunged to -18°C , then covering traces of their dig every morning, the scientists' persistence paid off. On p. 584, Zhang's team reports the first Denisovan ancient DNA found outside Denisova Cave: mitochondrial DNA (mtDNA) gleaned not from fossils, but from the cave sediments themselves. Precise dates show the Denisovans took shelter in the cave 100,000 years and 60,000 years ago, and possibly as recently as 45,000 years ago, when modern humans were flowing into eastern Asia.

The find shows that even though their bones are rare, "Denisovans were widespread in this hemisphere," says University of Oxford geochronologist Tom Higham, who was not part of the study. It also ends a long quest for Denisovan DNA outside Siberia. "Every year, I've said we will find this," says co-author Svante Pääbo of the Max Planck Institute for Evolutionary Anthropology (EVA). "It's been a decade."

The presence of Denisovan DNA in the genomes of living people across Asia suggested these ancient humans were widespread. But the partial jaw from Baishiya Karst Cave was the first fossil evidence. Zhang and her colleagues identified the jaw as Denisovan based on a new method that relies on variation in a protein (*Science*, 3 May 2019, p. 418). Some researchers questioned the claim, however, because the method was new, and no one knew where in the cave the jaw had been found.

Those questions are likely to fade. The dig, led by Zhang and Fahu Chen of the Institute of Tibetan Plateau Research, took many sediment samples and found charcoal from fires, 1310 simple stone tools, and 579 pieces of bone from animals including rhinos and hyenas. Paleogeneticist Qiaomei Fu of the Institute of Vertebrate Paleontology and Paleoanthropology in Beijing managed to extract hominin mtDNA from the sediment itself. The mtDNA, perhaps shed in poop or urine,

PHOTO: DONGJU ZHANG/LANZHOU UNIVERSITY

most closely matched that of Denisovans.

Meanwhile, geochronologists led by Bo Li and Zenobia Jacobs of the University of Wollongong dated material from those same sediment samples. They used optical dating to reveal when light last struck mineral grains in the samples, showing when each grain was buried. The four layers that yielded Denisovan mtDNA were laid down 100,000, 60,000, and as recently as 45,000 years ago, although the younger sediments were disturbed.

The dates for the older sediments seem highly reliable, says Higham, who dated Denisova Cave. And by showing DNA and dates can be gleaned from the same sediment samples, the work opens “a new era of molecular caving,” says Oxford geochronologist Katerina Douka.

The charcoal in the cave shows its occupants built fires. They also used simple stone tools, and, from the cave’s high opening, must have spied on animals grazing in the meadows below. Some may also have been on the lookout for modern humans, who were in the region by 40,000 years ago.

In a separate study on p. 579, Pääbo reports extracting modern human DNA, the oldest yet in Asia, from 34,000- and 40,000-year-old fossils from what is now Mongolia and from near Beijing, respectively. Those genomes included Denisovan DNA, the legacy of mating that happened roughly 50,000 years ago. But the Denisovan sequences differed from those found in living New Guineans and Australian Aboriginals. *Homo sapiens* must have met and mated with two populations of Denisovans—one in mainland Asia and one in Southeast Asia, says EVA paleogeneticist Diyendo Massilani—further evidence that they were once numerous and wide-ranging.

The Denisovans bequeathed a particular genetic gift to modern Tibetans: a “super-athlete” variant of a gene, called *EPAS1*, that helps red blood cells use oxygen efficiently and is found in Denisovans from Denisova Cave. Zhang and her colleagues think the Tibetan Plateau Denisovans may have been adapted to life at high altitude, and that *EPAS1* may have spread widely among them, before they handed it on to modern Tibetans.

But molecular dating suggests *EPAS1* spread rapidly only in the past 5000 years. And natural selection would have favored that gene variant only in people who lived at high altitude year-round, says archaeologist Mark Aldenderfer, professor emeritus at the University of California, Merced. The Denisovans may have lived only seasonally in the cave. Zhang’s team will need to find nuclear DNA to test its hunch.

Zhang expects more digs at the cave will clarify the issue with DNA and perhaps fossils. “The study of this cave is only beginning,” she says. ■

COVID-19

Herd immunity? India still has a long way to go, scientists say

Cases are dropping, but critics question optimistic forecast

By **Vaishnavi Chandrashekhar**, in Mumbai

Last week, a panel of leading scientists appointed by the Indian government delivered a startlingly optimistic message: The world’s second largest COVID-19 epidemic has rounded a corner. India’s daily number of cases has declined by roughly half in the past month, and a new mathematical model suggests “we may have reached herd immunity,” the panel wrote in a paper published online by *The Indian Journal of Medical Research*. Assuming measures such as social distancing, wearing masks, and hand washing remain in place, the group said the pandemic could be “controlled by early next year.”

But other scientists say the model overestimates the number of people already infected and warn that with colder temperatures and several religious holidays approaching, India may well see a second wave. The positive national trends hide a more complex picture, suggests Giridhar Babu, an epidemiologist with the Public Health Foundation of India. He believes the virus may have burned through large, densely packed populations but will continue to spread in rural areas, at a lower rate, for many months: “We still have large numbers of people for the virus to go through.”

The encouraging projections come from the National Supermodel Committee, which modeled the past and future of India’s epidemic at the government’s request. Its work suggests 380 million Indians had already been infected by mid-September and that there might be “minimal active symptomatic infections” by late February 2021 if control measures continue. (The study also concluded that by flattening the curve, India’s lockdown in the spring saved up to 2.6 million lives.)

The model did not take regional differences in viral spread into account. But Babu notes that serological surveys—which test

for antibodies in a population to gauge the fraction already infected—have found much higher infection rates in Indian cities, and in particular in slum areas. Studies in August, for example, found antibodies in 41% of residents of Mumbai’s slums, compared with 18% elsewhere in the city. “Many dense pockets that can be easily infected have [likely] already been infected,” Babu says. That could help explain the slowdown in new infections but suggests many more people remain vulnerable.

Gautam Menon of Ashoka University, a co-author on several COVID-19 modeling studies, adds that the model suffers from “a lack of epidemiological realism” because it assumes an unusually large fraction of infected people remains asymptomatic. He says 200 million to 300 million is a better estimate for the number of infected people.

Experts agree that spread in rural areas, home to more than half of the population, is a challenge to both monitor and fight. Testing isn’t easily available in many small towns and villages; serosurveys suggest official testing, now at 1 million per day, vastly undercounts actual cases. Some states rely heavily on so-called rapid antigen tests, which

range widely in sensitivity. Meanwhile, the health infrastructure in many of India’s rural states is weak, making it harder to treat patients.

More granular data might help scientists better understand the pandemic’s trajectory. The national epidemic is a “figment of statistical imagination,” says T. Jacob John, former head of the department of virology at Christian Medical College; instead, “There are 100 or more small epidemics in different states and cities, rising and falling at different times,” he says.

Babu warns against complacency. “The decline [in cases] is real and valid, but no one should rejoice yet,” he says. “It only means that the first set of formidable challenges is over, and the next set is beginning.” ■

“There are 100 or more small epidemics in different states and cities, rising and falling at different times.”

T. Jacob John,
Christian Medical College

Science’s COVID-19 reporting is supported by the Pulitzer Center and the Heising-Simons Foundation.

Vaishnavi Chandrashekhar is a journalist in Mumbai.

VOICES OF THE PANDEMIC

How an immunologist pivoted to tackle COVID-19

For Akiko Iwasaki, the pandemic has brought new research priorities—and new urgency

By Jennifer Couzin-Frankel

Until this year, Akiko Iwasaki's lab had never handled so many tubes of human blood. "We were mostly working with mouse models," says the Yale University immunologist, who speaks precisely and thoughtfully. "We used to look at the data and contemplate it." Then COVID-19 struck, and such unhurried musings flew out the window. In a matter of weeks, Iwasaki overhauled her research to launch a slew of studies on how the new virus, SARS-CoV-2, takes its toll on patients. She and her nearly two dozen lab members know their discoveries could impact people falling sick right now. "Every minute counts."

In the months since, she has produced a string of high-profile papers by redirecting her expertise in the immune system, honed in mice, to questions such as why men are more likely to get severely ill and how immune responses in hospitalized patients can help predict their prognosis. Now, she is turning her attention to long-haulers, people who suffer a bout with the virus and don't fully recover.

Iwasaki has had decades of practice adapting to new circumstances. As a child growing up in rural Japan, she dreamed of becoming a poet, turned off science by her physicist father's immersion in his profession. "We'd go on vacation and he'd bring papers with him," she says, laughing. "I thought, 'What kind of life is this?'" But when a high school teacher hooked her on math, she began to reconsider. Soon after, 9 months as an exchange student in Canada left her itching to escape the expectations for a woman in Japanese society—marry a nice man and have a family. Her mother, who worked at a local radio station, had endured jeers from co-workers for sticking with the job while raising three children. "Knowing how much she stood up for always has stayed with me," Iwasaki says.

So she reimagined her future, embracing science and leaving Japan. She enrolled as an undergraduate at the University of Toronto—falling hard for immunology her senior year there—and stayed on for graduate school. Twenty years ago she founded her lab at Yale,

where she studies how the body responds to and combats viruses. "Having to adapt to different situations throughout my life," she says, "prepared me [for] a different virus."

The shift called for new science, new collaborations, and new skills. In February, Iwasaki's lab joined a universitywide testing effort for SARS-CoV-2 led by Albert Ko, Nathan Grubaugh, and Anne Wyllie at Yale's School of Public Health. Alice Lu-Culligan, a

patients in Connecticut to map the spread there and across the United States. She also launched a separate study to examine patients' immune responses, recruiting 113 people with COVID-19 at Yale New Haven Hospital and redeploying skills in her lab to make the project happen. Post-doctoral fellow Carolina Lucas had been studying the mosquito-borne chikungunya virus, and her project was housed in a biosafety level 3 lab at the university, the kind used for hazardous pathogens. "Akiko asked me to coordinate this," says Lucas, who quickly agreed.

Every few days, the team collected samples from the nose, throat, and blood of patients. There were "all these weird immune responses being engaged," Iwasaki says. In severe cases, the immune system churned out a flood of cytokine proteins. Lucas, Iwasaki, and others found four immune signatures that appeared to correlate with later outcomes. That paper appeared in *Nature* in July.

Swiftly, the scientific questions mushroomed. In mid-March, the Yale hospital treated a woman with COVID-19 who was in her second trimester of pregnancy. The woman lost her fetus—and a private tragedy became interwoven with urgent questions about whether the virus could infect the placenta and pose a danger to a pregnancy. A collaborator of Iwasaki's secured permission to collect the placenta, and late one night, Lu-Culligan retrieved it. Until that moment, the only placentas Lu-Culligan had seen belonged to mice. "This is big and bloody," she says, and as she stared at it under a biosafety hood, "I'm thinking, 'I don't know what I'm doing here.'"

In that case, the virus had indeed infected the placenta, and Lu-Culligan began to collaborate with Yale obstetricians to recruit women delivering at the hospital who were positive for the virus to study their placentas, too. That paper is nearing completion.

Meanwhile, Iwasaki began to investigate sex differences and found the male immune system is more likely to spark a harmful inflammatory response to the virus, whereas in women, T cells that fight



"Having to adapt to different situations throughout my life prepared me [for] a different virus."

Akiko Iwasaki, Yale University

graduate student of Iwasaki's who had been studying the immune system during pregnancy in mice, recalls the scramble. Lab members scouted for supplies such as swabs and equipment. "We were going around our floor, to the neighboring labs, seeing how many PCR [polymerase chain reaction] machines they had," Lu-Culligan says. It was "full-on sprint mode, collaboration and chaos."

Iwasaki's lab began to help Grubaugh's group sequence viral genomes from early

it off are activated more robustly. These distinctions, she reported in an August paper in *Nature*, might help explain why men who are infected tend to fare worse than women.

Iwasaki's juggling act impresses her colleagues. "She's made it seem so effortless, even though I know it's probably not effortless at all," says Angela Rasmussen, a virologist at Columbia University's Mailman School of Public Health. Iwasaki's husband, Ruslan Medzhitov, is also a well-known Yale immunologist (they discuss COVID-19 while walking their dog), and the pair has two daughters, ages 11 and 13. Iwasaki fears the pandemic is widening the gender gap in science as women face disproportionate pressure to support their children when schools are closed. Her husband drives their daughters to in-person school each morning, but with COVID-19 cases climbing, she wonders how much longer schools will be open.

Solutions to a COVID-19-fueled gender gap in science are elusive, she says, "other than to really have a different mindset about evaluating progress in science during this time." Iwasaki has long advocated for female and minority scientists on Twitter, where she has 80,000 followers. In one post, she minced no words in advising female scientists who worry about pregnancy torpedoing a job interview: "If they don't welcome you with open arms and offer child care options, they don't deserve you."

Her advocacy goes beyond rhetoric. Lu-Culligan met Iwasaki at a luncheon for women in science at Yale, while struggling with bullying and harassment in another lab. Iwasaki said, "We have to get you out of there," Lu-Culligan recalls. A few months later, the young scientist abandoned more than 2 years of graduate work to start over with Iwasaki—later learning that she wasn't the first person her new mentor had rescued from a miserable experience elsewhere.

Some 8 months into the pandemic, lab life has settled down—somewhat. Iwasaki's latest passion is long-haulers who can't shake symptoms like fatigue and brain fog. Volunteers find her via word of mouth. The project faces hurdles, though: Iwasaki is hunting for a facility to draw blood from her volunteers, who are still symptomatic and potentially contagious. With many competing studies, such space is at a premium and she hasn't yet been able to secure any. She's also racing to apply for grants to fund the project.

"We really want to get to the bottom of what's going on," she says impatiently. Until then—along with so many other researchers—she'll be in overdrive. ■

PEER REVIEW

Edit reviews without permission? Some journal editors say it's OK

Deleting offensive language is common, but some editors alter reviewers' overall recommendation

By Cathleen O'Grady

Fiona Fidler, a metaresearcher at the University of Melbourne, was outraged. She had discovered that her appraisal of a submitted paper had been changed before being sent to the author, sometimes drastically. The words "very sympathetic" had become "generally sympathetic." "This one is a good example" ended up as "this one still needs work." Worst of all, she felt that the bottom line of her peer-review report to the journal *Educational and Psychological Measurement*, recommending that it accept the paper with minor revisions, was misrepresented in the editor's rejection letter to the author.

"I had never experienced anything like this before," Fidler says about the 2012 incident. She demanded explanations from the journal editor. And she later partnered with the snubbed paper author, Rink Hoekstra, a psychologist at the University of Groningen,

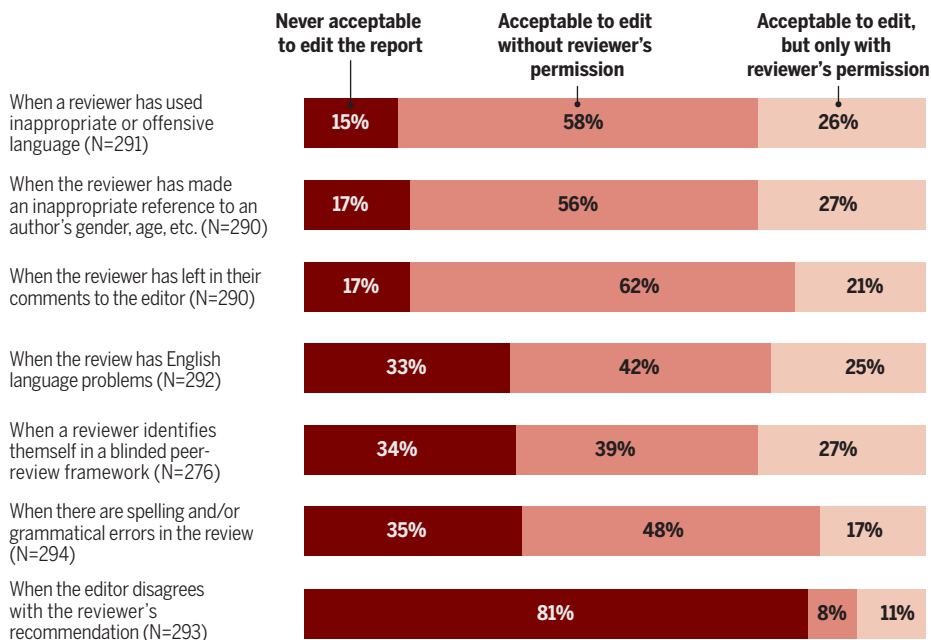
to find out how widespread this practice was.

With colleagues, they've now surveyed 322 editors at high-impact journals across ecology, economics, medicine, physics, and psychology on when they think altering peer-review reports is justified. Published as a preprint earlier this year at the Open Science Framework and now under review at *eLife*, the survey reports that 91% of the respondents identified at least one situation in which they would edit a report. More than 80% said they would do so if a reviewer used offensive language or made inappropriate personal comments about the authors. But 8% said they would change the reviewer's overall recommendation—even without their permission, a finding that shocked Hoekstra. He believes that's probably an underestimate, given the stigma of admitting to such dubious behavior in a survey. "I think there are probably even more who actually do it."

The survey comes amid a wider push for consistent guidelines. There's little dispute

Red lines

A survey of 322 editors at high-impact journals has found a wide range of attitudes toward altering peer-review reports. About 8% believe it's OK to change the reviewer's recommendation, even without permission.



that editors should intervene when peer-review comments are hostile. Such comments have a disproportionate impact on minorities and other marginalized groups, inducing self-doubt and harming productivity, according to a 2019 *PeerJ* paper. It's not just ethical to edit such review reports—it's essential, says Seth Leopold, editor-in-chief of *Clinical Orthopaedics and Related Research*.

Jane Alfred, director of Catalyst Editorial, which offers training on research integrity and publication ethics, thinks it's best to return reports containing hostile language and ask the reviewer to change it. Reviewers are often grateful for the chance to revise ill-considered comments, she says. But Leopold says this may be impractical at fast-paced journals, and trying to educate reviewers who make unprofessional comments is likely futile. It's better, he says, not to ask that person to review again.

Changing a reviewer's recommendation is another matter, says Howard Browman, a council member of the nonprofit Committee on Publication Ethics, which is developing guidance on how to navigate the ethics of editing reviewers' comments. "It's so obviously something you wouldn't do," he says.

Yet it apparently happened to Fidler, who stumbled on the changes to her report by accident. Although the review process was double-blinded, she recognized the paper as Hoekstra's because she had seen him present the work, on Ph.D. students' statistical reasoning, at a conference in Slovenia. So after sending in her review, she emailed him to congratulate him and tell him that her review was nit-picky but positive.

That same day, Hoekstra's paper was rejected by the journal. "The reviewers have spoken in nearly a single voice in their recommendation to me that I decline publication of the paper in its current form," wrote the editor-in-chief, George Marcoulides, a research methodologist at the University of California, Santa Barbara. Although editors may override reviewer recommendations, the normal practice is to explain this in the decision letter. Hoekstra wrote back to Fidler, attaching the two reviews he got from the journal and asking for suggestions of other journals that might publish the work.

When she saw the reviews, Fidler cottoned on. "I think we have a controversy on our hands," she wrote to Hoekstra. She downloaded her review from the online journal system, and compared it, word by word, with the review Hoekstra received. Her sentence saying her concerns were minor had

been deleted. Other sentences criticizing Hoekstra's methods had been added. She emailed Marcoulides, who replied that the journal computer system sometimes blended and distorted reviewer comments.

Fidler calls that explanation "preposterous," saying it would be impossible for a technical glitch to create "perfectly grammatical sentences that are exactly opposite in meaning." She reached out to contacts at SAGE, the journal's publisher, who said they would investigate, but she heard no more from them. In an email to *Science*, a SAGE spokesperson wrote that the publisher "addressed the issue directly with the editor at the time" and that Marcoulides now double-checks reviewer comments for consistency.

Later in 2012, Marcoulides invited Hoekstra to resubmit his paper. This time, Fidler's review made it through intact, and Hoekstra's paper was accepted—pivotal

for him winning a permanent job at Groningen. In an email to *Science*, Marcoulides wrote that the reviewer comments automatically attached to his decision letter were "distorted" and his edits were intended to clarify his interpretation of Fidler's assessment. "In hindsight, I should have contacted her rather than attempting to resolve the problem on my own," he wrote. He added that he still sometimes edits reports for clarity or to remove inappropriate language.

Few journals offer explicit guidance on when editing peer-review

reports is and isn't permissible. Alfred says they ought to, and should also allow reviewers to opt out from being edited. No matter how well-intentioned editors may be, she says, clear policies will ensure a transparent and unbiased process. Many journals have a safeguard: They share all reviews and the editorial decision with reviewers, allowing them to see how their comments were communicated to the authors. But about 20% of the editors in the survey report that their journals do not send out either the reports or the decision letters to reviewers.

Simine Vazire, editor-in-chief of *Collabra: Psychology* and a colleague of Fidler's, argues for a bright line on the question: no edits without reviewer permission. Her journal has no policy on the issue, but she is considering proposing one. Without clear boundaries, she says, it becomes easy to rationalize changes, adding that journal editors have lots of power and little accountability. "There's no one watching over editors," she says. "I think it's especially important that they have really hard and fast rules for themselves." ■

"I think it's especially important that [editors] have really hard and fast rules for themselves."

Simine Vazire,
Collabra: Psychology

NEUROSCIENCE

Efforts to control monkey brains get a boost

With open data project, primate optogenetics confronts stumbling blocks

By Kelly Servick

When neuroscientist Sébastien Tremblay set out to manipulate monkeys' brains with light, colleagues had sobering advice: "It's more difficult than it sounds."

Tremblay, who works in neuroscientist Michael Platt's lab at the University of Pennsylvania, uses light to activate or silence precise groups of neurons and probe their role in brain function. The method, called optogenetics, works well in rodents, but studies in nonhuman primates are critical if it's ever going to become a therapy for humans—to suppress seizures, for example, disrupt tremors in Parkinson's disease, or even project images into the brain of a blind person (*Science*, 8 November 2019, p. 671).

But in spite of more than 10 years of work, progress has been slow. The tools for rendering cells light sensitive were largely refined in rodents and behave unpredictably in monkeys. It's hard to illuminate enough tissue in large primate brains to reliably change animals' behavior. Researchers have devised their approaches by trial and error, often without knowing what had or hadn't worked for others.

Tremblay, Platt, and colleagues from 45 primate optogenetics labs in nine countries hope to change that with the Nonhuman Primate Optogenetics Open Database, which published its first results last week. The database contains minute details of successes and failures, many of which have gone unpublished. And if it can be sustained, it may soon include tests in monkeys of promising new optogenetic tools. The open-data approach "is tremendously powerful, tremendously useful to the community," says Hongkui Zeng, a neuroscientist who develops optogenetic tools for mice at the Allen Institute for Brain Science and was not involved in the project.

In optogenetics, researchers endow brain



cells with a gene for one of several opsins, light-sensitive proteins from microbes. These proteins can influence the flow of ions in and out of a neuron to control whether it fires an electrical signal. Depending on the opsin, researchers can excite or inhibit neurons by shining light on them, usually via an implanted optical fiber.

Strains of mice have been genetically engineered to express opsins in their brains from birth. But for now, getting an opsin into monkey neurons means infecting the cells with a virus injected through a hole in the skull. Along with opsin DNA, the virus typically carries a sequence called a promoter, which restricts the opsin's expression to certain cell types.

There's no proven formula for getting monkey brain cells to make opsins. In the hunt for the right combination of viral strains and promoters, "we kind of entered this voodoo land," says Arash Afraz, a neuroscientist at the U.S. National Institute of Mental Health. Scientists relied on rumors of other labs' successes and failures, he says, and were afraid to vary a recipe once they got it working. Unlike with plentiful mice, researchers couldn't afford to use lots of monkeys to hone their technique, he adds. "We value them more. They have names. We view them as our colleagues, in a sense."

Afraz hopes the database, which he contributed to, will minimize wasted effort by pooling the field's failures. It catalogs 1042 viral injections performed in non-human primates, 552 of them previously unpublished. Seven-tenths of the experiments were in rhesus macaque monkeys. Tremblay can't be sure the database is exhaustive, but the 66 groups he invited to contribute—identified through publications and referrals from colleagues—represent the majority of labs active in the field, he says.

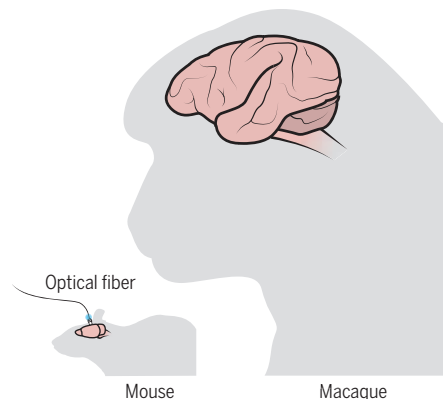
In a 19 October paper in *Neuron* introducing the database, the team estimates the success rate of the most commonly used vectors, promoters, and opsins in the data set. About half of the experiments in monkey brains looked for changes in neural activity after cells were hit with light; 69% found a strong effect. Of the 20% of experiments that aimed to influence an animal's behavior—to prompt an eye or hand movement, for example—nearly half saw a weak effect or none at all.

Failure likely discouraged some researchers from publishing studies, says Julio Martinez-Trujillo, a neurophysiologist at Western University and contributor to the project. His group has tried, without success, to evoke eye movements in one macaque and to impair working memory function in another. "This is the first paper that shows our experience," he says.

Such attempts probably fail in part because the virus doesn't reach enough of the

Scaling up

Monkey brains have been harder to manipulate with light than mouse brains, in part because they're much larger.



Methods to influence rodent neurons using light have not translated smoothly into primate brains.

brain, Tremblay says. A single injection can infect about 1 cubic millimeter of tissue—a broad swath of a mouse's brain, but a puny fraction of a monkey's. And scientists want to avoid multiple injections that could cause excessive tissue damage. Instead, some labs are trying to send the virus farther by injecting it at high volumes and pressures, a technique called convection-enhanced delivery.

Others hope to eliminate the need for brain injections by designing viruses that are small enough to cross into the brain via its tiny capillaries after being infused into a vein. In a June preprint on bioRxiv, neuroscientist and bioengineer Viviana Gradinaru and her team at the California Institute of Technology describe such an engineered virus that selectively infects the neurons of a marmoset.

Delivering light to large brains is a hurdle as well. "Say I am using a 200-micron-diameter fiber optic for stimulating my mouse brain," Afraz explains. "To scale that up, I'd have to stick a flashlight in the monkey's head." In a bioRxiv preprint last month, Afraz and colleagues describe a possible alternative: a 5-square-millimeter array of 24 light-emitting diodes (LEDs), each of which can produce as much light as a typical optical fiber. By laying this array over a monkey's cortex, researchers might illuminate a relatively broad brain area without multiple implanted fibers, Afraz says. They can also use individual LEDs to excite separate parts of the cortex in precise patterns.

Other groups are developing more sensitive opsins so that weaker light can affect more distant tissue. In a study in mice reported on 5 October in *Nature Biotechnology*, a group led by Stanford University neuroscientist Karl Deisseroth—one of the original developers of optogenetics—used a highly sensitive opsin called ChRmine to activate neurons several millimeters below the brain's surface with light from outside the rodent's skull.

"I can't wait to test them," Laval University molecular biologist Marie-Ève Paquet says of these ultrasensitive opsins. She's part of a Canadian collaboration that tests and disseminates emerging optogenetics tools. As opsins, promoters, and viruses make their way to participating research groups, Paquet's team plans to upload its results to the new database.

To keep the database up to date, she says, "the community really has to be motivated," especially because she expects the next few years to bring a boom in studies to influence and understand the brain circuits of some of our closest animal relatives. ■



The 110 members of the U.K. Climate Assembly were randomly selected.

FEATURES

POWER TO THE PEOPLE

Nations are turning to citizen assemblies to weigh up climate policies

Until recently, Sue Peachey, an apartment building manager in Bath, U.K., didn't think much about climate change. "I did my recycling," she says. "I just wasn't aware of how serious it was." She never imagined the U.K. Parliament asking for her advice on climate policy. But last year, a letter arrived in her mailbox inviting her to do just that, by joining the United Kingdom's first ever climate assembly. "I'd never done jury service—I'd never done anything like it before," Peachey says. She thought about her stepchildren and stepgrandchildren, and the legacy they would inherit. "That was the main reason why I thought, 'Well, if somebody is going to have input on it, why shouldn't it be me?'"

So, Peachey signed up to spend four weekends driving to Birmingham to listen to experts, deliberate with 109 fellow citizens, and recommend climate policies. The assembly

By **Cathleen O'Grady**

was more than a focus group or a town hall meeting: It was an experiment in handing political power to a random but representative set of citizens. Last month, it produced its final report, and its recommendations will shape debates in Parliament.

The U.K. Climate Assembly is one of a growing number of similar gatherings popping up across Europe, many of them charged with addressing climate change and other science-heavy issues. A citizens' assembly in Ireland that deliberated from 2016 to 2018 led to a referendum that legalized abortion and a government plan to quadruple its carbon tax by 2030. This year in France, an assembly made 149 climate policy recommendations, and President Emmanuel Macron has agreed to push for 146 of them, including making "ecocide" a crime and including climate goals in the French constitution. Spain, Denmark, and Scotland have

announced their own upcoming climate assemblies, although they have been delayed by the coronavirus pandemic. And at the regional and local level, dozens of citizens' juries and councils have drawn up policies on climate adaptation, air quality, and environmental protection.

Advocates say these carefully designed "minipublics" can break political stalemates by bringing together citizens to hear evidence and deliberate. They point to evidence of high-quality policy created by fresh and diverse perspectives, and to behavioral research showing the power of guided deliberation to change minds and reduce polarization. What began as an abstract, academic model has taken off in real-world settings, says David Van Reybrouck, a Belgian historian, author, and advocate of minipublics. "We've really seen a wave, a surge, of deliberative democracy."

Claudia Chwalisz, a policy analyst who has tracked the growing wave of mini-

PHOTO: FABIO DE PAOLA/PA WIRE

publics for the Organisation for Economic Co-operation and Development, a club of economically developed nations, says these deliberative bodies are well-suited to tackling long-term issues because citizens need not worry about the short-term incentives of electoral cycles, giving them more freedom than elected politicians. A minipublic can also embody a wide range of perspectives, an advantage for problems with complex trade-offs and value-based dilemmas. Climate policy, she adds, ticks both boxes.

Letting ordinary people determine climate policy might seem risky, but “if climate scientists have learned anything over the last couple of decades, it’s that they can’t just do the science and expect it to speak for itself,” says John Dryzek, a political scientist at the University of Canberra who wants to use the approach to tackle another fraught science issue. Last month, he and other minipublic advocates published a call for what would be the first global citizens’ assembly. He wants it to take on the issue of genome editing, on the grounds that its widespread ramifications demand international agreements. For this and many other science policy questions, he argues, minipublics are an excellent way to integrate public values with advice from scientists and ethicists. “Scientists don’t have a monopoly on public values,” he says.

RANDOMLY ASSIGNING CITIZENS to positions of political power has a history stretching back to ancient Greece, where the Athenians used the practice to select magistrates and members of their representative Council of Five Hundred. But the architects of electoral systems in postrevolution France and the United States preferred a republican system of professional politicians—an “elected aristocracy”—over outright rule by the masses, Van Reybrouck says. “They were as much afraid of democracy then as we would be of anarchy today.”

Now, however, electoral democracies are floundering in the face of partisanship and populism. In November 2018, Van Reybrouck attended a lunch during a French state visit to Belgium and was pulled into a conversation with Macron. At the time, France was on fire: Thousands of protesters furious at the prospect of a fuel tax hike were expressing their rage in the streets. “What did they want me to do?” Van Reybrouck recalls Macron asking in frustration. “They want me to save the environment and at the same time to keep petrol prices low.”

Bringing citizens into the discussion would help, Van Reybrouck says he told the French president. Not everyone in France has access to public transit, and those already struggling with the costs of

a car would now be further disadvantaged. When Van Reybrouck started to talk about minipublics, he says, “[Macron] puts down his fork and takes a ballpoint and starts taking notes.”

The next month, Macron’s administration walked back the tax hike and announced a series of public town hall meetings for citizens to air grievances. In April 2019, Macron announced the Citizens’ Convention on Climate. Among the assembly’s many eventual proposals were a fuel tax for recreational aviation, an insurance tax

commitments to the 2016 Paris agreement. But since then, there’s been no action to meet the goal, says Darren Jones, a Labour Party politician. “We’ve legislated this target, and that’s great,” he says. “But now how are we going to get there?” Climate policies need the backing of the public, Willis says, which means the public should be involved in creating them. And so, like Macron, Jones and other members of the U.K. Parliament asked the public to help.

Peachey was one of the 30,000 people whose postcode came up in a random lot-



The U.K. Climate Assembly deliberated in Birmingham, after hearing from climate scientists and advocacy groups.

based on vehicle emissions, and a tax on vehicle weight—but no fuel tax for drivers.

Macron’s situation was hardly unique. Governments are hamstringed by the problem of climate action, says Rebecca Willis, an environmental social scientist at Lancaster University who helped choose expert speakers for the U.K. Climate Assembly. She interviewed U.K. politicians and studied their speeches and found that they underestimated public support for climate action. “They didn’t feel under any pressure to act,” she says. Politicians also tend to overestimate the opposition of a vocal minority to some climate measures—such as onshore wind farms, she says. And they fear punishment at the ballot box, where citizens express their opinions about a multitude of policies at the same time: “Voting is such a blunt instrument.”

In mid-2019, the United Kingdom set a target of net zero carbon emissions by 2050, a first step toward keeping its com-

mittee to choose potential participants, and one of 1748 people who responded to the invitation. An algorithm whittled the sample down to 110 people who matched the overall U.K. population’s gender, age, ethnicity, education level, geographic region—and, crucially, their degree of concern about climate change. This two-stage process limits the influence of self-selection, which skews toward people who already have strong interests in a topic and drives polarization, says Stephen Elstub, a political scientist at Newcastle University who is studying how well the U.K. Climate Assembly worked.

The nonprofit Involve, contracted by Parliament to run the assembly, tried to eliminate barriers to entry, says Sarah Allan, Involve’s head of engagement. The team chose Birmingham as a geographically central city and arranged travel, accommodation, child care, and accessibility adaptations like large-print reading materials. They covered all the members’ costs, paying upfront when necessary so

as not to exclude members who couldn't pay out of pocket. And they paid all members an honorarium of £150 for each weekend, a lure to those with little preexisting interest in the subject. "It's really important to try and get the people who wouldn't normally participate," Elstob says.

ON A GLOOMY FRIDAY evening in late January, the assembly members gathered for the first time on the top floor of the glossy, high-rise Park Regis hotel. Famed naturalist David Attenborough made an appearance to greet the members and host a Q&A about his work. In the first sessions, U.K. academics explained climate science basics: the greenhouse effect, the impacts of rising temperatures. Peachey's understanding skyrocketed. "I just thought it was getting hotter, but that'd be nice," she says. "I bet you there's a lot of people out there who think the same thing."

Over three weekends—and a fourth weekend forced online and stretched over three weekends because of the pandemic—the assembly listened not only to scientists, but also to representatives of interest groups such as Greenpeace and industry body Energy UK. The goal was to provide both impartial information and explicitly labeled opinions from advocates, says Chris Stark, chief executive of the Committee on Climate Change, an independent body advising the government. Assembly member Ibrahim Wali, a doctor from Epsom, says that although some members didn't believe the scientists or think climate action was worthwhile, it was clear to everyone that they were not there "to argue about whether climate change is real." The assembly's clear task—identifying policies to reach net zero by 2050—kept discussions on track, he adds.

Sessions shifted between speaker presentations and small group discussions, where members deliberated on policy recommendations, then voted on them by secret ballot. The ballot papers combined predetermined policy suggestions—like a ban on the sale of petrol, diesel, and hybrid cars by 2035—with ideas and amendments suggested by members.

With facilitators making space for everyone to chime in, Wali says no one dominated in the small group discussions. A poll of members found that 94% felt their views were respected, even when others dis-

agreed, and 95% felt they were given "ample opportunity" to express their views. In an age of polarization, a willingness to respectfully hear other views, and the reasons people hold them, changes the hostile dynamic of politics entirely, says Alice Siu, a political scientist at Stanford University: "Something magical happens."

Some researchers contend that entrusting policymaking to ordinary citizens is risky. Cass Sunstein, a legal scholar at Harvard Law School, has argued that deliberating groups can be fertile ground for polarization, pointing to evidence from experiments showing that groups such as churchgoers or jury members can spiral into polarized conclusions.

But a long tradition of research suggests the discussions are often constructive. In 1994, Stanford political scientist James Fishkin polled 869 people—randomly

reform in Northern Ireland. A deliberative poll held in Dallas last year saw Republicans and Democrats moving toward a middle ground on divisive issues such as refugee resettlement and a minimum wage. Across the board, the researchers found participants could back up their opinions with clear reasoning. Siu, analyzing transcripts from these experiments, has found that men and women contributed equally to discussions, contrary to expectations that men would dominate.

Such good outcomes aren't guaranteed, says Sander van der Linden, a University of Cambridge social psychologist who studies decision-making. In some cases participants do become more polarized, rather than less, in the face of disagreement; that happened in group discussions of gay rights in Poland—not one of Fishkin's experiments—that had recruited partici-

pants from pools with strong opinions. And in groups with a majority and minority opinion, "polarization cascades" can override the minority, causing them to shift their opinion for reasons that have nothing to do with evidence and good arguments.

But with high-quality information, facilitators to keep discussions on track, and rules to enforce civility, it is possible to steer people away from group biases, van der Linden says. In 2017, Kim Strandberg, a behavioral scientist at Åbo Akademi University in Finland, and his colleagues found

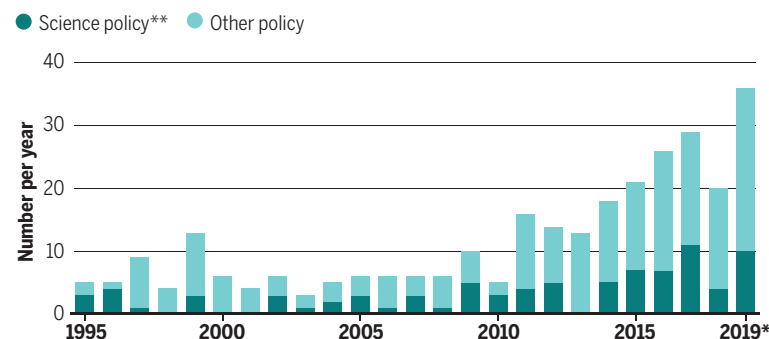
such measures prevented polarization in discussions about the status of the Swedish language, spoken by a minority in Finland. "The people who are polarized and tribal are actually a minority," Dryzek says. A good assembly may include some members of this minority, he says. "But most of the people in a minipublic have no history of activism or involvement with an issue, and so they're in a good position to reflect on what they hear."

What makes deliberation so powerful and positive seeming is still a bit of a black box, says André Bächtiger, a political scientist at the University of Stuttgart. Recent experiments like Strandberg's seem to demonstrate the effects of deliberation and good design. But a deliberative experience is made up of countless moving parts of human interaction, and "we just don't have everything under control," Bächtiger says. "We're still in the infancy of all of this."

And the growing popularity of mini-

Deliberative surge

The number of national and local minipublics—randomly selected assemblies that deliberate over policy recommendations—has risen in the past decade. Many dealt with scientific questions.



*The Organisation for Economic Co-operation and Development tallied minipublics through November 2019;

Science assembled data for the rest of 2019.

**Dealt with questions on climate, environment, health, biotechnology, neuroscience, research priorities, or food technology.

selected from the U.K. electoral register—on their attitudes toward rising levels of crime, asking questions such as whether more people should be sent to prison and whether sentences should be tougher. Next, 301 of them agreed to meet for 2 days in Manchester, where they heard from politicians, lawyers, and police. They asked questions and discussed policy proposals in small groups before being polled again. The results showed participants had changed their opinions, in some cases dramatically: For each question in the poll, at least one-third of the deliberators had shifted their stance to some degree, and on some questions, as many as two-thirds had changed their minds.

Fishkin and his collaborators around the world went on to conduct more than 100 of these experiments, which they call deliberative polls. They have explored opinions on flood management in Uganda, constitutional reform in Mongolia, and educational



An Extinction Rebellion climate protest in London in September. Social scientists have found that politicians underestimate public support for climate action.

publics carries another risk: Some will be constituted as cheap imitations, without the crucial design elements. Fishkin trademarked the term “deliberative poll” partly to maintain quality control, and Chwalisz and her colleagues have published guidance on best practices. But citizens’ panels, juries, and conferences that rely on self-selected participants still spring up. “Lots of people are calling a lot of different things citizens’ assemblies now, because it’s kind of a trendy method,” says Graham Smith, a political scientist at the University of Westminster.

THE U.K. CLIMATE ASSEMBLY’S final report, published on 10 September, says climate action should be based on cross-party leadership and principles of fairness to different groups of people. Its policy recommendations, supported by a majority of assembly members, include restoring public ownership for public transit—largely privatized in the United Kingdom—introducing air travel taxes that increase as people fly more often, and adding labeling to food that indicates its carbon footprint.

Will those recommendations be adopted in the end? Whereas the French assembly was set up by Macron himself, the U.K. assembly was run at arm’s length from Prime Minister Boris Johnson’s Conservative government, which means its impact will be

less direct, Smith says. The parliamentary committees that established it will use the report to introduce debates, inform their own recommendations, and scrutinize legislation coming from Johnson and his government, but “we’ve yet to see how the government will respond,” Smith says.

Even if a minipublic’s recommendations are discussed by politicians, it’s difficult to say how much they influence later policies, says Émilien Paulis, a political scientist at the French-speaking Free University of Brussels. But in an analysis of data on the policy impact of 55 minipublics, Chwalisz and her colleagues found that, 75% of the time, public authorities implemented more than half the citizens’ suggestions. Only six minipublics in the sample saw none of their recommendations implemented.

Some advocates have suggested imbuing minipublics with official legislative powers, but that would be antidemocratic because they were never given consent to govern, says Cristina Lafont, a political philosopher at Northwestern University who describes herself as a “frenemy” of the deliberative bodies. Keeping minipublics in an advisory role, rather than enabling them to produce binding recommendations, is more truly democratic, she says. Although she wants the public to engage in thoughtful deliberation, she worries that minipublic participants cannot be held to account by the

wider public. And being too enthralled by the results of “lottocratic” minipublics runs the risk of encouraging a lazy shortcut to good governance, she says: “Why not let the minipublics do the thinking and the deciding for us?”

Dryzek agrees that democracies should not blindly defer to the decisions of minipublics. But they still offer a critical piece of information for policymakers that experts can’t provide, he argues: a meaningful gauge of public values.

For Peachey, the climate assembly made the right course clear. Her new awareness of climate change and its impacts has galvanized her on behalf of her stepgrandchildren. “We’re not going to benefit,” she says. “But future generations will.” Since the U.K. assembly’s final meeting, Peachey has joined her local parish council and worked with fellow councilors to declare a climate emergency—a symbolic gesture that has led to local climate minipublics elsewhere in the country. She daydreams about bringing a wind farm to an old military airfield near Bath.

Her own habits have changed, too. She now shops at her local farmer’s market and plans to get solar panels for her home. Her new electric car is a great conversation starter, a chance for her to unpack what she learned at the assembly, she says. “I’ve been given this information. It seems a shame not to use it.” ■

INSIGHTS

PERSPECTIVES

EVOLUTIONARY GENETICS

Of dogs and men

Ancient genomes reveal the common history of human and dog

By Pavlos Pavlidis¹ and Mehmet Somel²

But Argos passed into the darkness of death, now that he had fulfilled his destiny of faith and seen his master once more after twenty years.” This quote from *The Odyssey* vividly illustrates the bond between dog and human. These two species, separated by 90 million years of evolutionary history, have spent much of their recent past in each other’s service. This interaction has remodeled both species’ environments and has modified the phenotypic and genetic composition of dog populations. On page 557 of this issue, Bergström *et al.* (1) use 27 ancient dog genomes from across Eurasia, going back 11,000 years, to resolve whether dog domestication happened once or mul-

tiplle times, whether dog dispersals and adaptations were coupled to those of humans, and how dogs interacted with their wild sisters, the wolves.

Dogs likely evolved from a wolf population that self-domesticated, scavenging for leftovers from Paleolithic hunter-gatherers in Eurasia (2, 3). However, the exact timing and geographic location where the dog lineage started remain unknown, owing to the scarcity of Paleolithic dogs in the archaeological record. Analyses of genetic data suggest that dog-wolf divergence took place ~25,000 to 40,000 years ago (4, 5), providing an earliest possible date for dog domestication.

Consistent with previous analyses (2–5), Bergström and colleagues support a scenario where dogs were domesticated 20,000 years ago, around the Last Glacial Maximum

(LGM). If so, the domestication of dogs predated other Neolithic domestications—such as sheep, pig, and cattle—and may have even facilitated them. Interestingly, many of these later domestications happened independently in multiple local wild populations. For example, there is evidence that pigs were domesticated in both Anatolia and China (6). For dogs, however, the story is different. Dogs and modern-day Eurasian gray wolves appear as monophyletic groups; that is, any dog is genetically closer to another dog than to a wolf, and vice versa (7). Monophyly supports a single origin of dogs from a possibly extinct wolf lineage.

¹Foundation for Research and Technology Hellas, Institute of Computer Science, 70013 Heraklion, Greece. ²Middle East Technical University, 06800 Ankara, Turkey. Email: pavlidisp@gmail.com; somel.mehmet@gmail.com

A hunter and dog are depicted in rock art that dated to between 5000 and 2000 BCE. It is located in Wadi Tashwinet, Tadrart Acacus, Libyan Sahara, Libya.

Although this remains speculative, Bergström and colleagues show that by 11,000 years ago dog lineages had already diversified and spread worldwide. But how did they spread? Even though human movement may have contributed to this expansion, LGM human migrations were probably not extensive enough to explain these patterns. Instead, dogs could have been exchanged between forager groups, or may have spread autonomously in a semi-feral state.

Bergström *et al.* show that during the next 10,000 years, diversified dog lineages interbred frequently over wide geographical areas. Further, they directly compare quantitative measures of population history of humans and of dogs. They show that the genetic relations between human populations largely match the genetic relations between proximal dog populations in Eurasia and the Americas, suggesting that movement patterns are correlated between dog and human. For instance, about half of the ancestry of European dogs originates from Paleolithic West Eurasia, and the other half from Southwest Asia; similarly, modern-day Europeans are a mixture between pre-Neolithic hunter-gatherers and Neolithic farmers from Anatolia. However, as shown by Bergström and colleagues, dogs have not always faithfully followed humans, resulting in cases of decoupling between dog history and human history. For instance, comparing Neolithic and Chalcolithic Iran, they find that people have remained, but indigenous dogs have been replaced by Levantine dogs. Conversely, in Neolithic Germany and Ireland, incoming farmers of Anatolian descent appear to have adopted dogs from local foragers.

In addition to sharing dispersal paths, dogs and humans have traced parallel paths of evolutionary adaptation. Variation in the copy number of genes encoding amylase, the enzyme required for breaking down starch, is such an example of convergent evolution. Humans carry extra salivary amylase copies compared to chimpanzees (8, 9), owing to high starch consumption that perhaps began before farming (10). Likewise, most dogs, compared to wolves, carry extra pancreatic amylase (*AMY2B*) copies, possibly facilitating starch digestion in their new environment (11). Bergström and colleagues show that early dogs already carried extra amylase copies compared to wolves, but amylase copy numbers further expanded following the increasing reliance on starch-rich agricultural diets in prehistoric Eurasia over the past 7000 years. Similarly, a recent

study on Arctic sled dogs reported genetic signatures of adaptation in their fatty acid metabolism genes (12), analogous to their Inuit masters who carry adaptive changes in the same metabolic pathways—a likely response to the high-fat Arctic diet (13).

After their split, dogs and wolves have continued to occasionally interbreed. For instance, it was shown that a black coat color allele passed from dogs to wolves in North America (14). Bergström *et al.* also confirm regional wolf-dog admixture. They show that Iberian wolves are genetically closer to European dogs than to Asian dogs, whereas Mongolian wolves are closer to Asian dogs. But did gene flow occur from wolf to dog, from dog to wolf, or both ways? Bergström *et al.* reason that in the case of wolf-to-dog gene flow, all wolves should be more similar to those wolf-like dogs than to non-admixed dogs. However, they do not find such a pattern. Specifically, a wolf from Xinjiang, China, was identified as equally distant from all dogs, past and present, suggesting mainly unidirectional gene flow from dog to wolf. This intriguing finding could be linked with wolf or dog behaviors (e.g., going feral) or asymmetry in population sizes. There might also be selection against such hybrids, perhaps as a result of their unbiddable or suboptimal behavior (12). Studying and dating the distribution of introgression signatures across wolf genomes would be interesting. If population size asymmetry is the reason, then we might expect admixture to intensify after agriculture began. Studying signatures of sex bias in dog-to-wolf gene flow could also provide insight into the behavioral background of the process. We further anticipate that successful genetic analyses of early dog-like fossils from Eurasia may help to resolve the long-standing debate surrounding the origins of dog domestication. ■

REFERENCES AND NOTES

1. A. Bergström *et al.*, *Science* **370**, 557 (2020).
2. G. D. Wang *et al.*, *Nat. Commun.* **4**, 1860 (2013).
3. L. M. Shannon *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13639 (2015).
4. P. Skoglund, E. Ersmark, E. Palkopoulou, L. Dalén, *Curr. Biol.* **25**, 1515 (2015).
5. L. R. Botigué *et al.*, *Nat. Commun.* **8**, 16082 (2017).
6. M. Price, H. Hongo, *J. Archaeol. Res.* **28**, 557 (2020).
7. A. H. Freedman *et al.*, *PLoS Genet.* **10**, e1004016 (2014).
8. G. H. Perry *et al.*, *Nat. Genet.* **39**, 1256 (2007).
9. P. Pajic *et al.*, *eLife* **8**, e44628 (2019).
10. S. Mathieson, I. Mathieson, *Mol. Biol. Evol.* **35**, 2957 (2018).
11. E. Axelsson *et al.*, *Nature* **495**, 360 (2013).
12. M. S. Sinding *et al.*, *Science* **368**, 1495 (2020).
13. M. Fumagalli *et al.*, *Science* **349**, 1343 (2015).
14. T. M. Anderson *et al.*, *Science* **323**, 1339 (2009).

ACKNOWLEDGMENTS

We thank E. Tsagioti and the METU CompEvo group for helpful suggestions.

10.1126/science.abe7823

NOISE POLLUTION

The quiet spring of 2020

Anthropogenic noise has no legacy effects on bird song

By Wouter Halfwerk

Sounds of the past can be easily forgotten, especially when soundscapes change gradually over long periods of time. This past spring, many people got a chance to experience how the outside world sounded in the 1950s. Global transport came to a halt as human activities decreased abruptly (either voluntarily or under direct order for lockdown) to stem the spread of coronavirus disease 2019 (COVID-19). With fewer people driving cars and hardly any airplanes traversing the skies, the amount of background noise across whole continents dropped substantially. On page 575 of this issue, Derryberry *et al.* (1) report the impact of the COVID-19 shutdown on animal behavior—namely, the songs of white-crowned sparrows. The findings suggest that mitigation measures against noise pollution could yield immediate beneficial effects on urban wildlife.

Derryberry *et al.* observed white-crowned sparrows in the San Francisco Bay Area. The males of this species have a beautiful, crystal-clear song, that starts with a long whistle, followed by a series of fast, down-swept notes. Males use their song to keep territorial intruders at bay (2, 3). In urban areas, with high amounts of anthropogenic background noise, males normally sing louder songs compared to their rural counterparts (4). Although increasing song amplitude (making songs louder) is an effective communication strategy to overcome the masking impact of urban noise (5), it is often traded against other song components such as reduced song complexity. In the case of Bay Area white-crowned sparrows, increasing song amplitude comes at the cost of reduced trill performance (4).

Derryberry *et al.* recorded white-crowned sparrow males during the shutdown in March and April of 2020 in San Francisco as well as in nearby rural areas. The authors measured background noise data from their recordings and used toll data since the opening of the Golden Gate Bridge in 1937 to extrapolate traffic data. During the lockdown period,



Urban white-crowned sparrows improved song performance during the coronavirus disease 2019 (COVID-19) shutdown.

traffic was reduced to amounts resembling those of the 1950s, which was accompanied by a large reduction in noise amounts in the urban areas (on average 7 dB), but not rural areas. Comparing songs from the 2016 breeding season with songs recorded in 2020, the authors found that urban males sang with higher trill performance (rapid and complex motor control) during the pandemic shutdown. By modeling detection distances, the authors further showed that urban birds communicated over twice the distance during lockdown when compared to preshutdown periods, despite a notable reduction of 4 dB in song amplitude. Rural birds did not experience changes in background noise and consequently did not alter their songs. The magnitude of change in effective communication was beyond expectation, suggesting that noise pollution is even more detrimental than has been so far assumed.

Previous work has shown that many songbirds immediately change their songs when they are experimentally exposed to short periods of traffic noise by either increasing their amplitude or pitch (6, 7) or changing the length and shape of individual song elements (8). Derryberry *et al.* reveal that the opposite is also true: When noise pollution is removed, birds sing immediately as if they have never been exposed to any traffic noise.

Such a fast return to preexposure conditions suggests that anthropogenic noise may not suffer from the same legacy effects as other environmental pollutants, particularly chemical compounds. Sensory pollution includes anthropogenic stimuli, such as vibrational or airborne noise caused by traffic and industry, or artificial light at night, which hampers animals in acquiring environmental information (9, 10). These stimuli generally leave no traces when the source of pollution is taken away, and we would therefore expect their disruptive impact to stop immediately, without legacy effects on animal behavior.

Sensory pollutants may, however, have other longer-lasting impacts through direct effects on the animal's development (11), through behaviors involved in important life-history decisions (9), or through changes in community composition (12). Animals exposed to high amounts of sensory pollutants during critical stages in their lives may develop specific morphological, physiological, or cognitive traits that remain fixed long after these pollutants have been reduced. Noise exposure over multiple days, for example, can influence learning and cognition in birds (13). Likewise, animals that made a poor decision by breeding in a place considered suboptimal because of sensory pollution may remain there for the rest of their lives, irrespective of changes in exposure amounts (9). Sensory pollution also may lead to long-lasting community-level changes, such as changes in vegetation as the result of the presence or

absence of seed predators and seed dispersers near noisy gas-compressor stations (12). It would therefore be very interesting to track the fates of urban animals that were born during the quiet spring of 2020 and assess whether they develop and behave similarly to their rural counterparts (but different from their urban predecessors).

The findings of Derryberry *et al.* suggest that temporary noise reduction immediately influences animal behavior, providing policy-makers with a possible solution to the harmful effects of this sensory pollution on animal populations, namely reducing exposure to pollutants during critical stages of an animal's life cycle (e.g., during migration or dispersal). Mitigation measures to reduce the impact of artificial light at night or anthropogenic noise should therefore be tailored to taxon-specific needs. Shutting down the 9/11 memorial lights in New York City for a few minutes during peak periods, for example, may greatly reduce fatalities among nocturnal migrating birds (9). Likewise, brief road closures during specific times of day or breeding season may prevent the development of chronic stress in endangered species, with long-lasting impacts.

Spring 2020 not only reminded us how the natural world of 1950s felt (14), smelt (15), and sounded like, but it also highlighted the impact of anthropogenic activities on our daily lives and the lives of organisms with whom we share this planet. It taught us that we can change the way we live in a rapid and drastic manner and that it is not too late to do so because species may still recover from anthropogenic stresses. It is an important lesson that we may need to remember again soon to cope with serious pressing challenges of the 21st century—namely mass extinction and climate change. ■

REFERENCES AND NOTES

1. E. P. Derryberry *et al.*, *Science* **370**, 575 (2020).
2. D. A. Luther, J. Phillips, E. P. Derryberry, *Behav. Ecol.* **27**, 332 (2016).
3. J. Podos, *Evolution* **51**, 537 (1997).
4. D. A. Luther, E. P. Derryberry, *Anim. Behav.* **83**, 1059 (2012).
5. H. Brumm, H. Slabbekoorn, in vol. 35 of *Advances in the Study of Behavior*, P. J. B. Slater *et al.*, Eds. (Academic Press, 2005), pp. 151–209.
6. W. Halfwerk, H. Slabbekoorn, *Anim. Behav.* **78**, 1301 (2009).
7. H. Brumm, D. Todt, *Anim. Behav.* **63**, 891 (2002).
8. M. J. Montague, M. Danek-Gontard, H. P. Kunc, *Behav. Ecol.* **24**, 343 (2013).
9. D. M. Dominoni *et al.*, *Nat. Ecol. Evol.* **4**, 502 (2020).
10. W. Halfwerk, H. Slabbekoorn, *Biol. Lett.* **11**, e20141051 (2015).
11. S. L. Nedelec *et al.*, *Sci. Rep.* **4**, 5891 (2014).
12. C. D. Francis, N. J. Kleist, C. P. Ortega, A. Cruz, *Proc. R. Soc. London Ser. B* **279**, 2727 (2012).
13. W. Halfwerk, K. van Oers, *Proc. R. Soc. London Ser. B* **287**, 20192951 (2020).
14. T. Lecocq *et al.*, *Science* **369**, 1338 (2020).
15. M. B. Karuppasamy *et al.*, *Air Qual. Atmos. Health* **10**, 1007/s11869-020-00892-w (2020).

10.1126/science.abe8026

PROTON MEMBRANES

Speeding protons with metal vacancies

Membranes made from nanosheets derived from CdPS_3 have high proton conductivity

By Fengmei Wang¹ and Jun He^{1,2}

Proton exchange membranes (PEMs) find applications not only in fuel cells and sensors but as chemical filters and in biological systems. In particular, the Nafion (Chemours Company) PEM, which consists of sulfonated tetrafluoroethylene-based fluoropolymer-copolymer, is widely used in practical electrochemical processes. It has a high proton conductivity up to 0.2 S/cm below 80°C and high relative humidity (RH) (>93%) (1), but its conductivity drops severely at higher temperatures or at low RH (below 50%). On page 596 of this issue, Qian *et al.* (2) report a new class of proton membranes assembled from two-dimensional (2D) layered transition-metal phosphorus trichalcogenide (TMPTC) nanosheets in which metal vacancies boost ion conductivity. These membranes exhibit a proton conductivity of ~0.95 S/cm at 90°C and 98% RH but still have a conductivity of 0.26 S/cm even at 60% RH.

Various materials under different conditions have been explored as PEMs, including organic polymers other than Nafion, inorganic-organic nanohybrids (3), metal organic frameworks (MOFs) (4), bioderived materials (5), and 2D layered materials such as graphene (6) and graphdiyne (7). However, these alternatives to Nafion have much lower proton conductivity (10^{-5} to 10^{-2} S/cm). To understand why the TMPTC-based assembled PEM fabricated by Qian *et al.* has much higher performance, it is necessary to look in general at the structures and mechanisms underlying proton conductivity in PEMs.

A high-performing PEM usually has a large density of optimally sized channels for proton trans-

port, with high selectivity. Delicate pores must be created that can be hydrated (3), and hydrophilic groups (6) are purposely introduced to attain improved functionality. For organic polymers, $-\text{SO}_3\text{H}$ groups in Nafion are much more hydrophilic than the fluorocarbon-based backbone, and these groups line the inner walls of nanochannels for proton conduction (1). The oxygenated functional groups (hydroxyl, carboxyl, and epoxide) of graphene oxide (GO) also play the same role to establish hydrogen-bonded channels in water-molecule networks for proton transport (6).

The TMPTCs explored by Qian *et al.* are layered van der Waals materials that

contain various metals M (Cd, Mn, Fe, Co, Ni, or Zn) and chalcogenides X (S or Se). For example, each layer of MPS_3 consists of infinite arrays of $\text{P}_2\text{S}_6^{4-}$ units coordinated to M^{2+} cations by S atoms (8). The fairly ionic M–S bonds, especially in MnPS_3 and CdPS_3 , enable an intercalation process that removes M^{2+} cations to create vacancies under mild aqueous conditions (9). A heterogeneous equilibrium that occurs between M^{2+} ions and solvated $\text{P}_2\text{S}_6^{4-}$, in which M^{2+} ions are then easily displaced by other alkali ions (Li^+ , Na^+ , and K^+) or protons (H^+) present in the aqueous solution, maintains charge balance (see the figure, top). After this intercalation and exchange

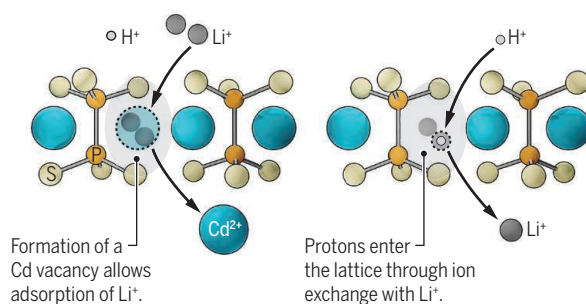
process, a complete chemical transformation of the entire structure is ensured and not limited by ion diffusion within the host lattice. Both $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ and $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ have different structural features, except for the valence state of each element, compared with the pristine CdPS_3 . Either Cd vacancies (or Mn vacancies in Mn analogs) in the matrix serve as a donor center for mobile H^+ or of lithium ions to boost ion conductivity.

In PEMs, two mechanisms for proton transport (10), the vehicle mechanism and Grotthuss mechanism, can operate. In the former, which operates in Nafion films, water molecules come together to form a cluster (or nanochannel) that act as a vehicle for transporting protons. Thus, ion conductivity starts to dramatically decrease in Nafion PEMs at high temperatures or at low RH as the water clusters become disconnected.

However, conditions of low RH and elevated temperatures have no obvious effect on the Grotthuss mechanism, in which the protons hop along the hydrogen-bonded water-molecular network, such as those in GO membranes and the 2D layered $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ stacked nanosheets Qian *et al.* assembled into membranes through vacuum filtration. In their system, the uniform Cd^{2+} vacancies negatively charge the surfaces of each layer to provide uniform donor centers

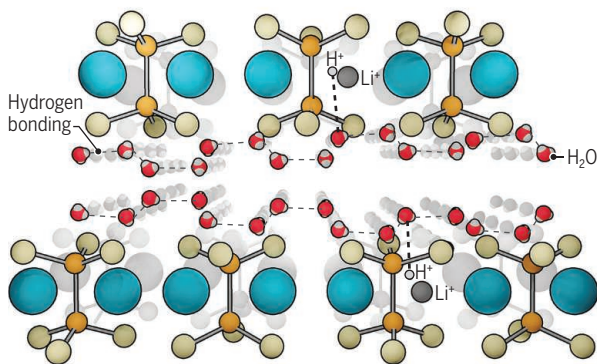
Ionic defects and water networks

The high proton conductivity of the $\text{Cd}_{1-x}\text{PS}_3\text{Li}_x\text{H}_x$ membranes fabricated by Qian *et al.* results from sites created by cadmium vacancies and the formation of water networks.



Ion exchange in layered materials

A heterogeneous equilibrium can occur between the solid CdPS_3 structure and Li^+ ions and protons (H^+) in aqueous solution to form $\text{Cd}_{1-x}\text{PS}_3\text{Li}_x\text{H}_x$.



Proton transport in the layered membrane

Ordered water networks between the monolayers, created through hydrogen bonding, along with protons in the Cd vacancies create proton transport channels.

¹Chinese Academy of Sciences (CAS) Center for Excellence in Nanoscience, CAS Key Laboratory of Nanosystem and Hierarchical Fabrication, National Center for Nanoscience and Technology, Beijing 100190, P. R. China.

²Key Laboratory of Artificial Micro- and Nano-structures of Ministry of Education, School of Physics and Technology, Wuhan University, Wuhan 430072, P. R. China. E-mail: wangfm@nanoctr.cn; hej@nanoctr.cn

to adsorb H^+ and form monolayers that are highly hydrophilic on both sides (see the figure, bottom). The ordered water molecular network between the monolayers provides a proton transport channel. The adsorbed H^+ in Cd vacancies join this network and accelerate proton conduction through the Grotthuss mechanism under humidified conditions to form a one-dimensional (1D) hydrogen-bonded chain (see the figure, bottom). A large capillary-like force develops in the 2D layered material assembled membrane that can also guide and accelerate ion transport (6, 11). Even at low RH (53%), the water clusters can expand and join each other to form a monolayer water network within the $Cd_{0.85}PSLi_{0.15}H_{0.15}$ framework.

As an emerging family of 2D layered materials, the MPX_3 series has promising features in coordination chemistry and solid-

“The ordered water molecular network between the monolayers provides a proton transport channel.”

state chemistry. Besides the univalent ion conductivity in an aqueous system induced by tailoring chemical coordination environment, the divalent ionic (Zn^{2+}) conductivity in the solid-state MPX_3 (like $ZnPS_3$) is also worth studying (12). The polarizable lattice provided by the sulfide anions and the layered structure offers pathways for ion conduction in the solid-state structure. This discovery of extremely high proton conductivity in a membrane based on 2D MPX_3 represents a step toward practical PEM in energy conversion and storage. ■

REFERENCES AND NOTES

1. S. Bose *et al.*, *Prog. Polym. Sci.* **36**, 813 (2011).
2. X. Qian *et al.*, *Science* **370**, 596 (2020).
3. S. Moghaddam *et al.*, *Nat. Nanotechnol.* **5**, 230 (2010).
4. M. Liu *et al.*, *Nat. Commun.* **7**, 12750 (2016).
5. C. Ma *et al.*, *Sci. Adv.* **6**, eabc0810 (2020).
6. M. R. Karim *et al.*, *J. Am. Chem. Soc.* **135**, 8097 (2013).
7. J. Xu *et al.*, *Nat. Commun.* **10**, 3971 (2019).
8. F. Wang *et al.*, *Adv. Funct. Mater.* **28**, 1802151 (2018).
9. R. Clement, O. Garnier, J. Jegoudez, *Inorg. Chem.* **25**, 1404 (1986).
10. J. Wei, in *Polymer-Based Multifunctional Nanocomposites and Their Applications*, K. Song, C. Liu, J. Guo, Eds. (Elsevier, 2019), pp. 245–260.
11. R. R. Nair, H. A. Wu, P. N. Jayaram, I. V. Grigorieva, A. K. Geim, *Science* **335**, 442 (2012).
12. A. J. Martinolich *et al.*, *Chem. Mater.* **31**, 3652 (2019).

ACKNOWLEDGMENTS

F. W. and J. H. thank the National Natural Science Foundation of China (nos. 21805057 and 61625401) and the National Key R&D Program of China (nos. 2018YFA0703700 and 2016YFA0200700).

10.1126/science.abe6166

ECOLOGY

Should I stay or should I go?

Pinpointing genetic alleles that influence the timing of salmon migration might aid restoration efforts

By Garrett McKinney

A common assumption in genetics is that complex traits are influenced by many genes, each with small effect. However, a growing number of examples reveal that single genes or gene regions can have a strong influence on phenotypic traits (1–3). The genetic architecture of a trait has implications for how traits are inherited and how trait variation is viewed in the context of conservation (maintenance and recovery). On page 609 of this issue, Thompson *et al.* (4) examine the genetic basis of migration timing in Chinook salmon and find that timing is controlled by a single genomic region containing two genes, *GREBIL* (GREB1-like retinoic acid receptor coactivator) and *ROCK1* (Rho-associated coiled-coil containing protein kinase 1). Genetic variants in this region include single-nucleotide polymorphisms (SNPs) and short duplications that together form distinct haplogroups with linked alleles (inherited together from one parent) that effectively function as small supergenes.

Chinook salmon complete their ocean migration and return to streams to spawn as either early migrants (usually called spring- or summer-run Chinook) or late migrants (usually called fall- or late fall-run Chinook). Despite this difference in return timing, individuals of each type spawn at relatively similar times, with spring-migrating salmon remaining in streams for several months before spawning and fall-migrating salmon spawning soon after returning to streams. The difference in migration timing can be associated with a number of traits, including body size, fat content, difference in spawn timing, and, in many cases, use of a different habitat. Spring-migrating salmon are highly prized by anglers because of their high fat content. These fish also fill a vital ecological function by serving as a food source for predators when other salmon species are not in streams and by transporting marine-derived nutrients into an up-river habitat that is inaccessible to fall-migrating salmon (5).

Unfortunately, spring-migrating Chinook

are threatened or even extirpated in many regions. An understanding of the genetic basis of migration timing can illuminate conservation actions that are most appropriate for protecting and restoring threatened populations. Previous research identified a small genomic region associated with migration timing; however, these studies examined subsets of the genome (6, 7), excluded individuals with intermediate run timing (8), or used pooled-genome resequencing, which cannot reconstruct individual haplotypes (9). The new study answers multiple lingering questions, including the genetic variants associated with differences in run timing and the degree to which there is gene flow (mating) among different run types.

Although declining populations have been observed for both spring- and fall-run Chinook, spring-run Chinook have exhibited particularly steep declines, resulting at least in part from habitat degradation. Spring-run salmon return to freshwater months before spawning and rely on cold-water pools for survival during the summer. This makes them particularly vulnerable to anthropogenic changes that alter summer water flows or increase water temperature, such as logging, mining, water diversion, and dam building (6). Before the discovery that the *GREBIL-ROCK1* genomic region influences run timing, researchers noted that spring- and fall-run populations within coastal drainages were often closely related genetically. This led to the conclusion that run timing had evolved independently many times and the corollary that a specific run-time trait that is lost could reevolve (10). However, the results of Thompson *et al.* and other recent studies on the *GREBIL-ROCK1* region (4, 6, 8) suggest that loss of spring-run alleles in a population equates to loss of the spring-run trait and that reevolution of the trait through mutation is exceedingly unlikely (11).

This conclusion raises an obvious question: Can fall-run Chinook act as a reservoir for spring-run alleles (4, 6), thereby allowing individuals with the spring-run phenotype to reemerge from fall-run populations? Thus far, two studies have examined this question; both found that spring-run alleles are either very rare or absent in fall-run populations where the spring-run has been extirpated. This suggests that recovery

National Research Council Research Associateship Program, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, WA 98112, USA. Email: garrett.mckinney@noaa.gov



Chinook salmon return to spawn as early migrants in spring or late migrants in fall. This important life-history variation is under the control of a single genetic region.

of extirpated spring-run populations will require either natural immigration of spring-run individuals or human intervention to reintroduce fish with the spring-run alleles, likely through translocation of individuals. A frequent concern with salmon restoration efforts is that translocated individuals might not have the appropriate adaptations to thrive under local conditions. Thompson *et al.* showed evidence of natural historic interbreeding between run types in the Klamath River that is sufficient to homogenize the genome outside of the region associated with run-timing differentiation. This finding suggests that in these populations, locally adapted alleles for other traits still exist in the fall-run population. Rebuilding of spring-run Chinook populations will most likely be achieved through restoration or improvement of upriver habitats traditionally used by these salmon.

The study by Thompson *et al.* provides a deeper view into the genetic basis of migration in Chinook salmon. However, additional avenues must be explored to fully decipher the mechanisms underlying migration timing. The new study focused only on Chinook salmon from California and Oregon. Also, although previous studies found SNPs or haplotype blocks within the *GREBIL-ROCK1* region that are associated with migration throughout the species range, these studies either examined individual SNPs or used pooled rather than individual-level whole-genome sequences. Thus, it is unclear if the specific haplotypes identified in the new study exist throughout the species range or if populations from Washington to Alaska harbor additional ge-

netic variants or distinct haplotypes related to migration.

Differential migration timing has also been documented in other salmon species, and the *GREBIL-ROCK1* genomic region has been found to influence migration timing in steelhead trout (8, 12). Although Micheletti *et al.* (12) used whole-genome sequencing, these results have not been directly compared with studies of Chinook salmon. Thus, it is not yet known if the same genetic variants within this region influence run timing in both species or if run timing in each species is the result of parallel evolution using the same molecular pathway. Comparative analyses of these two species as well as other salmon that exhibit run-timing variation offer appealing opportunities to study how similar traits can evolve across populations and species. ■

REFERENCES AND NOTES

1. N. J. Barson *et al.*, *Nature* **528**, 405 (2015).
2. M. R. Jones *et al.*, *Science* **360**, 1355 (2018).
3. S. E. Johnston *et al.*, *Mol. Ecol.* **20**, 2555 (2011).
4. N. F. Thompson *et al.*, *Science* **370**, XXX (2020).
5. T. P. Quinn, P. McGinnity, T. E. Reed, *Can. J. Fish. Aquat. Sci.* **73**, 1015 (2015).
6. T. Q. Thompson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 177 (2019).
7. I. J. Koch, S. R. Narum, *Evol. Appl.* **13**, 2316 (2020).
8. D. J. Prince *et al.*, *Sci. Adv.* **3**, e1603198 (2017).
9. S. R. Narum, A. Di Genova, S. J. Micheletti, A. Maass, *Proc. Biol. Sci.* **285**, 20180935 (2018).
10. R. S. Waples, D. J. Teel, J. M. Myers, A. R. Marshall, *Evolution* **58**, 386 (2004).
11. M. Ford *et al.*, "Reviewing and synthesizing the state of the science regarding associations between adult run timing and specific genotypes in Chinook salmon and steelhead" (NOAA Processed Report NMFS-NWFS-PR-2020-06, U.S. Department of Commerce, 2020).
12. S. J. Micheletti, J. E. Hess, J. S. Zandt, S. R. Narum, *BMC Evol. Biol.* **18**, 140 (2018).

10.1126/science.abe9169

VIEWPOINT: COVID-19

Will SARS-CoV-2 become endemic?

Reinfection, seasonality, and viral competition will shape endemic transmission patterns

By Jeffrey Shaman and Marta Galanti

Reinfection, in which an individual is subject to multiple, distinct infections from the same virus species throughout their lifetime, is a salient feature of many respiratory viruses. Indeed, the persistence and ubiquity in human society of common respiratory viruses—including influenza viruses, respiratory syncytial virus (RSV), rhinovirus, and the endemic coronaviruses—are largely due to their ability to produce repeat infection. Since the emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus responsible for the ongoing coronavirus disease 2019 (COVID-19) pandemic, a critical concern has been whether humans will experience reinfections with this pathogen, which might enable it to become endemic.

Typically, following an initial infection, the human adaptive immune system develops a suite of defenses, including memory B lymphocytes capable of producing neutralizing antibodies targeted to bind to that particular pathogen, and memory T lymphocytes that help regulate immune responses and induce death of infected cells. These adaptive immune components, particularly B cells, can produce sterilizing immunity in which the pathogen, if reintroduced to the host, is prevented from replicating within the body.

However, for many viruses, a number of processes, particularly insufficient adaptive immune response, waning immunity, and immune escape, can undermine or circumvent the sterilizing character of immunity and allow subsequent reinfection. In the first instance, an initial infection with a particular agent may not engender an adaptive immune response sufficient to confer sterilizing immunity. Serological studies indicate that most SARS-CoV-2 infections, regardless of severity, induce de-

Factors influencing postpandemic transmission of SARS-CoV-2

Rates of repeat infection, factors modulating seasonality, competition with other circulating respiratory viruses, and control measures will influence the endemic pattern of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) transmission.

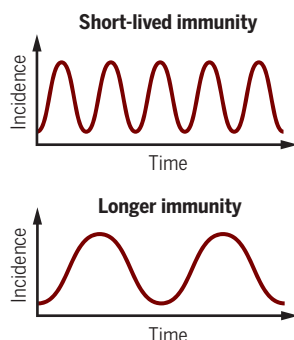
Immunity

Adaptive immune response following infection

Severity of infection

Waning of immunity

Immune escape

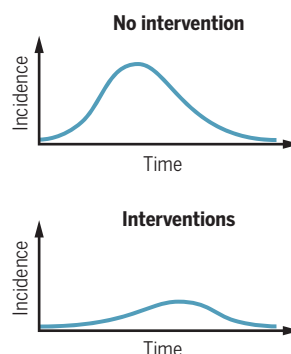


Interventions

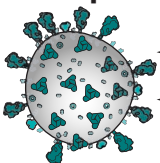
Vaccination (timing efficiency, coverage)

Drug treatments

Nonpharmaceutical interventions (personal protective equipment, social distancing, travel restrictions, hygiene measures)



Endemic pattern

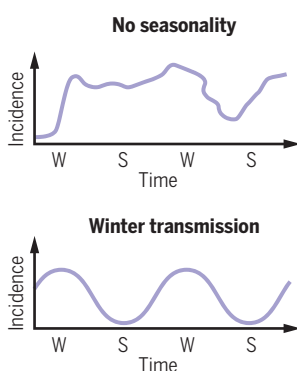


Seasonality

Impact of environmental conditions on transmissibility

Seasonal changes in host behavior, school calendar, and indoor/outdoor gatherings

Seasonal changes in immune function

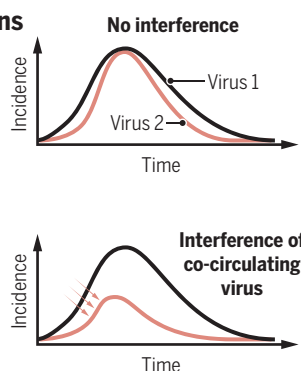


Virus interactions

Transmissibility of each virus

Order of exposure

Degree of cross protection



development of some specific antibodies (1); however, despite encouraging results from the experimental vaccination of primates, it remains unclear whether those antibodies are sufficient to provide long-term effective protection or if other adaptive immune components are present and functional. Furthermore, immune response to SARS-CoV-2 infection is heterogeneous, with individuals who experience asymptomatic infections manifesting a weaker immune response than those experiencing more severe disease (1). It is possible that some individuals never develop sterilizing immunity following infection with SARS-CoV-2, or that multiple exposures will be needed for affinity maturation and development of long-lasting protection.

Waning immunity, in which the initial adaptive immune response is robust and protective but dissipates over time, leav-

ing the host vulnerable to reinfection, may also undermine sterilizing immunity. Immune escape is a third process that can facilitate reinfection, in particular by viruses. Here, a virus, during its continued serial passage through a host population, accumulates point mutations. This accumulation, termed antigenic drift, may lead to conformational changes of viral surface proteins that disrupt the binding of antibodies previously generated against an earlier variant. Immune escape is a consequence of this antigenic drift that enables reinfection through the evasion of adaptive protection.

The time scales of waning immunity and immune escape differ by pathogen and have yet to be defined for SARS-CoV-2. Thus far, the mutation rate of the SARS-CoV-2 genome appears to be slower than that of influenza viruses. This lower rate may be a consequence of proof-reading during replication, which is exclusive to coronaviruses among RNA viruses. Conversely, human coronavirus (HCoV) OC43 is highly variable, particularly

in genes encoding surface proteins such as the spike protein, indicating that considerable diversification can occur. To date, some evidence of SARS-CoV-2-specific antibody waning has been captured in a longitudinal study (2), and a few verified repeat SARS-CoV-2 infections have been documented (3). Although reinfections can occur, the number of reinfection cases is not currently sufficient to generalize the duration of immunity at population scales or the severity of repeat infection. Whether reinfections will be commonplace, how often they will occur, how contagious reinfected individuals will be, and whether the risk of severe clinical outcomes changes with subsequent infection remain to be understood.

Insight from other respiratory viruses points to the possibility of reinfection with SARS-CoV-2. Naturally acquired infections with the four endemic HCoVs (OC43, HKU1, 229E, and NL63) indicate that reinfections with the same HCoV type are common within 1 year (4); sequential infections with the same influenza virus strain can occur in less than 2 years (5); and reinfections of adults with RSV within 1 year have also been documented (6).

By contrast, more pathogenic viruses that induce systemic effects on the host may elicit a longer-lasting adaptive immune response. For example, longitudinal immune profiles from SARS survivors showed a stronger immune response with neutralizing antibodies persisting for 2 to 5 years (7). However, it could not be confirmed if and for how long this response conferred immunity because the SARS outbreak lasted less than 1 year.

In addition to duration of protective immunity, the long-term effects of SARS-CoV-2 on humans will depend on the severity of reinfection. Sequential infections with influenza virus have been associated with less severe symptoms (8), whereas no association between reinfection and symptom severity was found in recurring endemic HCoV infections (4). In addition, for other viruses (e.g., RSV and dengue), suboptimal binding of naturally induced or vaccine-induced antibodies can enhance infection severity upon subsequent exposure, a phenomenon called antibody-dependent enhancement (ADE) (9). To date,

Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY 10032, USA. Email: jls106@cumc.columbia.edu

responses among the few patients with verified SARS-CoV-2 reinfection have been heterogeneous with one apparent repeat infection requiring hospitalization. Thus, thorough serological and prospective studies are needed to determine whether ADE manifests among SARS-CoV-2 infections, either because of prior homologous infection or cross-reactive antibodies from other HCoVs. This will have particular relevance for vaccines and convalescent plasma therapy.

Should reinfection prove commonplace, and barring a highly effective vaccine delivered to most of the world's population, SARS-CoV-2 will likely become endemic (10). The typical time scale at which individuals experience reinfection and seasonal differences in transmissibility will determine the pattern of endemicity. Outside the tropics, the incidence of many common respiratory virus infections increases during particular times of the year. This phase-locked behavior is due to accumulated susceptibility to reinfection, which increases over time because of immune escape and waning immunity, and seasonal modulation of virus transmissibility derived from environmental conditions, changing behavior (e.g., mixing indoors in cold weather), or altered immune function. For example, influenza incidence is greatest during winter in temperate regions. Once expelled from an infectious host, the influenza virus appears to be more stable in low-humidity conditions (11), which are prevalent both indoors and outdoors during winter. Further, during colder months, people spend more time indoors and school is in session, which may facilitate transmission, and shorter days and less sunlight exposure may suppress immune function.

The endemic HCoVs (OC43, HKU1, NL63, and 229E) all exhibit a seasonality in temperate regions, similar to influenza viruses (12). Consequently, numerous studies have sought to determine whether conditions such as temperature, sunlight, humidity, ozone, and pollution affect SARS-CoV-2 viability and transmissibility. The results are not currently conclusive, although it appears that environmental conditions, such as sunlight and humidity, may modulate SARS-CoV-2 transmissibility—not enough to preclude transmission during the first waves of the pandemic when immunity is generally low—but perhaps sufficient to favor seasonal, phase-locked transmission during winter in temperate regions, similar to influenza virus, once immunity increases.

Like the 2009 influenza pandemic, the continued circulation by SARS-CoV-2 following this initial pandemic period will

manifest as a function of reinfection rates, vaccine availability and efficacy, and social, immune, and innate factors that modulate virus transmissibility (see the figure). In addition, the cyclic persistence of SARS-CoV-2 in human populations may be affected by ongoing opportunities for interaction with other respiratory pathogens.

Co-circulating respiratory viruses may interfere with one another while competing for the same resources, and their interactions have been studied at population and individual levels, in reconstructed human tissues and in animal models. The outcomes in individuals experiencing serial exposure to different viruses vary and in general appear to depend on the order and timing of exposures. Many studies have documented evidence of negative interference between viruses caused by short-lived (days) protection elicited from the first infection. Host antiviral interferon responses are often regarded as the main mechanism by which interference manifests; that is, as a result of a recent infection, the host cells up-regulate the synthesis of interferons, potentially inhibiting a secondary infection. Even though it is short-lived, this effect can be strong at population scales and temporarily reduce the prevalence of a virus or shift the timing of its circulation. For example, it is hypothesized that a large summer 2009 rhinovirus outbreak delayed pandemic influenza virus emergence in Europe (13).

The clinical and population-scale interactions of SARS-CoV-2 with other respiratory viruses, particularly influenza viruses and other HCoVs, need to be monitored in the coming years. To date, some SARS-CoV-2 coinfections have been documented (14), including coinfections with influenza and RSV; however, testing for multiple pathogens has not been routinely carried out, and the scarce data that do exist, mostly for older adults with high rates of preexisting medical conditions, do not support a definitive evaluation of coinfection likelihood or severity. Studies prior to the pandemic indicate that simultaneous infections with multiple respiratory viruses are not uncommon but are not associated with increased disease severity.

At the population scale, a possible overlap between influenza and SARS-CoV-2 outbreaks poses a serious threat to public health systems. Seasonal influenza produces millions of severe infections worldwide every year, and this additional burden could be catastrophic on systems already challenged by the COVID-19 pandemic. Conversely, given similar modes of transmission among different respiratory viruses, the nonpharmaceutical interventions adopted to mitigate SARS-CoV-2 transmis-

sion (personal protective equipment, social distancing, increased hygiene, limited indoor gatherings) may reduce the magnitude of seasonal influenza outbreaks. Such increased use of nonpharmaceutical measures, and possible virus interference, could be responsible for the reduced incidence of influenza during the recent winter of the Southern Hemisphere (15).

The phases and magnitudes of different outbreaks in a multipathogen system are dictated by the interaction dynamics between those pathogens: from large overlapping phases when pathogens enhance one another's transmission, to complete inhibition of a strain by the neutralizing cross-reactivity of a more transmissible one (9). Several postpandemic scenarios for SARS-CoV-2 have been modeled (10), postulated on duration of immunity and cross-immunity between SARS-CoV-2 and the other betacoronaviruses (OC43 and HKU1). A duration of immunity similar to that of the other betacoronaviruses (~40 weeks) could lead to yearly outbreaks of SARS-CoV-2, whereas a longer immunity profile, coupled with a small degree of protective cross-immunity from other betacoronaviruses, could lead to the apparent elimination of the virus followed by resurgence after a few years. Other scenarios are, of course, possible, because there are many processes at play and much that remains unresolved. ■

REFERENCES AND NOTES

1. Q.-X. Long *et al.*, *Nat. Med.* **26**, 1200 (2020).
2. J. Seow *et al.*, medRxiv 10.1101/2020.07.09.20148429 (2020).
3. K. K.-W. To *et al.*, *Clin. Infect. Dis.* 10.1093/cid/ciaa1275 (2020).
4. M. Galanti, J. Shaman, *J. Infect. Dis.* 10.1093/infdis/jiaa392 (2020).
5. M. J. Memoli *et al.*, *Clin. Infect. Dis.* **70**, 748 (2020).
6. C. B. Hall, E. E. Walsh, C. E. Long, K. C. Schnabel, *J. Infect. Dis.* **163**, 693 (1991).
7. H. Mo *et al.*, *Respirology* **11**, 49 (2006).
8. J. R. Davies, E. A. Grilli, A. J. Smith, *J. Hyg. (Lond.)* **92**, 125 (1984).
9. N. Ferguson, R. Anderson, S. Gupta, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 790 (1999).
10. S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, *Science* **368**, 860 (2020).
11. J. Shaman, M. Kohn, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3243 (2009).
12. A. S. Monto *et al.*, *J. Infect. Dis.* **222**, 9 (2020).
13. A. Wu *et al.*, *Lancet Microbe* 10.1016/S2666-5247(20)30114-2 (2020).
14. D. Kim, J. Quinn, B. Pinsky, N. H. Shah, I. Brown, *JAMA* **323**, 2085 (2020).
15. WHO, www.who.int/influenza/surveillance_monitoring/updates/2020_08_31_update_GIP_surveillance/en/ (2020).

ACKNOWLEDGMENTS

J.S. and Columbia University partially own SK Analytics. J.S. consults for BNI. This work was supported by U.S. National Science Foundation grant DMS-2027369 and a gift from the Morris-Singer Foundation.

Published online 14 October 2020

10.1126/science.abe5960

LANDMARK: ANIMAL INTELLIGENCE

Charles H. Turner, pioneer in animal cognition

An African American scientist's early discoveries are forgotten for all the wrong reasons

By **Hiruni Samadi Galpayage Dona**
and **Lars Chittka**

In the late 19th and early 20th centuries, Charles Henry Turner (1867–1923) established a research program that was in sharp contrast to prevailing ideas regarding animal behavior and cognition. Despite facing almost insurmountable barriers because of his African American ethnicity, he published more than 70 papers, including several in *Science* (1–3), on comparative brain anatomy in birds and invertebrates, individual variation of behavior and learning competences, and intelligent problem-solving in a large variety of animals, at a time when the dominant ideas only credited animals with the simplest of learning abilities. But his discoveries and conceptual advances failed to gain the recognition they deserved, and his works were later all but forgotten—indeed, some recent animal cognition research has reinvented wheels that had already been fashioned by Turner.

Charles Darwin (1809–1882) and George Romanes (1848–1894) were famously generous in attributing intelligent behavior and mental abilities to animals, but their musings were largely based on observation and inference. The predominant experimentalist theories of animal behavior in the early 20th century, however, largely rejected notions of advanced animal intelligence or insight. Early ethologists such as Oskar Heinroth, Charles Whitman, and Wallace Craig focused instead on innate behavior and imprinting, a simple form of learning. Where problem-solving was observed, such as when animals open puzzle boxes, behaviorists such as Edward Thorndike proposed that this materialized as a result of trial and error; not insight or understanding of the nature of the challenge. None of these scientists were interested in individual variation of behavior. Enter Charles H. Turner, who took seriously Darwin's assertion of the importance of individual variation as well as the idea that humans were not the only intelligent animal species. But Turner backed up this possibility with a rigorous experimental approach.

Turner was born just 2 years after the end of slavery in the United States in 1865.

Department of Psychology, School of Biological and Chemical Sciences, Queen Mary University of London, London, UK. Email: h.s.galpayagedona@qmul.ac.uk; l.chittka@qmul.ac.uk

He obtained his M.Sc. from the University of Cincinnati in 1892 (4). In the same year, the 25-year-old published two single-author papers in *Science*—one of which was a short version of his B.Sc. work on the comparative anatomy of bird brains, whose relative size and complexity he compared with those of reptiles (1). Turner's verdict was, "When we compare the brain of a crow or a titmouse with the brain of a snake or a turtle, it is no longer a marvel that birds bear towards their reptilian cousins the relation of intellectual giants to intellectual dwarfs" [(1), p. 16]. The same year also saw the publication of another remarkable study on variations in web building by gallery spiders (5) that contained key ingredients of Turner's distinct interpretation of animal behavior that was to accompany his entire body of work. Like many of his future papers, the study interfaces careful field observations with meticulously controlled laboratory work. Contrary to the still-popular view that spider web construction is a prime example of invertebrates' robotic, repetitive action patterns, Turner reported variation between individuals in adapting their construction to the geometry of available space and the functionality in capturing prey: "we may safely conclude that an instinctive impulse prompts gallery spiders to weave gallery webs, but the details of the construction are the products of intelligent action" [(5), p. 110]. In the year of his Ph.D. (1907), Turner published on associative and spatial learning in ants and reported individual learning curves of their performance (6). Turner's focus on individual differences in behavior is a constant theme in his studies. It is deplorable that the now-popular field of "animal personality" has taken so little notice of Turner's trailblazing approach.

The list of Turner's discoveries and insights that should have garnered attention, but did not, is long. Every student of animal behavior knows Nikolaas Tinbergen's study from 1932 on spatial learning, in which the later Nobel laureate (awarded for studies of individual and social behavior in animals) first marked a beewolf's nest entrance with pine cones, then moved them to demonstrate that the insect was guided by a memory of the landmarks (7). But it is mostly unknown that Turner had already published similar findings in 1908, observing a solitary burrowing bee whose nest entrance was close to a discarded Coca Cola bottle cap. When the

cap was moved to a nearby location next to an artificial burrow that Turner had made, the bee crawled into that burrow without hesitation—indicating, just as in Tinbergen's experiments, that the insect had a memory for landmarks rather than, for example, being guided by an instinct to follow the scent of the nest (8).

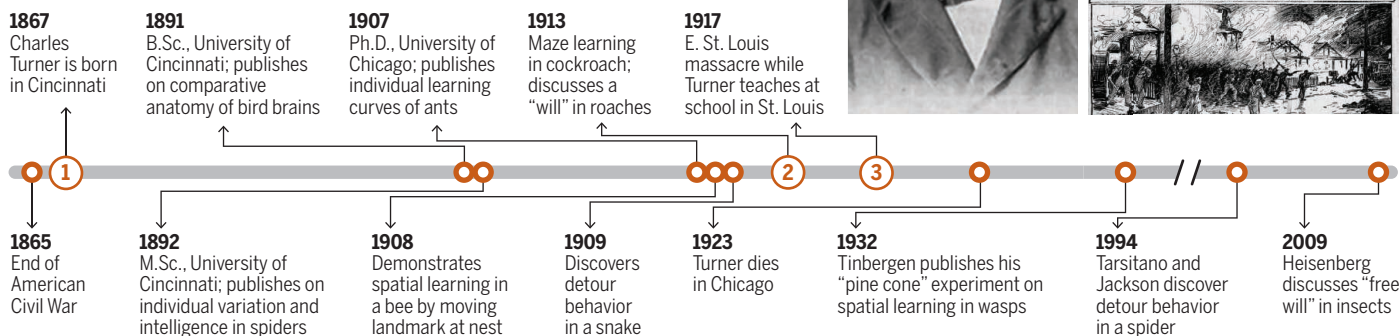
In 1912, in a study that explored how a prey-carrying walking wasp finds its way home around obstacles in the path, Turner explicitly confronted Thorndike, affirming that the wasp's behavior is not explicable by trial-and-error learning and is instead consistent with a form of intentionality and an awareness of the desired outcome of the wasp's actions (9). Moreover, Turner found that an ant stuck on a small island began assembling a bridge to the "mainland," using three different materials (10). The ant's behavior could not easily be explained by then-popular notions of instinct or trial-and-error learning; instead, the ant appeared to appreciate the nature of the problem, imagined a solution, and then worked toward this goal.

The view that animals are capable of insightful problem-solving was also apparent in Turner's interpretation of his field observations of the hunting behavior of a snake pursuing a lizard (3). The lizard had escaped up a tree and looked downward where it expected the snake to launch the next attack. The snake, which had been pursuing the lizard for some time, instead ascended another tree, crossed over when it had reached a point higher than the lizard, and then attacked from behind. These observations are reminiscent of the detour behavior seen when jumping spiders hunt—discovered in the 1990s (11). It is remarkable that Turner's views on animal intentionality preceded present-day explorations of the same topic by a century. Even though his experimental work was known to contemporary giants such as John Watson and Thorndike (4) and across the Atlantic by later Nobel laureate Karl von Frisch, Turner's visionary ideas about animal intelligence did not resonate in the field; perhaps they were simply too far ahead of the time. Accordingly, they are almost completely unrecognized in the current literature.

Further highlighting the importance and insightful nature of Turner's work, in 1913 he reported on the effects of age and sex on cockroaches trained to navigate mazes (12). Turner found that individuals placed an em-

From Charles H. Turner to comparative cognition: 1850–2020

Charles H. Turner made important observations about animal cognition, which went against the leading paradigms of the time. His ideas have stood the test of history, but Turner's work has largely been forgotten, likely because his ethnicity prevented him from becoming a research team leader and so he could not train scientists who might have continued his approach. Turner was active in the U.S. civil rights movement and advocated that education is key to overcoming ethnic barriers in society.



phasis on either speed or accuracy: Older cockroaches choose slowly but more precisely. Extraordinarily, Turner suggests that the hesitation that cockroaches display when evaluating their options bears the hallmarks of will, a facet of consciousness. The question of whether humans and other animals exhibit free will continues to generate controversy among neuroscientists and philosophers. That insights from insect behavior could contribute to this debate has only recently been suggested again by neuroscientist Martin Heisenberg (13), who proposed that insects display an awareness of the consequences of their actions and evidence of free will in deciding between options.

Why is Turner not more widely credited as a major luminary in research on animal intelligence? Turner faced substantial obstacles because of his ethnicity. Despite publishing many important papers, he was not given a post at a major U.S. research university. Turner's work was thus conducted without access to state-of-the-art laboratory facilities or library resources. One reason for Turner's relative obscurity today may be that he had no possibility of mentoring research students who would have carried his ideas into subsequent generations. For comparison, Russian Nobel laureate Ivan Pavlov (1849–1936), famed for his studies on classical conditioning, trained more than 140 co-workers. One cannot help but wonder what Turner might have achieved if he had had comparable resources and manpower. The entire field of animal cognition may have developed differently. Would a "cognitive revolution" have been needed against the dominant ideas of behaviorism that ruled psychology for the first half of the 20th century (postulating

that learning largely happens in the form of simple associations), if Turner's ideas about advanced cognition in animals had generated a movement at the time he expressed them?

African American historian William Du Bois (1868–1963) lamented that "C. H. Turner, one of the great world authorities on insects, nearly entered the faculty of Chicago University; but the head professor who called him died, and his successor would not have a "N----," despite a reputation which was European; Turner died in a high school of neglect and overwork" [(4), p. 348]. The institution at which Turner taught from 1908 to 1922 was Sumner High School, a school for African American children in St. Louis. During his time there, he and his pupils would have witnessed the East St. Louis massacre in 1917, during which white mobs murdered more than 100 African Americans; another 6000 lost their homes as a result of arson attacks on their neighborhoods (14). Turner was active in the U.S. civil rights movement, and years before coming to St. Louis, he wrote that an emphasis on high-quality education and a conscious effort to abandon prejudices might eliminate barriers between Blacks and whites within a few decades (15).

One would hope that nowadays, a person of Turner's caliber might not face similar adversity in terms of academic employment opportunities or long-term recognition of their contribution to science. But even today, very few scholars in animal cognition, or indeed across biology, are Black. Turner clearly recognized the importance of ethnic-minority role models from the earliest stages of education; their near-complete absence in a field of scholarly study will require concerted counterefforts. Funded summer schools for ethnic

minority students can also make a substantial difference to inspire budding scientists. Institutions must make still-stronger efforts to eliminate biases in hiring, promotions, and salary decisions and to celebrate the successes of ethnic minority scholars. Even where they do (and there is likely plenty of room for improvement), overt or poorly concealed racism is still commonly experienced by underrepresented ethnic groups, even in academia. This will likely discourage many aspiring scientists from venturing further. A hopeful development is that some conference organizers are taking steps in the right direction to increase inclusivity; for example, the Animal Behavior Society annually supplies the Charles H. Turner award that prioritizes traditionally underrepresented groups for conference travel funding. More than ever, humanity needs to be inclusive to confront current and future challenges. Diversity increases the pool of talent and, as Turner's example shows, has the potential to transform entire fields. ■

REFERENCES AND NOTES

1. C. H. Turner, *Science* **19**, 16 (1892).
2. C. H. Turner, *Science* **20**, 39 (1892).
3. C. H. Turner, *Science* **30**, 563 (1909).
4. C. I. Abramson, *Annu. Rev. Entomol.* **54**, 343 (2009).
5. C. H. Turner, *J. Comp. Neurol.* **2**, 95 (1892).
6. C. H. Turner, *J. Comp. Neurol. Psychol.* **17**, 367 (1907).
7. N. Tinbergen, *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **16**, 305 (1932).
8. C. H. Turner, *Biol. Bull.* **15**, 247 (1908).
9. C. H. Turner, *Psyche* **19**, 100 (1912).
10. C. H. Turner, *Biol. Bull.* **13**, 333 (1907).
11. M. S. Tarsitano, R. R. Jackson, *Behaviour* **131**, 65 (1994).
12. C. H. Turner, *Biol. Bull.* **25**, 348 (1913).
13. M. Heisenberg, *Nature* **459**, 164 (2009).
14. J. N. Harrington, *Buzzing with Questions: The Inquisitive Mind of Charles Henry Turner* (Boyd's Mills Press, 2019).
15. C. H. Turner, in *Twentieth Century Negro Literature* (J. L. Nichols, 1902), pp. 162–166.

10.1126/science.abd8754

RETROSPECTIVE

Paolo Sassone-Corsi (1956–2020)

Groundbreaking molecular biologist and epigeneticist

By **Eric Verdin**

Paolo Sassone-Corsi, eminent molecular biologist, died unexpectedly at the age of 64 on 22 July in Laguna Beach, California. Paolo made seminal contributions in the fields of transcriptional regulation, epigenetics, circadian biology, and metabolic regulation, and he pioneered the discovery of key links between these disciplines.

Born in Naples, Italy, in 1956, Paolo devoted himself to science and discovery from an early age. At just 12 years old, he founded an amateur astronomy club with his brother Emilio in Naples. (Ten years later, the two brothers would publish a paper on the bands of Saturn.) Although he retained his passion for astronomy, Paolo decided to focus his graduate studies on yeast biology; he received his doctorate in biological sciences in 1979 from the University of Naples Federico II. The university would honor him as a most meritorious alumnus in 2016.

During a postdoc under the mentorship of molecular biologist Pierre Chambon at the Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC) in Strasbourg, France, Paolo was drawn to the basic mechanisms of transcription. He contributed key work on promoters, enhancers, and transactivating factors as well as the identification and characterization of a promoter sequence of the adenovirus major late gene that became known as the TATA box, a hallmark regulatory element of most eukaryotic genes. He then moved to the Salk Institute for Biological Studies in San Diego, California, where he studied the intracellular pathway leading to the activation of the proto-oncogenes *c-fos* and *c-jun* and demonstrated their mechanism of action in the regulation of transcription; he showed that *c-fos* and *c-jun* interact with each other, and he identified the regulation of *c-fos* by cyclic adenosine monophosphate (cAMP).

In 1989, Paolo established his first independent laboratory at the IGBMC, where he rapidly progressed to the level of director of research at the Centre National de la Recherche Scientifique. During this

period, he identified a previously unknown messenger in the cAMP signaling pathway—the cAMP-responsive element modulator (CREM)—and characterized its role in spermiogenesis and germ cell apoptosis. The identification of a distinctive circadian fluctuation of CREM expression in the pineal gland led Paolo and his team into chronobiology, a fitting reconnection with astronomy, as this field of biology explores the consequences of Earth's rotation and the 24-hour cycle of light and darkness that it imposes on all living organisms.

In 2006, the University of California, Irvine (UCI), recruited Paolo and his wife, prominent neurobiologist Emiliana Borrelli. Paolo was hired as a distinguished



professor and led the Department of Pharmacology for 5 years. He then founded the university's Center for Epigenetics and Metabolism, an interdisciplinary center reflecting his increasing interest in metabolism and circadian biology in relation to epigenetics. Paolo's singular understanding of transcriptional regulation led to a series of additional groundbreaking discoveries, including the dissection of transcriptional and epigenetic mechanisms that regulate circadian rhythms and an explanation of how they respond to metabolic states. He was first to demonstrate that the key oscillating transcription factor CLOCK is a histone acetyltransferase, and he characterized how the deacetylase SIRT1 and fluctuations in concentrations of nicotinamide adenine dinucleotide and its reduced form participate in the control of circadian rhythms. Through these pio-

neering studies, Paolo became a leader in the fields of circadian rhythms, epigenetics, and metabolism.

Having worked in similar fields, I came to know Paolo at scientific meetings and through service on advisory boards, visits to UCI, and collaborative efforts. I will remember him for his beautiful lectures, his charm, his warmth, the twinkle in his eye when he would gently tease you, his passion for science, and his love and pride for his students and postdocs. More than anything, many of us will remember his zest for life, food, wine, and science and his incredibly generous hospitality.

Paolo and Emiliana excelled at organizing memorable meetings. Their biannual Epigenetic Control and Cellular Plasticity International Symposium was known not only for the caliber of its science but also for Paolo and Emiliana's warmth and hospitality. They had moved to UCI in hopes of living somewhere climatically and geographically similar to their native Italy. In the process, they brought a big slice of Italy and its dolce vita to California for everyone to enjoy. Paolo made all of us, his guests, feel truly special. His pride and joy in wanting to share with me a beautiful sunset at the Pelican Inn in Newport Beach remains to this day a special moment with a cherished colleague and friend.

Paolo's creativity, originality, and scientific courage were recognized by his election as a fellow of the American Association for the Advancement of Science (AAAS, the publisher of *Science*) in 2014, as well as numerous awards. He received the Grand Prix Liliane Bettencourt in 1997 and the Grand Prix Charles-Léopold Mayer de l'Académie des Sciences in 2003. He also served on various scientific advisory boards and journal editorial boards.

The scientific community has lost in Paolo a trailblazing scientist, a superb and generous colleague and mentor, and a witty, adventurous, and warm friend. Paolo is survived by Emiliana and his brothers Lucio and Emilio Sassone-Corsi. For all of us who counted Paolo as a close friend, his early departure is hard to accept. Our loss is balanced only by the gratitude we feel for having had the chance to cross paths with and draw inspiration from this brilliant, generous colleague. ■

10.1126/science.abe9149

Buck Institute for Research on Aging, Novato, CA 94945, USA. Email: everdin@buckinstitute.org

Political sectarianism in America

A poisonous cocktail of othering, aversion, and moralization poses a threat to democracy

By Eli J. Finkel¹, Christopher A. Bail², Mina Cikara³, Peter H. Ditto⁴, Shanto Iyengar⁵, Samara Klar⁶, Lilliana Mason⁷, Mary C. McGrath¹, Brendan Nyhan⁸, David G. Rand⁹, Linda J. Skitka¹⁰, Joshua A. Tucker¹¹, Jay J. Van Bavel¹², Cynthia S. Wang¹, James N. Druckman¹

Political polarization, a concern in many countries, is especially acrimonious in the United States (see the first box). For decades, scholars have studied polarization as an ideological matter—how strongly Democrats and Republicans diverge vis-à-vis political ideals and policy goals. Such competition among groups in the marketplace of ideas is a hallmark of a healthy democracy. But more recently, researchers have identified a second type of polarization, one focusing less on triumphs of ideas than on dominating the abhorrent supporters of the opposing party (1). This literature has produced a proliferation of insights and constructs but few interdisciplinary efforts to integrate them. We offer such an integration, pinpointing the superordinate construct of political sectarianism and identifying its three core ingredients: othering, aversion, and moralization. We then consider the causes of political sectarianism and its consequences for U.S. society—especially the threat it poses to democracy. Finally, we propose interventions for minimizing its most corrosive aspects.

ASCENDANCE OF POLITICAL HATRED

Democrats and Republicans—the 85% of U.S. citizens who do not identify as pure independents—have grown more contemptuous of opposing partisans for decades, and at similar rates [see supplementary materials (SM)]. Only recently, however, has this aversion exceeded their affection for copartisans. On a “feeling thermometer” scale ranging from cold (0°) to neutral (50°) to warm (100°), affect toward copartisans has consistently hovered in the 70° to 75° range. By contrast, affect toward opposing partisans has plummeted from 48° in the 1970s to 20° today (see the figure, top panel). And cold feelings toward the out-party now exceed warm feelings to-

ward the in-party (see the figure, bottom panel). Out-party hate has also become more powerful than in-party love as a predictor of voting behavior (2), and by some metrics, it exceeds long-standing antipathies around race and religion (SM).

This aversion to opposing partisans might make strategic sense if partisan identity served as a strong proxy for political ideas. But given that sectarianism is not driven primarily by such ideas (SM), holding opposing partisans in contempt on the basis of their identity alone precludes innovative cross-party solutions and mutually beneficial compromises. This preclusion is unfortunate, as common ground remains plentiful. Indeed, despite the clear evidence that partisans have grown increasingly disdainful of one another, the evidence that they have polarized in terms of policy preferences is equivocal (3).

Along the way, the causal connection between policy preferences and party loyalty has become warped, with partisans adjusting their policy preferences to align with their party identity (SM). For example, a recent experiment demonstrated that Republicans exhibit a liberal attitude shift after exposure to a clip of President Donald Trump voicing a liberal policy position (SM); there is little evidence to suggest that Democrats are immune to analogous shifts in response to their own political leaders. Overall, the severity of political conflict has grown increasingly divorced from the magnitude of policy disagreement (4).

POLITICAL SECTARIANISM

In the past decade, political scientists have introduced various constructs to capture this nonideological type of polarization, including “affective polarization” (1) and “social polarization” (4). Scholars from psychology and other disciplines have introduced constructs, such as “tribalism” (SM), to flesh out its social-psychological properties.

We propose here a superordinate construct, political sectarianism—the tendency to adopt a moralized identification with one political group and against another. Whereas the foundational metaphor for tribalism is kinship, the foundational metaphor for political sectarianism is religion, which evokes analogies focusing less on genetic relatedness than on strong faith in the moral correctness and superiority of one’s sect.

Political identity is secondary to religion in traditional forms of sectarianism, but it is primary in political sectarianism. In the United States today, even though Democrats and Republicans differ on average in terms of religious affiliation, their schism is fundamentally political rather than religious. It is, in this sense, quite distinct from the Sunni-versus-Shia sectarian schisms that characterize politics in some Muslim-majority nations.

Political sectarianism consists of three core ingredients: othering—the tendency to view opposing partisans as essentially different or alien to oneself; aversion—the tendency to dislike and distrust opposing partisans; and moralization—the tendency to view opposing partisans as iniquitous. It is the confluence of these ingredients that makes sectarianism so corrosive in the political sphere. Viewing opposing partisans as different, or even as dislikable or immoral, may not be problematic in isolation. But when all three converge, political losses can feel like existential threats that must be averted—whatever the cost.

WHY SECTARIANISM IS SURGING

Rising political sectarianism in the United States is multiply determined; here we con-

On American exceptionalism

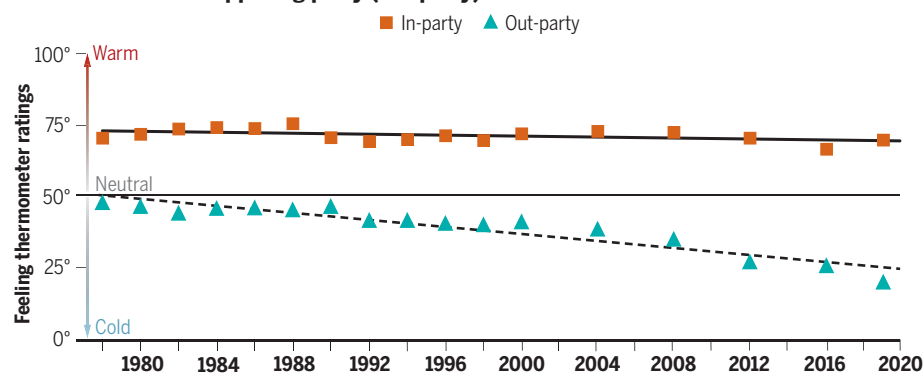
A recent study offers valuable international perspective on political polarization, leveraging data from 1975 through 2017 in nine Western democracies to examine feelings toward copartisans and opposing partisans. The study controls statistically for the number of parties and offers a valuable, albeit noncomprehensive, comparison set (13).

Four nations—America, Canada, New Zealand, and Switzerland—exhibit increasing sectarianism over time, with the rate steepest in America. By contrast, Australia, Britain, Norway, Sweden, and Germany exhibit decreasing sectarianism over time. The most notable findings pertain to out-party hate [increasingly “frigid” evaluations of opposing partisans, via a “feeling thermometer” (see main text)]. Across the eight other nations, the mean rate of change in out-party hate was 0.004° per year (range: −0.2° to +0.2°) on the 0°-to-100° scale. In the United States, the rate of change was −0.6° per year. By 2017, out-party hate was stronger in America than in any other nation.

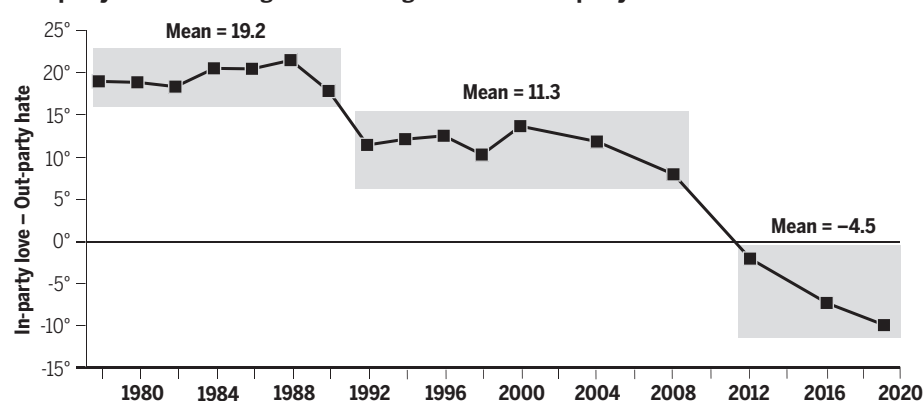
The rise of out-party hate

With the exception of 2020, all data come from the American National Election Study (ANES), as reported in (1). To calculate the estimates for the lower panel, we used upper-panel estimates to compute, relative to the neutral point on the feeling thermometer, the strength of in-party love (in-party score – 50) and out-party hate (50 – out-party score), and then took the difference of those two scores. See supplementary materials for details.

Warmth toward the opposing party (out-party) has diminished for decades



Out-party hate has emerged as a stronger force than in-party love



sider three crucial causes. First, in recent decades, the nation's major political parties have sorted in terms of ideological identity and demography. Whereas self-identified liberals and conservatives used to be distributed broadly between the two parties, today the former are overwhelmingly Democrats and the latter are overwhelmingly Republicans (SM). The parties also have sorted along racial, religious, educational, and geographic lines. Although far from absolute, such alignment of ideological identities and demography transforms political orientation into a mega-identity that renders opposing partisans different from, even incomprehensible to, one another (4). This mega-identity can grow so powerful that it changes other identities, as when partisans alter their self-identified religion, class, or sexual orientation to align with their political identity (SM).

As distinct as Democrats and Republicans actually are today, partisans nevertheless vastly overestimate such differences. They view opposing partisans as more socially distant, ideologically extreme, po-

litically engaged, contemptuous, and uncooperative than is actually the case (5) (SM), thereby exacerbating political sectarianism. For example, Republicans estimate that 32% of Democrats are LGBT when in reality it is 6%; Democrats estimate that 38% of Republicans earn over \$250,000 per year when in reality it is 2% (6).

Second, as Americans have grown more receptive to consuming information slanted through a partisan lens, the media ecosystem has inflamed political sectarianism. The decline of the broadcast news era, during which impartiality was prized, began in the 1980s, driven in part by the Reagan administration's termination of the Federal Communications Commission (FCC) "fairness doctrine" in 1987. This doctrine, introduced in 1949, required that broadcasters discuss controversial topics in a manner that the FCC assesses as unbiased. Among the first media figures to leverage the demise of the fairness doctrine was Rush Limbaugh, whose influential conservative radio program went into national syndication in 1988 (SM). The ethos of impartiality that CNN espoused when intro-

ducing cable news faltered with the launch of the conservative Fox News in 1996 and the liberal pivot of MSNBC a decade later. People who are already sectarian selectively seek out congenial news, but consuming such content also amplifies their sectarianism (SM).

In recent years, social media companies like Facebook and Twitter have played an influential role in political discourse, intensifying political sectarianism. Scholars from sociology, political science, economics, psychology, and computational social science debate whether such web platforms create polarizing echo chambers (7) (SM). However, a recent field experiment offers intriguing evidence that Americans who deactivate their Facebook account become less politically polarized (8). In addition, emotional and moralized posts—those containing words like "hate," "shame," or "greed"—are especially likely to be retweeted within rather than between partisan networks (9). Social-media technology employs popularity-based algorithms that tailor content to maximize user engagement, increasing sectarianism within homogeneous networks (SM), in part because of the contagious power of content that elicits sectarian fear or indignation.

Third, in contrast to the equivocal ideological-polarization trends among the public, politicians and other political elites have unambiguously polarized recently on ideological grounds, with Republican politicians moving further to the right than Democratic politicians have moved to the left (SM). This ideological divergence is driven in part by extreme economic inequality in America today, especially in conjunction with candidates becoming increasingly reliant on ideologically extreme donors. As politicians chase campaign dollars, these extreme voices garner disproportionate influence (SM).

The ideological divergence of political elites contributes to political sectarianism, especially as these individuals increasingly use disciplined messaging to discuss their preferred topics in their preferred manner (SM). Such messaging leads the public to perceive sharper ideological distinctions between the parties than actually exists, which inflames sectarianism (SM). In addition, Newt Gingrich and his followers achieved electoral success with strongly moralized language in the 1980s and 1990s, inspiring political elites on both sides to double down on the rhetoric

¹Northwestern University, Evanston, IL, USA. ²Duke University, Durham, NC, USA. ³Harvard University, Cambridge, MA, USA. ⁴University of California, Irvine, Irvine, CA, USA. ⁵Stanford University, Stanford, CA, USA. ⁶University of Arizona, Tucson, AZ, USA. ⁷University of Maryland, College Park, MD, USA. ⁸Dartmouth College, Hanover, NH, USA. ⁹Massachusetts Institute of Technology, Cambridge, MA, USA. ¹⁰University of Illinois at Chicago, Chicago, IL, USA. ¹¹New York University, New York, NY, USA. Email: finkel@northwestern.edu; druckman@northwestern.edu

of moral outrage (e.g., “disgraceful,” “shameful”), further exacerbating sectarianism (SM).

These three trends—identity alignment, the rise of partisan media, and elite ideological polarization—have contributed to radically different sectarian narratives about American society and politics. Although the content of these narratives is entirely different across the political divide, their structure is similar: The other side cheats, so our side would be foolish to adhere to long-standing democratic norms. These narratives, which partisans experience less as stories than as truth (SM), increase their willingness to sacrifice those norms in pursuit of partisan ends.

DARK CONSEQUENCES

Rising political sectarianism has, not surprisingly, increased the social distance between Democrats and Republicans. Compared to a few decades ago, Americans today are much more opposed to dating or marrying an opposing partisan; they are also wary of living near or working for one. They tend to discriminate, as when paying an opposing partisan less than a copartisan for identical job performance or recommending that an opposing partisan be denied a scholarship despite being the more qualified applicant (1). They are also susceptible to motivated partisan cognition—seeking out, believing, and approving of information more readily when it reflects positively on copartisans or negatively on opposing partisans (10) (SM)—although scholars debate whether such tendencies are equally strong among Democrats and Republicans (see the second box).

These manifestations of political sectarianism echo those that emerge from religious sectarianism. What is distinctive about political sectarianism—beyond its largely non-theological foundation—is the immediacy of its links to governance. Political sectarianism compromises the core government function of representation. Because sectarian partisans almost never vote for the opposition, politicians lack the incentive to represent all of their constituents. Straight-ticket voting has grown increasingly widespread. In contested districts, the correlation of the Democratic share of the House vote and the Democratic share of the presidential vote—the association of district-level with national representation—surged from 0.54 in the 1970s to 0.94 by the 2010s (2).

Perhaps most troubling of all, the political sectarianism of the public incentivizes politicians to adopt antidemocratic tactics when pursuing electoral or political victories. A recent experiment shows that, today, a majority-party candidate in most U.S. House districts—Democrat or Republican—could get elected despite openly violating democratic principles like electoral fairness,

checks and balances, or civil liberties (17). Voters’ decisions to support such a candidate may seem sensible if they believe the harm to democracy from any such decision is small while the consequences of having the vile opposition win the election are catastrophic. However, the accumulation of such choices undermines representative democracy. And a society that pretends to adhere to democratic principles but actually does not is one in which people who possess resources and influence can leverage democratic gray zones to impose their will on those who do not.

Sectarianism stimulates activism (SM), but also a willingness to inflict collateral

gating climate change, reducing the federal debt, and safeguarding democratic rights.

America’s response to the coronavirus disease 2019 (COVID-19) pandemic highlights the perils of political sectarianism. An October 2019 report from Johns Hopkins University suggested that America was better prepared for a pandemic than any other nation (SM), but that report failed to account for the sort of political sectarianism that would, months later, make mask-wearing a partisan symbol, one favored more by Democrats than by Republicans. Democrats were also more likely to prioritize stay-at-home orders despite their massive, immediate eco-

Is motivated partisan cognition bipartisan?

The extent to which each side exhibits motivated partisan (or biased) cognition is a focus of ongoing debate. Some scholars argue for symmetry (SM). For example, a recent meta-analysis demonstrates equivalent levels of motivated partisan cognition across 51 experiments investigating the tendency to evaluate otherwise identical information more favorably when it supports versus challenges one’s political beliefs or allegiances (14). In an illustrative experiment, liberals and conservatives viewed a film clip of a political demonstration in which protestors clashed with police. Despite viewing the identical clip, liberals rated the protestors as more violent when they believed it was an anti-abortion protest (a conservative cause) rather than a gay-rights protest (a liberal cause), whereas conservatives exhibited the opposite pattern (SM).

Other scholars argue for asymmetry. For example, some evidence suggests that, relative to Democrats, Republicans have a higher need for order and greater trust in their gut-level intuitions. Such tendencies appear to motivate them to favor explanations that are straightforward and intuitive rather than complex and abstract, even when the latter types of explanation might be more accurate (15) (SM). Such findings are representative of the existing evidence, but conclusions remain tentative.

damage in pursuit of political goals (SM) and to view copartisans who compromise as apostates (SM). As political sectarianism has surged in recent years, so too has support for violent tactics (SM). In addition, highly sectarian partisans are vulnerable to exploitation. In 2016, Russia sought to stoke partisan outrage during America’s election by creating fake social-media avatars with names like “Blacktivist” and “army_of_jesus.” These efforts succeeded in duping sectarian extremists—especially those who were older and more conservative than average—into amplifying the avatars’ memes about the depravity of opposing partisans (SM). In doing so, these partisans served as pawns in Russia’s efforts to weaken America.

Political sectarianism also undermines the core government function of competence—of providing for and protecting the people. Members of Congress increasingly prioritize partisan purity over the sorts of compromises that appeal to a large proportion of the population, a tendency that creates legislative gridlock. Issues that are not inherently partisan become politicized, impeding the ability to make progress on goals like miti-

nomic cost—a pattern that was especially prominent among highly sectarian partisans (SM). This schism, fomented in part by President Trump, pushed toward a disequilibrium in which too few people engaged sufficiently in commerce to stimulate economic growth while too few social-distanced sufficiently to contain the pandemic. The result has been lethal and expensive for Americans across the political spectrum.

MITIGATING SECTARIANISM

Political sectarianism is neither inevitable nor irreversible. When considering promising avenues for intervention, the goal is not to restore America to some halcyon republic of yore. As exemplified by the 1870s transition from the relatively antiracist Reconstruction era to the deeply racist Redemption era, many historic episodes of partisan comity rested upon bipartisan support for (or at least acquiescence to) antidemocratic institutions and behaviors, including the marginalization and disenfranchisement of women and racial minorities. The current divide is so potent in part because battles surrounding sexism and racism have grown strongly partisan.

Rather, the goal of these interventions is to move toward a system in which the public forcefully debates political ideals and policies while resisting tendencies that undermine democracy and human rights. Given that substantial swaths of American society (including many who identify as Democrat or Republican) are fed up with surging sectarianism (SM), dedicated efforts to mitigate it may well land in fertile soil. Such efforts must circumvent the sectarian true believers, profiteers, and chaos-seekers who benefit from stoking sectarianism. These actors contribute directly to political sectarianism, and they leverage the government sclerosis caused by political sectarianism to derail efforts to address structural sources of that sectarianism, such as economic inequities and biased electoral procedures (SM).

Nonetheless, scholars have begun to identify procedures that can potentially mitigate political sectarianism. These include efforts to help Americans comprehend opposing partisans regardless of their level of agreement, such as by focusing on commonalities rather than differences (e.g., “we’re all Americans”; SM) or communicating in the moral language of the other side (e.g., when liberals frame the consequences of climate change in terms of sanctity violations; SM).

Here, we consider three avenues for intervention that hold particular promise for ameliorating political sectarianism. The first addresses people’s faulty perceptions or intuitions. For example, correcting misperceptions of opposing partisans, such as their level of hostility toward one’s copartisans, reduces sectarianism (5) (SM). Such correction efforts can encourage people to engage in cross-party interactions (SM) or to consider their own positive experiences with opposing partisans, especially a friend, family member, or neighbor. Doing so can reduce the role of motivated partisan reasoning in the formation of policy opinions (SM).

A related idea is to instill intellectual humility, such as by asking people to explain policy preferences at a mechanistic level—for example, why do they favor their position on a national flat tax or on carbon emissions. According to a recent study, relative to people assigned to the more lawyerly approach of justifying their preexisting policy preferences, those asked to provide mechanistic explanations gain appreciation for the complexities involved (SM). Leaders of civic, religious, and media organizations committed to bridging divides can look to such strategies to reduce intellectual self-righteousness that can contribute to political sectarianism.

A second avenue involves altering social-media platforms, although some popular ideas along these lines may be counterproductive. Echo chambers are widely blamed

for surging sectarianism, but simply tweaking algorithms to show partisans more content from the opposition may aggravate sectarianism rather than reducing it (7). More promising are interventions that encourage people to deliberate about the accuracy of claims on social media, which causes them to evaluate the substance of arguments and reduces their likelihood of sharing false or hyperpartisan content (12) (SM). Another option is to use crowdsourcing to identify such content and the outlets that emit it, relying on users’ ratings of trustworthiness to augment the efforts of professional fact-checkers. Such information can be incorporated into algorithmic rankings to reduce the presence of false or hyperpartisan content in people’s news feeds (SM).

A third avenue involves creating incentives for politicians and other elites to reduce their sectarianizing behaviors. People become less divided after observing politicians treating opposing partisans warmly, and nonpartisan statements from leaders can reduce violence. Campaign finance reform may help, especially by eliminating huge contributions from ideological extremists (SM). Reducing partisan gerrymandering likely would make representation fairer, generate more robust competition in the marketplace of political ideas, and send fewer extremists to the House of Representatives (SM).

A FIERCE URGENCY

In 1950, the American Political Science Association issued a report expressing concern that America was insufficiently polarized, a perspective that remained dominant in the ensuing decades (SM). Ideological differentiation is an essential feature of party-based democracy, sharpening debates on important topics. Because most people lack the expertise required to make informed judgments on specific policies, divergent and internally coherent party platforms function as helpful heuristics that voters can use to prioritize their preferred policies and hold politicians accountable.

But the ideological polarization the American Political Science Association had in mind has, in recent decades, been eclipsed among the public by political sectarianism. When politics becomes an identity-based struggle against depraved opponents—when ideals and policies matter less than dominating foes—government becomes dysfunctional. Viable political strategies focus less on policy-based arguments and more on marginalizing the opposition, with false smears and antidemocratic tactics if necessary. Insofar as politicians are pursuing unpopular policies, they are incentivized to destroy the idea of objectivity altogether, undermining the reputation of

fact-checkers and mobilizing sectarian loyalists to believe “alternative facts.”

As political sectarianism grows more extreme, pushing strong partisans deeper into congenial media enclaves that reinforce their narratives of moral righteousness, it may also become self-reinforcing, rendering mitigation efforts more difficult. Scholars have long argued that a shared threat can bring people together; indeed, some suggest that rising sectarianism in America is due in part to the loss of the Soviet Union as a unifying arch-nemesis. But such threats may do the opposite when sectarianism is extreme. COVID-19 offered a test case (SM). By the summer of 2020, 77% of Americans believed that the nation had grown more divided since the pandemic arrived that winter, a response 2.8 standard deviations higher than the mean of the 13 other nations in the study and 1.6 standard deviations higher than the second-highest nation (Spain). Such findings underscore the urgent need to counteract sectarianism before it grows more poisonous.

Political sectarianism cripples a nation’s ability to confront challenges. Bolstering the emphasis on political ideas rather than political adversaries is not a sufficient solution, but it is likely to be a major step in the right direction. The interventions proposed above offer some promising leads, but any serious effort will require multifaceted efforts to change leadership, media, and democratic systems in ways that are sensitive to human psychology. There are no silver bullets. ■

REFERENCES AND NOTES

1. S. Iyengar, Y. Lelkes, M. Levendusky, N. Malhotra, S. J. Westwood, *Annu. Rev. Polit. Sci.* **22**, 129 (2019).
2. A. I. Abramowitz, S. Webster, *Elect. Stud.* **41**, 12 (2016).
3. Y. Lelkes, *Public Opin. Q.* **80** (S1), 392 (2016).
4. L. Mason, *Uncivil Agreement: How Politics Became Our Identity* (Univ. of Chicago Press, 2018).
5. J. Lees, M. Cikara, *Nat. Hum. Behav.* **4**, 279 (2020).
6. D. J. Ahler, G. Sood, *J. Polit.* **80**, 964 (2018).
7. C. A. Bail et al., *Proc. Natl. Acad. Sci. U.S.A.* **115**, 9216 (2018).
8. H. Allcott, L. Braghieri, S. Eichmeyer, M. Gentzkow, *Am. Econ. Rev.* **110**, 629 (2020).
9. W. J. Brady, J. A. Wills, J. T. Jost, J. A. Tucker, J. J. Van Bavel, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 7313 (2017).
10. J. J. Van Bavel, A. Pereira, *Trends Cogn. Sci.* **22**, 213 (2018).
11. M. H. Graham, M. W. Svoblik, *Am. Polit. Sci. Rev.* **114**, 392 (2020).
12. G. Pennycook, J. McPhetres, Y. Zhang, J. G. Lu, D. G. Rand, *Psychol. Sci.* **31**, 770 (2020).
13. L. Boxell et al., “Cross-country trends in affective polarization (no. w26669),” *National Bureau of Economic Research*; www.nber.org/papers/w26669 (2020).
14. P. H. Ditto et al., *Perspect. Psychol. Sci.* **14**, 273 (2019).
15. J. Baron, J. T. Jost, *Perspect. Psychol. Sci.* **14**, 292 (2019).

ACKNOWLEDGMENTS

We thank the Kellogg School of Management’s Dispute Resolution Research Center and the Institute for Policy Research, both at Northwestern University. We thank S. Matz and A. Wilson for feedback on an earlier draft and T. Brader, D. Costanzo, M. DeBell, L. Harbridge-Yong, E. Groenendyk, M. Levendusky, and S. Westwood for responding to questions.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/533/suppl/DC1
10.1126/science.abe1715



A woman examines a model of a Neanderthal man at the Natural History Museum in London in 2014.

BOOKS *et al.*

PALEOANTHROPOLOGY

The vibrant lives of Neanderthals

A nuanced portrait of our evolutionary cousins encourages empathy and understanding

By Emma Pomeroy

Over the past several decades, many academic and popular writers have attempted to narrow the long-entrenched gulf between humans and Neanderthals, focusing, for example, on the misinterpretations and racist presumptions of the 19th and early 20th centuries out of which the dim view of our evolutionary cousin arose, or on more recent paleoanthropological, archaeological, and genetic evidence suggesting that they interbred with our ancestors and displayed a range of sophisticated behaviors. In her new book, *Kindred*, Rebecca Wragg Sykes nevertheless brings something new to this discussion.

The book's scope is impressive, spanning from the initial discoveries and interpretations of Neanderthals to the diverse aspects of their biology and behavior, including their childhoods, lifestyles, technologies, art, and approaches to death. All of these individual components are beautifully detailed, combining overviews of state-of-the-

art evidence with parallels from everyday experience, which frame the interpretations and challenges inherent in the data in more familiar terms.

A recurring perspective in *Kindred* is that diversity can and should be expected in Neanderthal behavior, given their existence over hundreds of thousands of years and their vast geographical range. The culture of groups separated by space or time would likely have seemed just as foreign (had they met) as it does when we encounter unfamiliar cultures today, argues Wragg Sykes. "Those living in Wales would have been surprised—even turned their noses up—at what others of their own kind from Palestine consumed with relish," she suggests, for example, about the population's varied food cultures and preferences.

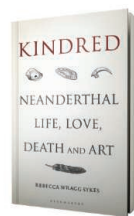
Wragg Sykes evaluates the available evidence on Neanderthals with empathy and even-handedness, revealing the group to be less "them" and more "us." She rejects aggression-centered narratives that have previously shaped our interpretations of

Neanderthal interactions with each other and with human ancestors, and she convincingly argues that we must recognize the potential role of other characteristics, including cooperation, sharing, and social bonds, in shaping their lives.

Wragg Sykes takes a similarly considered approach in discussing why Neanderthals went extinct. Rather than framing the question in terms of winners and losers, or superior and inferior species, she evaluates modern humans and Neanderthals in parallel, comparing the successes and shortcomings of each. While it is true that Neanderthals are the ones that went extinct, our story is not one of unmitigated success, she notes. Modern humans did not successfully settle in Europe until long after we evolved in Africa and reached Asia, for example, and those earliest modern humans in Europe were eventually replaced by later waves of migration. Meanwhile, the fact that Neanderthals existed for around 350,000 years hardly suggests failure. We, too, may yet succumb to environmental challenges or bring about our own downfall.

In *Kindred's* final chapter, Wragg Sykes tackles broader ethical questions faced in paleoanthropology and adjacent fields, which have recently come into sharp focus. She highlights the problematic ways in which human remains and cultures have been treated in research, arguing that these practices must change and, where possible, be redressed. These are uncomfortable topics to confront, but it is vital that we do so in order to understand the full legacies of the field and to hold future research to higher ethical standards.

She also addresses the possibility of reviving Neanderthal tissues in the near future and the major ethical concerns this could entail, arguing that the recent production of Neanderthal brain organoids (1), together with the unsanctioned gene editing of human babies (2), makes Neanderthal de-extinction not as far-fetched a notion as it may initially seem. Without forcing the point, *Kindred* closes by returning to the value of empathy and compassion, arguing that both deserve a more prominent place in our theories about Neanderthals and in our attitudes toward our fellow humans and other sentient creatures. ■



Kindred
Rebecca Wragg Sykes
Bloomsbury Sigma,
2020. 400 pp.

REFERENCES AND NOTES

1. J. Cohen, *Science* **360**, 1284 (2018).
2. D. Normile, *Science* **360**, 7347 (2019).

10.1126/science.abe5143

MARINE CONSERVATION

A profound plan to save the seas

Rejecting piecemeal strategies, a conservationist encourages total ocean protection

By **Mary Ellen Hannibal**

A long-time marine conservation advocate, Deborah Rowan Wright has spent decades supporting piecemeal strategies to protect the oceans. Yet despite her efforts and those of many others, our ocean life-support system continues to buckle under human pressures. We have been approaching marine conservation backward, she argues at the outset of her new book, *Future Sea*. Instead of regulating individual fisheries or putting boundaries around select areas of the ocean, we need to protect the whole thing.

The ocean waters that make Earth habitable for humans are becoming increasingly polluted with plastics and chemicals; they are acidifying and warming. They are also losing species at a rate that could drive many to extinction. The profound interconnectivity of ocean organisms means that negative impacts are amplified throughout the food web. But, according to Rowan Wright, the biggest problem many ocean denizens face is a shortage of food. Humanity is eating it all. In 2018, the Food and Agriculture Organization of the United Nations reported 4.6 million fishing vessels in operation, harvesting two-thirds of commercial stocks to capacity and the final third in biologically unsustainable numbers.

Science has long shown the negative impacts of lower-trophic-level overharvests on the rest of the food web (1). In Antarctica, we harvest krill—small planktonic crustaceans near the bottom of the food chain—for use in health supplements and fish-farm feed. Krill are eaten by a wide panoply of seabirds, by invertebrates such as squid, and by fish. They are also an important part of the whale diet. Take out the krill, and many other species go hungry; eventually, the ecosystem will collapse.

Removal of top predators likewise has negative impacts on the process, function, and resilience of the ecosystem (2). Humans like to eat bluefin tuna, for example, which have been hunted nearly to extinction.

As Rowan Wright probes catch limits putatively designed to prevent collapse of the fishery, she finds that, often, these limits are set well above those recommended by ex-

perts. “It’s a familiar and tedious tale,” she writes, “authorities giving way to the short-term economic demands of the fishing industry rather than ensuring the long-term viability of wild fish populations.”

Oceans are vulnerable to what Garrett Hardin memorably called “the tragedy of the commons,” wherein individuals fail to act in accordance with the common good and collectively exploit a shared resource (3). This is particularly true of the high seas, where both the lack of sovereign jurisdiction and technological advances have allowed humanity to penetrate the ocean’s farthest reaches, leading to the decimation of harvests even in the most remote waters.

Rowan Wright’s solution to the current plight of the seas is to reinvigorate existing international laws according to economist Elinor Ostrom’s principles for good commons governance. These include a “common-purpose, cross-sector, tiered (but non-hierarchical) system...rooted in the ecosystem-based and precautionary approaches and coordinated by a central agency.”

As she sets out to aggregate the various treaties and agreements that address human impacts on the ocean, Rowan Wright is surprised to learn that “virtually all the seas and oceans in the whole world are already protected by international law.” The comprehensive United Nations effort—called, simply, the Law of the

Sea—is an existing treaty signed by 180 countries (the United States is notably absent), which covers most of the bases necessary to mitigate damage to the ocean, from catch-size limits to guidelines for minimizing pollution.

The problem, argues Rowan Wright, is lax enforcement and apathy. Sovereignities are inconsistent and self-serving when it comes to upholding the rules to which they have agreed, and we have not put pressure on our governments to do better. “Most of us simply don’t see the connection between healthy oceans and a better future (in fact, any future)” she writes.

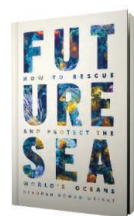
Last month, the sun did not rise in San Francisco. Or at least that is what it felt like, as smoke from fires around the West blotted out the light. Amid the helplessness and panic, I kept thinking: if only. If only we could go back 20 years and listen to the natural resource managers, scientists, and Indigenous peoples who predicted this exact scenario and offered strategies to mitigate such

a tragedy. We know the disaster that awaits us when the ocean reaches its tipping point. Will this ecosystem meet the same fate? ■

REFERENCES AND NOTES

1. A. D. M. Smith *et al.*, *Science* **333**, 1147 (2011).
2. J. A. Estes *et al.*, *Science* **333**, 301 (2011).
3. G. Hardin, *Science* **162**, 1243 (1968).

10.1126/science.abe0339



Future Sea: How to Rescue and Protect the World's Oceans

Deborah Rowan Wright
University of Chicago
Press, 2020. 200 pp.



Sourdough offers a lesson in microbial fermentation.

PODCAST

Science and Cooking: Physics Meets Food, From Homemade to Haute Cuisine

Michael Brenner, Pia Sörensen, David Weitz
Norton, 2020. 320 pp.

From melting temperature to phase transformation, cooking holds great pedagogical potential for communicating scientific principles. This week on the *Science* podcast, Pia Sörensen reveals how experiments in the kitchen can serve as an easy entry point for understanding a range of physical and chemical concepts. <http://scim.ag/2Hko12K>

10.1126/science.abf0517

The reviewer is the author of *Citizen Scientist: Searching for Heroes and Hope in an Age of Extinction* (The Experiment, 2016). Email: maryellenhannibal@gmail.com



The mink industry is associated with human disease, animal cruelty, and environmental pollution.

Edited by Jennifer Sills

Ban unsustainable mink production

More than 50 million minks are produced annually—primarily in China, Denmark, the Netherlands, and Poland (1)—to sustain the demand from the international fur industry (2, 3). In June, the Netherlands adopted a proposal to ban mink farming by the end of 2020 instead of phasing out the industry through 2024 (1). The Netherlands' decision was a step toward better protections for humans, animals, and the environment. China, Denmark, and Poland should support and extend the immediate and complete ban of mink production.

Minks harbor zoonotic infectious diseases such as leptospirosis, which leads to kidney failure in humans (4), and coronavirus disease 2019 (COVID-19), the illness caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). COVID-19 has been found in minks in the Netherlands, Denmark, Spain, and the United States (5, 6). In October, the spread of COVID-19 in Danish mink farms spun out of control, threatening the country's disease control and public health and resulting in the culling of millions of minks (7). Similar cases are likely in China, which harbors a multispecies fur industry (8). A recent study of

the SARS-CoV-2 genome and mutations suggests that the virus might be able to spread from minks to humans (9).

The mink industry is problematic beyond its potential to spread disease. Animal welfare practices fall short, with minks showing signs of fearfulness, self-mutilation, infanticide, and breeding difficulties (10). Mink production also leads to high N_2O emissions, eutrophication, and water consumption (11); the climate footprint of producing 1 kg of mink fur is 5 times as high as the footprint of producing 1 kg of wool (11).

The unsustainable use of natural resources by the mink industry impedes our response and recovery from the pandemic, puts animals in harmful conditions, and jeopardizes our ability to achieve the United Nations' Sustainable Development Goals (12). It is urgent to monitor, restrict, and—where possible—ban mink production.

Changlei Xia¹, Su Shiung Lam^{2,1}, Christian Sonne^{3,1*}

¹Nanjing Forestry University, Nanjing, China.

²Universiti Malaysia Terengganu, Terengganu, Malaysia. ³Aarhus University, Roskilde, Denmark.

*Corresponding author. Email: cs@bios.au.dk

REFERENCES AND NOTES

1. Party for the Animals, "The end of the Dutch mink farming industry—one of the largest in the world" (2020); www.partyfortheanimals.com/the-end-of-the-dutch-mink-farming-industry-one-of-the-largest-in-the-world.
2. Copenhagen Fur, "World production of mink drops" (2016); www.kopenhagenfur.com/en/news/2016/june/world-production-of-mink-drops/.
3. Danish Agriculture and Food Council, "Mink and Fur" (2020); <https://agricultureandfood.dk/danish-agriculture-and-food/mink-and-fur>.

4. F. Ramirez-Pizarro *et al.*, *Vector Borne Zoonotic Dis.* **19**, 908 (2019).
5. E. Cahan, *Science* **10.1126/science.abe3870** (2020).
6. S. Mallapaty, *Nature* **10.1038/d41586-020-01574-4** (2020).
7. "Coronavirus: Denmark culls 2.5 million mink amid outbreak on dozens of fur farms," *Independent* (2020).
8. J. Dalton, "Coronavirus fears as China plans to protect vast mink and fox fur industry from backlash at wildlife trade," *Independent* (2020).
9. B. B. O. Munnink *et al.*, *bioRxiv* **10.1101/2020.09.01.277152** (2020).
10. R. Bale, "Fur farms still unfashionably cruel, critics say," *National Geographic* (2016).
11. CE Delft, "The environmental impact of mink fur production" (2011).
12. J. Tollefson, *Nature* **584**, 175 (2020).

10.1126/science.abf0461

Support U.S. research during COVID-19

Colleges and universities are critical components of the U.S. innovation ecosystem. These institutions are called upon to play ever-evolving roles in building talent for a changing workforce; achieving scientific breakthroughs; creating new technologies, products, and companies; and contributing to local economic development. Yet, as the pace of change accelerates across our economy, federal and state budget constraints have made meeting these expectations increasingly challenging. The federal commitment to research and development stands at a multidecadal low as a percentage of GDP (1). Now, the coronavirus disease

2019 (COVID-19) pandemic has disrupted almost all aspects of higher education, including the ability to keep laboratories open, conduct research in a timely manner, collect and process data, and collaborate with colleagues and students.

As colleges and universities across the nation make difficult decisions to advance their vital missions this fall, the \$55 billion in federal support for university-performed R&D (i.e., on-campus research) is at risk (2). Maintaining the strength of the U.S. research enterprise—the same research enterprise that has enabled the rapid sequencing of the COVID-19 genome and launched numerous treatment and vaccine studies—must be a national priority. Laboratories must remain open, and researchers must be allowed to continue data collection and analysis, with all the necessary health protocols in place.

We cannot afford to shut down critical projects with long-term national benefits or to postpone projects that provide the hands-on graduate and undergraduate student research experiences necessary to train the next generation of scientists and engineers. In these difficult times, we call upon the federal government to provide the leadership, critical funding, and programmatic flexibility necessary to enable the nation's colleges and universities to continue the U.S. commitment to research, exploration, and new knowledge creation that will power our economy and provide opportunity for all.

Michael M. Crow^{1,2}, Jonathan Alger³, Michael Amiridis⁴, Dennis Assanis⁵, Eric Barron⁶, Mark P. Becker⁷, Rebecca M. Blank⁸, Gene D. Block⁹, Lee C. Bollinger¹⁰, Robert A. Brown¹¹, Sylvia M. Burwell¹², C. Michael Cassidy¹³, James P. Clements¹⁴, Steven Currall¹⁵, John J. DeGioia¹⁶, Julio Frenk¹⁷, W. Kent Fuchs¹⁸, Joan T. A. Gabel¹⁹, Patrick D. Gallagher²⁰, E. Gordon Gee²¹, Eric Isaacs²², Lloyd A. Jacobs²³, Farnam Jahanian²⁴, John Jenkins Sr.²⁵, Kristina Johnson²⁶, Martha Kanter²⁷, Paul K. Kearns²⁸, Mark Kennedy²⁹, Pradeep K. Khosla³⁰, Laurie Leshin³¹, Michael Lovell³², Gary S. May³³, Jere Morehead³⁴, Richard B. Myers³⁵, Luis M. Proenza³⁶, Clayton Rose³⁷, M. David Rudd³⁸, Timothy D. Sands³⁹, Mark S. Schlissel⁴⁰, Elisa Stephens⁴¹, Elizabeth Stroble⁴², Kumble Subbaswamy⁴³, Satish Tripathi⁴⁴, Ruth Watkins⁴⁵, Adam Weinberg⁴⁶, Kim Wilcox⁴⁷, Deborah Wince-Smith^{48*}, Wendy Wintersteen⁴⁹, W. Randolph Woodson⁵⁰, Robert J. Zimmer⁵¹

¹University Vice-Chair, Council on Competitiveness, Washington, DC 20006, USA. ²President, Arizona State University, Tempe, AZ 85281, USA. ³President, James Madison University, Harrisonburg, VA 22807, USA. ⁴Chancellor, University of Illinois at Chicago, Chicago, IL 60607, USA. ⁵President, University of Delaware, Newark, DE 19716, USA. ⁶President, Pennsylvania State University, University Park, PA 16802, USA. ⁷President, Georgia State University, Atlanta, GA 30302, USA. ⁸Chancellor, University of Wisconsin–Madison, Madison, WI 53706, USA. ⁹Chancellor, University of California, Los Angeles, Los Angeles, CA 90095, USA. ¹⁰President,

Columbia University, New York, NY 10027, USA. ¹¹President, Boston University, Boston, MA 02215, USA. ¹²President, American University, Washington, DC 20016, USA. ¹³Director, Emory Biomedical Catalyst, Emory University, Atlanta, GA 30322, USA. ¹⁴President, Clemson University, Clemson, SC 29634, USA. ¹⁵President, University of South Florida, Tampa, FL 33620, USA. ¹⁶President, Georgetown University, Washington, DC 20057, USA. ¹⁷President, University of Miami, Miami, FL 33123, USA. ¹⁸President, University of Florida, Gainesville, FL 32611, USA. ¹⁹President, University of Minnesota, Minneapolis, MN 55455, USA. ²⁰Chancellor, University of Pittsburgh, Pittsburgh, PA 15213, USA. ²¹President, West Virginia University, Morgantown, WV 26506, USA. ²²President, Carnegie Institute for Science, Washington, DC 20005, USA. ²³Former President, University of Toledo, Toledo, OH 43606, USA. ²⁴President, Carnegie Mellon University, Pittsburgh, PA 15213, USA. ²⁵President, University of Notre Dame, Notre Dame, IN 56446, USA. ²⁶President, The Ohio State University, Columbus, OH 43210, USA. ²⁷Executive Director, College Promise Campaign, Washington, DC 20009, USA. ²⁸Laboratory Director, Argonne National Laboratory, Lemont, IL 60439, USA. ²⁹President, University of Colorado, Boulder, CO 80309, USA. ³⁰Chancellor, University of California, San Diego, La Jolla, CA 92093, USA. ³¹President, Worcester Polytechnic Institute, Worcester, MA 01609, USA. ³²President, Marquette University, Milwaukee, WI 53233, USA. ³³Chancellor, University of California, Davis, Davis, CA 95616, USA. ³⁴President, University of Georgia, Athens, GA 30602, USA. ³⁵President, Kansas State University, Manhattan, KS 66506, USA. ³⁶President Emeritus, The University of Akron, Akron, OH 44325, USA. ³⁷President, Bowdoin College, Brunswick, ME 04011, USA. ³⁸President, University of Memphis, Memphis, TN 38152, USA. ³⁹President, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA. ⁴⁰President, University of Michigan, Ann Arbor, MI 48109, USA. ⁴¹President, Academy of Art University, San Francisco, CA 94105, USA. ⁴²Chancellor, Webster University, St. Louis, MO 63119, USA. ⁴³Chancellor, University of Massachusetts, Amherst, Amherst, MA 01003, USA. ⁴⁴President, University at Buffalo, Buffalo, NY 14260, USA. ⁴⁵President, University of Utah, Salt Lake City, UT 84112, USA. ⁴⁶President, Denison University, Granville, OH 43023, USA. ⁴⁷Chancellor, University of California, Riverside, Riverside, CA 92521, USA. ⁴⁸President & CEO, Council on Competitiveness, Washington, DC 20006, USA. ⁴⁹President, Iowa State University, Ames, IA 50011, USA. ⁵⁰Chancellor, North Carolina State University, Raleigh, NC 27607, USA. ⁵¹President, The University of Chicago, Chicago, IL 60637, USA.

*Corresponding author.

Email: deborahws@compete.org

All authors are members of the Council on Competitiveness.

REFERENCES AND NOTES

1. AAAS, "Federal R&D as a Percent of GDP" (2020); www.aaas.org/sites/default/files/2020-07/RDGDGP.png.
2. AAAS, "R&D at colleges and universities" (2018); www.aaas.org/programs/r-d-budget-and-policy/rd-colleges-and-universities.

10.1126/science.abf1225

Hainan peacock-pheasant needs focused protection

The Hainan peacock-pheasant (*Polyplectron katsumatae*), a distinct species endemic to Hainan island (1, 2), requires protection. Since the 1950s (3), the Hainan peacock-pheasant population has dispersed into small, partially isolated subpopulations. As

their habitat of tropical and seasonal rainforests suffered from overexploitation and deforestation (2, 4), the peacock-pheasant population declined from about 2700 in 1990 (4) to only about 300 individuals in 2000 (3). With such small numbers, Hainan peacock-pheasants could lose their ability to maintain a self-sustaining population (5).

The Hainan peacock-pheasant has been categorized as Endangered on the International Union for Conservation of Nature (IUCN) Red List since 2016 (6) and is classified as Category I in China's list of nationally protected animals (3). Despite this recognition, current conservation efforts are insufficient (7). The Chinese government should continue to strengthen the protection of mountain rainforest vegetation and prohibit recreational hunting, especially at the altitudes the peacock-pheasants inhabit [up to about 1500 meters (8)]. Fortunately, Hainan Tropical Rainforest National Park, now under construction, could provide the peacock-pheasant with a more stable habitat (9, 10). If the peacock-pheasant's original habitat can no longer meet the needs of survival and reproduction, rescue measures should be taken, including breeding bases, gene banks, germplasm resource banks, and relocation to protected areas where they are more likely to thrive.

Song Guan^{1,2*}, Hanlin Zhou^{1,2}, Yanfu He³

¹Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, China. ²Institute of Zoology, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, China. ³College of Food Science and Engineering, Hainan University, Haikou 570228, China.

*Corresponding author.

Email: guansong79@126.com

REFERENCES AND NOTES

1. G. M. Zheng, *A Checklist on the Classification and Distribution of the Birds of China* (Science Press, Beijing, 2005).
2. J. Chang et al., *Zool. Sci.* **25**, 30 (2008).
3. G. M. Zheng, Q. S. Wang, *China Red Data Book of Endangered Animals (Aves)* (Science Press, Beijing, 1998).
4. Y. R. Gao, D. Q. Yu, *Chin. J. Zool.* **25**, 42 (1990).
5. J. Chang et al., *Chin. Sci. Bull.* **58**, 2185 (2013).
6. BirdLife International, *Polyplectron katsumatae* (IUCN Red List of Threatened Species, 2016).
7. W. Liang, Z. W. Zhang, *Chin. Birds* **2**, 116 (2011).
8. xzwu.com, "Hainan peacock-pheasant" (2015); www.xzwu.com/article-19556-1.html [in Chinese].
9. State Forestry and Grassland Administration, National Park Administration, "Protect this precious virgin tropical rainforest," *Guang-ming Daily* (2020); www.forestry.gov.cn/main/6004/20201014/085431339375989.html.
10. Hainan Provincial People's Congress Standing Committee, "Hainan Tropical Rainforest National Park regulations (trial)" (2020); http://lyj.hainan.gov.cn/xxgk/0200/0202/202009/t20200906_2845900.html [in Chinese].

10.1126/science.abf0489



The diverse 2019–2020 Science & Technology Policy Fellows “look like America,” said AAAS CEO Sudip Parikh.

AAAS drafts plan to address systemic racism in sciences

Diversity assessment within the organization’s programs among plan’s first steps

By **Becky Ham**

The American Association for the Advancement of Science has released two parts of a draft plan that directs the organization to strengthen its advocacy on behalf of diversity, equity, and inclusion—while taking a hard look at its own demographics and related policies.

AAAS’s “Addressing Systemic Racism in the Sciences” plans were developed by AAAS CEO Sudip Parikh and AAAS leadership, after discussions inside and outside the association prompted by a year that Parikh calls “an opportunity born of tragedy.”

Protests over police brutality against Black people in the United States and the disproportionate impact of COVID-19 on Black and brown communities are sharp reminders that no part of society is immune from discrimination, Parikh said. “We haven’t yet held a mirror to the scientific enterprise, and we may not like everything we see.”

AAAS is holding up that mirror with a report that will be published in late October, compiling demographic data on the association’s Fellows, Science & Technology Policy Fellows, award recipients, *Science* authors, and AAAS governing bodies. These groups are critical to supporting diversity, Parikh said, because they represent the ways in which “AAAS functions as a career advancement enabler in science and engineering.”

The report by Jen Sargent and Marietta Damond in the Office of Membership provides a baseline accounting of diversity in terms of race, ethnicity, and gender. Because most of the data are self-reported by individuals, there are gaps in coverage, said Damond. For instance, the *Science* family of journals authors and reviewers had the smallest amount of available data, with gender data available for 13% of the group (before using a statistical estimate that brought coverage up to 53%) and race and ethnicity data available for only 12% of the group.

Although the data are “messy,” Parikh acknowledged, there are still some notable trends that shed light on how AAAS’s practices may encourage or discourage diversity. For instance, “the S&T Policy Fellows look like America. It’s extremely diverse, it’s got great gender parity, has good geographic diversity,” he said. “For the honorary Fellows, the East and West Coasts and white males are overrepresented.”

S&T Policy Fellows apply for the fellowship program, while honorary Fellows are nominated by colleagues, which could account for

the difference, Parikh said, noting that the data offer a glimpse at how “baked-in” policies often can exclude minorities in science.

“Those policies, they do their work even if the people who are carrying them out have no racist tendencies, and it’s the system itself that actually causes the problem,” he said.

To bring these policies to light, AAAS will undergo the SEA Change self-assessment, said Shirley Malcom, AAAS senior adviser and the program’s director. SEA Change provides support to universities and colleges as they transform campus cultures, policies, and procedures that disadvantage or exclude participation in science, technology, engineering, mathematics, and medicine.

Malcom and her colleagues are seeking support to modify and adapt SEA Change for educational and professional organizations, she said. “Some of the societies we work with in SEA Change—disciplinary groups and professional societies—have said, ‘we don’t think we can subject our community to something we’re not willing to undergo ourselves.’”

SEA Change will move the AAAS demographic reports “beyond the numbers to how did we get to those numbers,” Malcom said. “What are our processes in place that would support either diversifying those numbers or may be a barrier to diversifying those numbers?”

“This requires a certain amount of reflection,” she added. “We’re not going to move forward unless we’re honest about what we find and what it requires us to do.”

The second part of the draft plan, released in September, discusses how AAAS programs and initiatives can help increase diversity and equity throughout the scientific enterprise. Under the plan, AAAS will advocate for increased salaries for graduate students and postdoctoral researchers, mentor training, and funding for government and industry science programs for minority students.

The need for a financial safety net for young scientists has grown with COVID-19. The pandemic is having a disproportionate effect on women and people of color in the sciences, who are most likely to bear the brunt of a collision between career advancement and family and health responsibilities, said Malcom.

At the same time, more scientists are speaking bluntly and publicly about the impacts of racism in their fields. “We have not had that in the past, and so I live in hope that we are at a different place,” she said.

Where Science Gets Social.

AAAS.ORG/COMMUNITY



AAAS' Member Community is a one-stop destination for scientists and STEM enthusiasts alike. It's "Where Science Gets Social": a community where facts matter, ideas are big and there's always a reason to come hang out, share, discuss and explore.

**Member
COMMUNITY**
AAAS

AAAS NEWS & NOTES

Encouraging diversity is more than a moral issue, Parikh said. Countries like China and India draw on "huge human capital" for their scientific enterprises, and it makes competitive sense that the United States should expand its own pool of talent, he noted. "If we're only pulling from the East and West Coasts, and we're only pulling from communities that can afford to go into the sciences, then we are fighting with one hand tied behind our back in that global competition."

The third part of the AAAS plan, which looks at diversity within the AAAS staff, will be shared in late 2020. For the AAAS leadership, the most important part of the plans is to keep them moving forward beyond the draft stage.

Parikh recalled that AAAS had a similar moment of opportunity in the late 1960s and early 1970s, when students and younger scientists protested at the association's Annual Meetings for more inclusive policies. After the 1969 meeting, a Committee on Young Scientists was formed. Students began to attend Board meetings, and AAAS developed its first signature programs in education and advocacy.

But the committee disbanded within a few years, leaving many of the diversity issues it raised unaddressed, said Parikh. "My takeaway from that is that you have to keep up momentum, you can't stop talking about the issues at hand."

AAAS annual election: Preliminary announcement

The 2020 AAAS election of general and section officers is scheduled to begin later this fall. All members will receive a ballot for election of the president-elect, members of the Board of Directors, and members of the Committee on Nominations. Additionally, members registered in sections (up to three) will receive ballots for the specified section elections. Biographical information for the candidates will be provided along with ballots. The general election slate is listed below. The list of section candidates is still being compiled.

Notice to our international members:

In an effort to conserve resources, AAAS will be sending electronic election ballots to our non-U.S.-based members. To ensure you receive your ballot, please make sure your email is up-to-date with AAAS by logging on to www.aaas.org. 1) Click on "Member Login" (if you have not yet created an account, you will be prompted to do so); 2) After you log in, click on the red "My Profile" button in the upper right-hand corner of the page; 3) Click on "Edit My Contact Information" in the left-hand side bar; 4) Update your email and click on the "Save" button.

If you would like to request a special paper ballot, please send an email with your name and address with your request to elections@aaas.org.

General Election Slate

PRESIDENT-ELECT

Gilda Barabino, Ph.D., President, Olin College of Engineering

David Ho, M.D., Director, Aaron Diamond AIDS Research Center, Clyde '56 and Helen Wu Professor of Medicine, Columbia University

DIRECTOR

R. Alta Charo, J.D., Warren P. Knowles Professor of Law and Bioethics, University of Wisconsin-Madison

Janine Clayton, M.D., Associate Director for Research on Women's Health and Director of the Office of Research on Women's Health, National Institutes of Health


Kaye Husbands Fealing, Ph.D., Dean and Ivan Allen Jr. Chair, Ivan Allen College of Liberal Arts, Georgia Tech

Nora Newcombe, Ph.D., Laura H. Carnell Professor of Psychology, Temple University

RESEARCH

IN SCIENCE JOURNALS

Edited by Michael Funk



Transmission electron microscopy image of a white blood cell infected with the parasite *Trypanosoma cruzi*, which causes Chagas disease

CHAGAS DISEASE

Trouncing trypanosomes

Trypanosoma cruzi infection causes Chagas disease in millions of individuals in Latin America, and intensive drug treatment is frequently unsuccessful. Bustamante *et al.* demonstrated that high weekly doses of oral benznidazole over 30 weeks, rather than the current treatment of smaller twice-daily doses over 2 months, resulted in better clearance of both actively replicating and dormant trypanosomes in mouse models of Chagas disease. The clearance of dormant parasites was confirmed by light sheet fluorescence microscopy, which allowed the authors to image whole organs and intact tissues of infected mice. Further studies will determine whether this drug regimen will be successful in treating patients with Chagas disease. —MN *Sci. Transl. Med.* **12**, eabb7656 (2020).

GEOPHYSICS

Mapping out backward motion

Most deformation associated with an earthquake is, not surprisingly, in the same direction as the fault rupture. Xu *et al.* used satellite imaging to find areas of deformation associated with the 2019 Ridgecrest earthquake sequence

that moved in the opposite direction. These regions moved in this direction because of inelastic deformation, which helped to accommodate the overall fault rupture for the sequence. The observations were possible because of improved radar imaging and are likely more common than previously believed. —BG

Science, this issue p. 605

CORONAVIRUS

A series of unfortunate events

The history of how severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spread around the planet has been far from clear. Several narratives have been propagated by social media and, in some cases,

national policies were forged in response. Now that many thousands of virus sequences are available, two studies analyzed some of the key early events in the spread of SARS-CoV-2. Bedford *et al.* found that the virus arrived in Washington state in late January or early February. The viral genome from the first case detected had mutations similar to those found in Chinese samples and rapidly spread and dominated subsequent undetected community transmission. The other viruses detected had origins in Europe. Worobey *et al.* found that early introductions into Germany and the west coast of the United States were extinguished by vigorous public health efforts, but these successes were largely unrecognized. Unfortunately, several major travel events occurred in February, including repatriations from China, with lax public health follow-up. Serial, independent introductions triggered the major outbreaks in the United States and Europe that still hold us in the grip of control measures. —CA

Science, this issue p. 571, p. 564

PLANT SCIENCE

Hormone-guided self-organization

Auxin signals travel through pathways in developing or regenerating plant tissues to guide tissue formation. Hajný *et al.* now show that an auxin-regulated receptor in turn regulates phosphorylation and thus subcellular localization of an auxin transporter. Through this signaling pathway, auxin builds the pathways through which it is transported, guiding development as it goes. —PJH

Science, this issue p. 550

MAGNETISM

Lightening the load

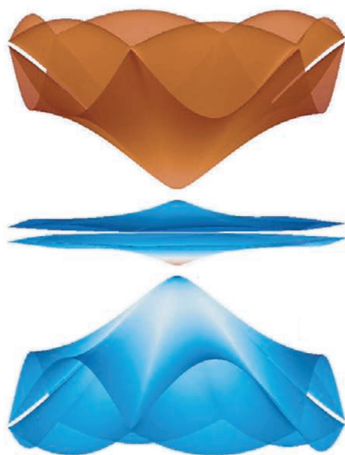
Permanent magnets are generally produced from solid metals or alloys. Less dense compositions involving lighter elements tend to demagnetize well below room temperature or under modest

applied external fields. Perlepe *et al.* now report that chemical reduction of a low-density chromium-pyrazine network produces a magnet that remains stable above 200°C and resists demagnetization with 7500-oersted coercivity at room temperature. The straightforward synthetic route to the material shows promise for broad exploration of potential applications. —JSY

Science, this issue p. 587

TOPOLOGICAL SYSTEMS Polaritons with a twist

The ability to design and fabricate optical systems with tunable topological features makes them especially attractive for developing analogs of topological condensed matter systems, which by themselves tend to be fixed or limited in their tunability. Liu *et al.* now show that the combination of a two-dimensional material with a photonic crystal can be used to develop an analogous quantum spin Hall system. The strong coupling between the monolayer tungsten disulfide excitons with a nontrivial hexagonal photonic crystal gives rise to helical topological polaritons observed at up to 200 kelvin. The topological polaritons can be actively tuned by temperature and may further be manipulated with electric or magnetic fields, thereby providing a flexible platform with which



Calculated band structure of a topological polaritonic lattice with odd (brown) or even (blue) spatial parity of the bands

to explore exotic topological phenomena and new phases of quantum matter. —ISO

Science, this issue p. 600

PANDEMIC PAUSE Songbirds reclaim favored frequencies

When severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic lockdowns were instituted across entire countries, human activities ceased in an unprecedented way. Derryberry *et al.* found that the reduction in traffic sound in the San Francisco Bay Area of California to levels not seen for half a century led to a shift in song frequency in white-crowned sparrows (see the Perspective by Halfwerk). This shift was especially notable because the frequency of human-produced traffic noise occurs within a range that interferes with the highest performance and most effective song. Thus, our “quiet” allowed the birds to quickly fill the most effective song space. —SNV

Science, this issue p. 575;
see also p. 523

ANTHROPOLOGY Temporal changes in body temperature

For people living in the United States, normal body temperature (BT) has decreased from 37.0° to 36.4°C over the past 150 years, a decline attributed to the reduced incidence of infectious diseases and to improvements in socioeconomic conditions. To understand whether this phenomenon is found in a very different context, Gurven *et al.* analyzed BT data from a rural population of indigenous farmers in rural Bolivia over a 16-year period. They determined that BT decreased by 0.5°C in less than two decades. Better access to health care contributed to improved health status but does not explain the rapid rate of BT decline. —MSA

Sci. Adv. 10.1126/sciadv.abc6599 (2020).

IN OTHER JOURNALS

Edited by Caroline Ash
and Jesse Smith



NEURODEVELOPMENT Finding the way through a complicated world

As the nervous system develops, axons grow out from immature neurons to build connections. The fruit fly, *Drosophila melanogaster*, is used as a model organism for tracking the development of the nervous system in bilaterally symmetrical animals. In the *Drosophila* midline, some axons will stay on their originating side and others will cross the midline. The secreted protein Slit prevents axons from crossing over. Things get complicated, however, as Kellermeyer *et al.* reveal with their identification of the protease Tolkin, which cleaves Slit into fragments. Although full-length Slit prohibits growing axons from crossing the midline, the amino-terminal fragment of Slit, which is also generated by Tolkin, induces the growing axons to elongate. Thus, little is left to chance in building the *Drosophila* nervous system: A negative signal on one route is itself the root of a positive signal for another route. —PJH

Development 10.1242/dev.196055 (2020).

INFLUENZA VIRUS Influenza A virus matrix protein

Influenza spreads in seasonal outbreaks that result in hundreds of thousands of deaths worldwide. The culprit, influenza A virus, has occasionally caused historic pandemics such as the 1918 flu pandemic, which killed tens of millions of people in 2 years. Despite more than a century of study, we still do not have a clear picture of many components of the virus, including the matrix protein M1, which is one of its most abundant proteins. Selzer *et al.* provide a high-resolution structure of M1 in a single-layered assembly using cryo-electron microscopy. The study offers new insight into the role of the matrix protein in maintaining the integrity of the influenza virus and may help in the development of new antivirals targeting the matrix layer of the virus. —DJ

PLOS Biol. 18, e3000827 (2020).

GENOMICS Decoding the great ape Y

The Y chromosome confers maleness in most mammals and in other species, including flies

CREDITS: LIU ET AL.; JANROZ/ALAMY STOCK PHOTO

PESTICIDES

No easy fix

Evidence has mounted over the years that neonicotinoid insecticides used for crop protection have dangerous sublethal effects. Continued use of the chemicals has led to beneficial insect decline, and many governments have taken action to ban their use. However, bans have not come with systemic changes in agricultural practices, but rather they have prompted replacement by different types of pesticides, including flupyradifurone and sulfoxaflor. Across published studies, Siviter and Muth found that these compounds have similar modes of action as neonicotinoids and similar sublethal impacts on beneficial insects. The authors argue that sublethal effects should be measured not just for pollinators but also for non-bee beneficial insects, and wider environmental interactions for any proposed pesticides must be assessed during the approval processes. —SNV *Proc. Biol. Sci.* 10.1098/rspb.2020.1265 (2020).

Beneficial insects, such as beetle predators of aphids (*Coccinella septempunctata* is shown here), succumb to the sublethal effects of a new generation of pesticides designed to replace banned neonicotinoids.

and some plants. Studies of this chromosome and its evolutionary history can inform upon evolutionary lineages and the genes involved in male fertility. However, the Y chromosome has been difficult to sequence; because of its high content of repetitive DNA and low gene number, it is often absent from genomic assemblies. Cechova *et al.* generated draft sequences of the bonobo and orangutan Y chromosomes and compared them with previously sequenced human, gorilla, and chimpanzee Y chromosomes. By comparing the Y chromosomes of all great ape species, this study provides resources for understanding the evolution of this chromosome, which can help explain population genetic differences within and between these species and may aid in conservation for our endangered closest relatives. —LMZ

Proc. Natl. Acad. Sci. U.S.A. **117**, 26273 (2020).

EXOPLANETS

Oceans protect circumbinary planets

Exoplanets that orbit a binary star, known as circumbinary planets, can experience large

variations in stellar heating over less than the planet's orbital period. Wolf *et al.* performed three-dimensional simulations of how Earth's climate would differ if it orbited in the habitable zone around a Sun-like star with a smaller stellar companion. Even in extreme cases, they found that Earth's oceans provided enough thermal inertia and negative feedback through cloud formation to buffer the planet against catastrophic climate variations. However,

temperatures on land can experience complex additional season-like variations. The authors conclude that circumbinary exoplanets can remain habitable if they have Earth-like oceans. —KTS

J. Geophys. Res. Planets
10.1029/2020JE006576 (2020).

EDUCATION

Competencies get their chance to shine

Transferrable skills, or competencies, are a critical part of an undergraduate degree in STEM (science, technology, engineering, and mathematics) but are often an underdeveloped facet of curricula. A possible remedy for this would be the transformation of a competency framework into a set of measurable learning outcomes. Using an iterative process, Clemmons *et al.* did just this by transforming and validating the core competencies from the *Vision and Change* framework into the BioSkills Guide. Compared with other science competency frameworks, the BioSkills Guide revealed some gaps and ambiguities, likely reflecting areas where still-evolving competencies in undergraduate biology merit future research. The BioSkills Guide has potential for a wide variety of applications,

including the backward design of lessons and courses, competency assessment development, and curriculum mapping and planning. —MMc

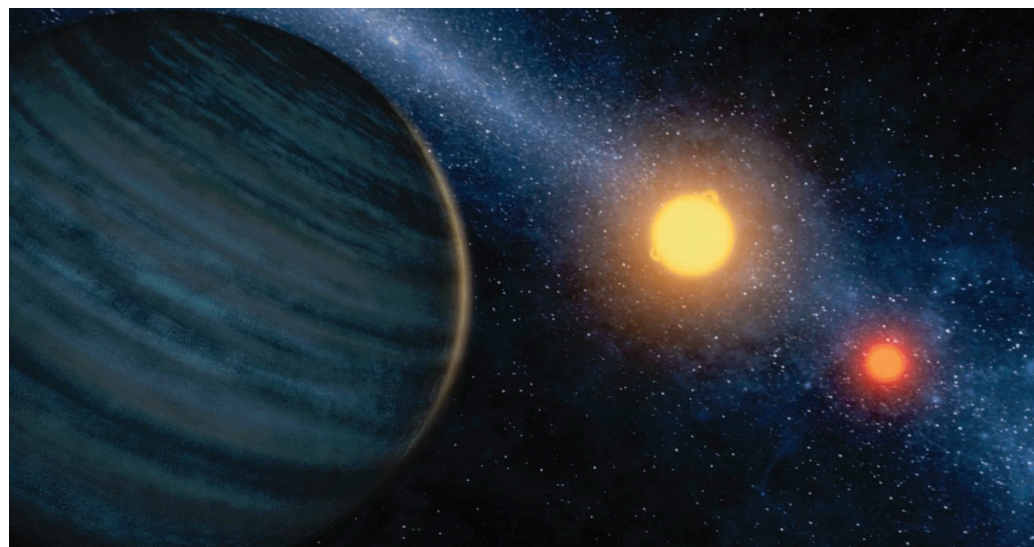
CBE Life Sci. Educ. 10.1187/cbe.19-11-0259 (2020).

BIOMATERIALS

It's all in the inks

For any printing process, the quality and properties of the final product depend on the availability of suitable inks. In bioprinting for tissue engineering, inks need to have the right combination of printability as well as biological compatibility and suitable mechanical properties for the printed parts. Lee *et al.* developed an extracellular matrix-based polymer through the combination of gelatin methacryloyl- and methacryloyl-substituted recombinant human tropoelastin. In addition to providing elasticity and resilience, tropoelastin provides cell-binding motifs and cell-signaling pathways. Printed vascularized tissue constructs were tested in vitro and in vivo and presented endothelium barrier function and spontaneous beating of cardiac cells, as well as minimal inflammatory response and in vivo degradation. —MSL

Adv. Mat. 10.1002/adma.202003915 (2020).



Planets orbiting binary stars, like the system in this artist rendering, are more likely to be habitable if they have Earth-like oceans.

ALSO IN SCIENCE JOURNALS

Edited by Michael Funk

CORONAVIRUS

Factors determining SARS-CoV-2 endemicity

Numerous factors determine whether a respiratory virus will become endemic and remain in circulation around the world. In a Perspective, Shaman and Galanti discuss the features that might contribute to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) endemicity. These include whether people can become reinfected after recovering from infection, whether there will be seasonality in the incidence of infection, and how SARS-CoV-2 competes with other respiratory viruses. The authors discuss the importance of waning immunity and immune escape and what can be learned from other endemic respiratory viruses, including the “common cold” coronaviruses and influenza virus. —GKA

Science, this issue p. 527

ANIMAL COGNITION

Forgotten ideas from a Black zoologist

Charles H. Turner (1867–1923) studied animal cognition, and despite an impressive body of work that challenged the main concepts of animal behavior, intelligence, and personality at the time, his work has largely been forgotten. As an African American, Turner was unable to obtain a university position, which likely explains why his work and ideas have not been continued and have indeed been repeated decades later. In a Perspective, Galpayage Dona and Chittka discuss Turner’s work, the challenges he faced as an African American then, and the racial disparities that continue in research today. —GKA

Science, this issue p. 530

IMMUNOLOGY

Rethinking metabolic manipulations in MS

Activation of the metabolic enzyme pyruvate kinase (PK), which catalyzes a critical step in glycolysis, has been proposed as a potential treatment for multiple sclerosis (MS). Seki *et al.* confirmed that activators of the PK isoform PKM2 limited the differentiation of T helper 17 cells, which are thought to be involved in MS disease, as previously reported. However, in a mouse model of MS, the authors found that PKM2 activators redirected disease pathology from the spinal cord to the brain instead of suppressing symptoms. These results suggest that the therapeutic potential of PKM2 activators for MS must be reassessed. —WW

Sci. Signal. **13**, eaay9217 (2020).

T CELLS

Poised for pathogenicity

The production of the cytokine granulocyte-macrophage colony-stimulating factor (GM-CSF) by T helper cells not only guides adaptive immune responses but can also contribute to the pathogenesis of autoimmune disease. Rasouli *et al.* used single-cell RNA sequencing, mass cytometry, and GM-CSF fate reporter mice to identify a distinct subset of T helper cells in humans and mice that specialize in sustained GM-CSF production but lack key phenotypic features of other T helper lineages. In a murine model of experimental autoimmune encephalomyelitis, these cells were poised to up-regulate interferon- γ within the central nervous system and promote the development of encephalitis, processes that required the transcription factor T-bet. These results provide insight into the identity of GM-CSF-producing CD4⁺ T cells and their function during autoimmune neuroinflammation. —CO

Sci. Immunol. **5**, eaba9953 (2020).

STRUCTURAL BIOLOGY

Secrets of a proton pumping machine

Mitochondrial complex I serves as a primary entry point for electrons from the tricarboxylic acid cycle into the mitochondrial electron transport chain. This massive, membrane-embedded protein complex must couple quinone reduction to conformational changes across more than 150 angstroms within four separate proton pumps. Kampjut *et al.* determined five structures of complex I in states along the catalytic cycle, a deactive conformation, and one with the inhibitor rotenone bound. The resolution of some structures was sufficient to see water molecules and to trace putative paths for proton transfer within the proton-pumping membrane domain. The structures add valuable details that provide a basis for generating mechanistic hypotheses for this crucial complex. —MAF

Science, this issue p. 547

CANCER

Physical traits of cancer

The biological and physical properties of tumors contribute to their growth and to treatment outcome. Although intense research efforts have helped to delineate cancer biology, the physics of cancer has only emerged in relatively recent times as a key area of research. Nia *et al.* reviewed the physical features that are common to tumors and that limit successful treatment: solid stresses, interstitial fluid pressure, stiffness (rigidity), and architecture and organization of tumor constituents. The authors provide a conceptual framework and discuss the origins of these distinct physical traits of cancer and how they enable and synergize with aberrant cancer biology to fuel cancer initiation, progression, immune evasion, and treatment resistance. —PNK

Science, this issue p. 546

HYBRID PEROVSKITES

Structural secrets of hybrid perovskites

The optoelectronic and photovoltaic applications of polycrystalline hybrid metal halide perovskite films are notable because grain boundaries in most materials cause scattering of charge carriers that decreases performance. Electron microscopy studies of these materials have been hindered by their rapid structural degradation under intense electron beams. Rothmann *et al.* now present an atomic crystallographic structure of formamidinium lead triiodide (FAPbI₃) polycrystalline thin films obtained by low-electron-dose scanning transmission electron microscopy with advanced image processing. The crystal structure sustains substoichiometry in the A-site cation, has a nearly perfect crystallographic alignment between PbI₂ impurity phases and the FAPbI₃ perovskite, and has atomically clean grain boundaries between polycrystalline domains. These features help to explain the films’ surprising regenerative ability, their benign grain boundaries where strain and dislocations appear mostly absent, and why excess lead-iodide precursor can be counterintuitively beneficial. —PDS

Science, this issue p. 548

MICROBIOTA

Radioprotective bacteria

A common symptom of radiation treatment for cancer is gastrointestinal disruption. The damage caused can become so severe and debilitating that it interrupts treatment. Guo *et al.* noticed that mice surviving experimental radiation exposure had distinctive taxonomic representation in their gut microbiota. A similar correlation was also observed in a small group of human subjects. Further experiments in mice revealed that some strains of

bacteria produced high levels of short-chain fatty acids, which seemed to be dampening inflammatory responses and alleviating the damage caused by reactive oxygen species released by the radiation. A metabolomics analysis also implicated a role for tryptophan metabolic pathways in radiation survivorship. —CA

Science, this issue p. 549

ANCIENT DOG GENOMICS

Dog domestication was multifaceted

Dogs were the first domesticated animal, likely originating from human-associated wolves, but their origin remains unclear. Bergstrom *et al.* sequenced 27 ancient dog genomes from multiple locations near to and corresponding in time to comparable human ancient DNA sites (see the Perspective by Pavlidis and Somel). By analyzing these genomes, along with other ancient and modern dog genomes, the authors found that dogs likely arose once from a now-extinct wolf population. They also found that at least five different dog populations ~10,000 years before the present show replacement in Europe at later dates. Furthermore, some dog population genetics are similar to those of humans, whereas others differ, inferring a complex ancestral history for humanity's best friend. —LMZ

Science, this issue p. 557;
see also p. 522

MIGRATION

Unexpectedly simple

Chinook salmon are known to return to spawn at two distinct times of the year: spring and fall. Individuals that return during these times have generally been referred to as parts of distinct groups, or ecotypes, with traits specific to their timing and presumed divergence being caused by the lack of interbreeding. By looking at genomes across fish from both runs, Thompson *et al.* found that a single genomic region of interest was nearly

perfectly associated with run timing but not with other traits such as maturity and fat reserves (see the Perspective by McKinney). Further, they conclude that the region operates as a Mendelian trait, with assortment dictating run timing and associated phenotypes being caused by the migration environment rather than genetics. —SNV

Science, this issue p. 609;
see also p. 526

PROTON MEMBRANES

Vacancies enhance proton conductivity

Proton-exchange membranes (PEMs) allow for the transport of protons while acting as electrical insulators and ensuring that reactants are kept apart; therefore, they are a key component in devices such as low-temperature fuel cells. PEMs are typically made from polymers or materials embedded in a polymer matrix and need to operate in conditions of very high humidity. Starting with an inorganic, layered material, CdPS₃, Qian *et al.* show that the removal of a small amount of cadmium introduces vacancies that greatly increase the proton conductivity of the PEMs (see the Perspective by Wang and He). The process works for manganese-based membranes as well, and high lithium ion transport was also observed. —MSL

Science, this issue p. 596;
see also p. 525

HUMAN EVOLUTION

DNA analyses of an early East Asian

Ancient, anatomically modern humans interbred with the archaic hominins Neanderthals and Denisovans. However, the extent of this interbreeding and how it affects modern populations is not well understood. Massilani *et al.* generated genome-wide data from a 34,000-year-old female individual from the Salkhit Valley in eastern Mongolia and

conducted a detailed modeling of her ancestry with regard to other Pleistocene human genomes. They found evidence for Denisovan ancestry in ancient human genomes from at least 6000 years before the Salkhit individual lived and determined that the Denisovan contribution differed from that of another ancient Asian individual, as well as from the ancient Denisovan contribution to extant Australasians. This reference point helps us to understand the early history of our species in Eurasia, especially Eastern Eurasia, for which genomic evidence remains scarce. —LMZ

Science, this issue p. 579

PALEOGENOMICS

A timeline of cave dwellers in sediment

Two archaic lineages overlapped with modern humans outside of Africa: the well-studied Neanderthals and their more mysterious cousins, the Denisovans. Denisovan remains are rare, being limited to Denisovan Cave in Siberia and a putative, undated jaw from Tibet. However, there is evidence for multiple introgressions from Denisovans into modern-day humans, especially in Australasian populations. By examining the sediment of Baishiya Karst Cave located on a high plateau in Tibet, Zhang *et al.* identified ancient mitochondrial DNA from Denisovans indicating their presence at about 100 thousand, 60 thousand, and possibly 45 thousand years ago. This finding provides insight into the timing and distribution of Denisovans in Asia and extends the time of occupation of the Tibetan plateau by hominins. —LMZ

Science, this issue p. 584

QUANTUM SYSTEMS

Addressing the many and the individual

The ability to coherently manipulate the quantum state of atomic defects in solid-state systems is

a promising route to developing a platform for quantum technologies. A successful platform requires the interaction of many qubits in close proximity, as well as the ability to address each qubit individually, and, to date, such requirements have run counter to each other. Chen *et al.* devised an optical frequency-domain method with which they were able to simultaneously address many individual rare-earth ion defects (six at this point) with separations all within the diffraction limit of the control light. Because the approach is scalable to tens or hundreds of defects, it provides the prospect of realizing truly large-scale quantum processors. —ISO

Science, this issue p. 592

REVIEW SUMMARY

CANCER

Physical traits of cancer

Hadi T. Nia, Lance L. Munn, Rakesh K. Jain*

BACKGROUND: Historically, cancer has been considered a disease of the cell, caused by mutations in genes that control proliferation, differentiation, and death. In recent decades, however, the microenvironment surrounding the cancer cell has gained notoriety as a co-conspirator in tumor initiation, progression, immune evasion, and treatment response. As tumors grow, they disrupt the structure and function of the surrounding tissue via physical and biochemical mechanisms. The resulting physical abnormalities affect both cancer cells and their microenvironment and fuel tumorigenesis and treatment resistance. The links between cancer biology and physics have provided opportunities for the discovery of new drugs and treatment strategies.

ADVANCES: Here, we propose four distinct physical cancer traits that capture the biomechanical abnormalities in tumors: (i) elevated solid stress, (ii) elevated interstitial fluid pressure, (iii) increased stiffness and altered material properties, and (iv) altered tissue microarchitecture. Solid stresses are created as proliferat-

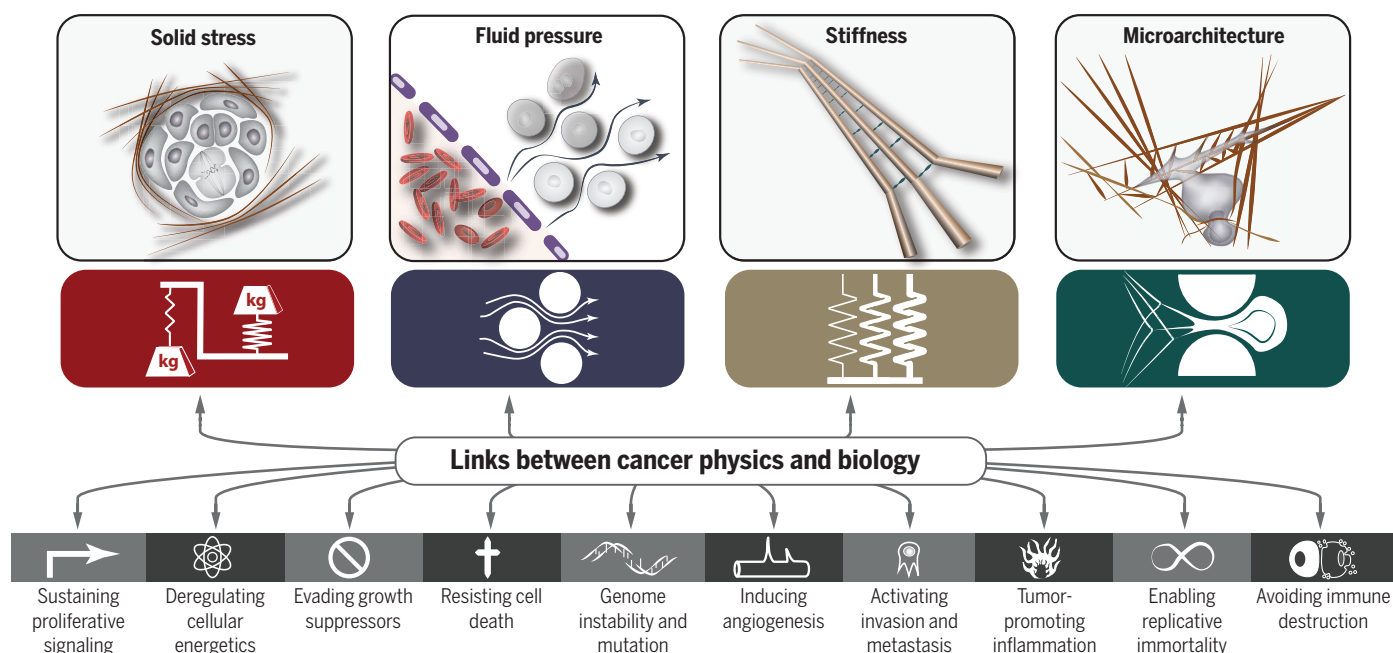
ing and migrating cells push and stretch solid components of the surrounding tissue. Being distinct from fluid pressure and close to zero in most normal tissues, solid stresses are large enough to compress blood and lymphatic vessels in and around tumors, impairing blood flow and the delivery of oxygen, drugs, and immune cells. Acting at organ, tissue, and cellular levels, solid stresses activate signaling pathways that promote tumorigenesis and invasiveness and induce treatment resistance. Elevated interstitial fluid pressure is caused by leakage of plasma from abnormally permeable tumor blood vessels and insufficient lymphatic drainage. As a result, the interstitial fluid leaks out of the tumor into the peritumor tissue, causing edema and elution of drugs and growth factors and facilitating invasion and metastasis through flow-induced shear stresses. Increased stiffness is caused by matrix deposition and remodeling. Traditionally used as a diagnostic marker, and more recently as a prognostic factor, increased stiffness activates signaling pathways that promote proliferation, invasiveness, and metastasis of cancer

cells. Finally, when normal tissue architecture is disrupted by cancer growth and invasion, microarchitecture is altered. Stromal and cancer cells and extracellular matrix adopt new organization. This changes the interactions between an individual cell and its surrounding matrix and cells, which affects signaling pathways associated with invasion and metastasis.

OUTLOOK: The tumor microenvironment is characterized by both biological and physical abnormalities. The growing appreciation of the role of tumor-stromal interactions in cancer has led to seminal discoveries that have resulted in previously unexplored targets and strategies for treatment. Understanding the key principles underlying the origins and consequences of the physical traits of cancer will be critical for improving treatment. Many of the concepts involved are nonintuitive and require deep and broad understanding of both the physical and biological aspects of cancer. Therefore, a rigorous but accessible description of physical cancer traits will assist research into the physical sciences of cancer—a highly multidisciplinary area—and help it remain an active and progressive subfield of cancer research. ■

The list of author affiliations is available in the full article online.
†Corresponding author. Email: jain@stele.mgh.harvard.edu
Cite this article as H. T. Nia et al., *Science* 370, eaaz0868 (2020). DOI: 10.1126/science.aaz0868

S READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.aaz0868>



Physical traits of cancer. To provide a comprehensive framework for understanding the links between the physics of cancer and signaling pathways in cancer biology in terms of a small number of underlying principles, we propose four physical traits of cancer that characterize the major physical abnormalities shared by most if not all tumors.

REVIEW

CANCER

Physical traits of cancer

Hadi T. Nia^{1,2}, Lance L. Munn¹, Rakesh K. Jain^{1,3*}

The role of the physical microenvironment in tumor development, progression, metastasis, and treatment is gaining appreciation. The emerging multidisciplinary field of the physical sciences of cancer is now embraced by engineers, physicists, cell biologists, developmental biologists, tumor biologists, and oncologists attempting to understand how physical parameters and processes affect cancer progression and treatment. Discoveries in this field are starting to be translated into new therapeutic strategies for cancer. In this Review, we propose four physical traits of tumors that contribute to tumor progression and treatment resistance: (i) elevated solid stresses (compression and tension), (ii) elevated interstitial fluid pressure, (iii) altered material properties (for example, increased tissue stiffness, which historically has been used to detect cancer by palpation), and (iv) altered physical microarchitecture. After defining these physical traits, we discuss their causes, consequences, and how they complement the biological hallmarks of cancer.

Cancer is generally considered a disease of the cell, caused by mutations in genes that control cell proliferation, death, metabolism, and DNA repair. To create a unified conceptual framework for understanding the various manifestations of cancer, Hanahan and Weinberg proposed eight biological hallmarks that delineate the key features and properties of cancer cells (1). These biological hallmarks are useful for conceptualizing cancer at the cellular level, but we now know that the microenvironment surrounding the cancer cell acts as a coconspirator in tumor initiation and progression. As tumors grow, they disrupt the surrounding tissue biochemically and physically. They also recruit normal cells from the surrounding tissue, which further alter the matrix and cellular compositions of the tumor. These perturbations result in physical abnormalities associated with both cancer cells and the microenvironment in which they grow that influence tumor biology and response to treatment (2).

To provide a more comprehensive conceptual framework for cancer, we propose four additional traits stemming from the physical abnormalities of tumors. These are (i) elevated solid stress, (ii) elevated interstitial fluid pressure (IFP) and the resulting fluid flow in the interstitium, (iii) increased stiffness and altered material properties, and (iv) altered microarchitecture (Fig. 1). As discussed in this Review, these four physical traits are conceptually distinct but can interact synergistically. They also enable and exacerbate many of the

biological hallmarks of cancer, thus facilitating cancer cell proliferation and invasion, immune system evasion, and resistance to therapies.

Solid stresses and elastic energy

Solid stresses, also known as residual stresses, are the mechanical forces (compressive, tensile, and shear) contained in—and transmitted by—solid and elastic elements of the extracellular matrix (ECM) and cells (3). Reported in pascals or millimeters of mercury (1 mmHg \cong 133.3 Pa), solid stress values range from <100 Pa (0.7 mmHg) in glioblastomas to 10,000 Pa (75 mmHg) in pancreatic ductal adenocarcinomas (PDACs) (4). Multiple mechanisms, summarized in Fig. 2 and discussed below, are responsible for generating solid stress in tumors.

1) Increased tissue volume caused by cell infiltration, cell proliferation, and matrix deposition. The added volume pushes and displaces existing viscoelastic structures inside and outside the tumor and gives rise to solid stresses in the tumor and the surrounding tissue (4, 5). As a result, when tumor cells are depleted through anticancer therapeutics, solid stress is decreased, and blood vessels are decompressed (6, 7).

2) Concerted displacement of normal tissue (8, 9). Some tumors grow as well-circumscribed, nodular masses, in which the tumor remains cohesive and pushes the surrounding tissue, generating considerable mechanical stresses. Other tumors are less cohesive and more infiltrative, interdigitating through the normal tissue by finding the path of least resistance or by creating space by virtue of cytotoxic and protease activities. In the latter case, there is less production of solid stress (5, 10).

3) Swelling of existing glycosaminoglycan matrix components such as hyaluronic acid (HA) owing to (electro)osmotic water absorption (11, 12). These components take up avail-

able water and swell, generating solid stress that is distinct from fluid pressure (12).

4) Actomyosin-mediated cell contractions. Fibroblasts, immune cells, and cancer cells can all contract matrix elements as they move around in a tumor or try to repair structural damage. Cell contraction generates tensile forces that contract ECM components (13), creating tension in some parts of the tumor, which are generally balanced by compression in other elements (4). Cancer-associated fibroblasts (CAFs) that are activated with transforming growth factor- β (TGF- β) become myofibroblasts and can generate large contractile forces (14).

The impact of solid stress on cancer cell biology was first recognized in 1997, when Helmlinger *et al.* found that accumulated solid stress inhibits the growth of tumor spheroids (3). These stresses are sufficiently large to compress and even collapse blood and lymphatic vessels (6, 7, 15). Vessel compression contributes to hypoxia (15, 16) and interferes with the delivery and/or efficacy of chemo-, radio-, and immunotherapies (17, 18). Solid stress may also have additional, direct effects on tumor biology, such as promoting the invasiveness of cancer cells (19) and stimulating tumorigenic pathways in colon epithelia (20) (Fig. 3).

Cancer and normal cells have mechanosensitive machinery, such as cell-ECM (21) and cell-cell (22) adhesions and stretch-sensitive ion channels (23), that allows them to respond to applied forces. Solid stresses can also act on cells indirectly by deforming ECM components. For example, matrix-bound latent TGF- β , inactive upon synthesis and unable to bind to its cognate cell surface receptors, can be activated by myofibroblast-induced tensile forces on ECM (24). Other examples of ECM sensitivity to tensile forces include the unfolding of fibronectin in response to tensile forces (25), the enzymatic resistance of collagen fibers (26), and the tension-regulated interactions of fibronectin with collagen fibers (27).

The nucleus is also a mechanosensitive organelle capable of responding to solid stresses through the activity of nuclear pore complexes and associated proteins, which modulate the nuclear import of transcription factors when the nucleus is deformed (28–30). Nuclear perturbations, such as those generated by cells migrating through small pores, cause changes in gene expression and induction of DNA repair programs (31, 32).

YAP (Yes-associated protein) and TAZ (transcriptional coactivator with PDZ-binding motif) have been identified as potent mechanoresponsive factors (33) that respond to physical cues, such as stretching (22, 34) and cell crowding (34), by translocating from the cytoplasm to the nucleus. YAP/TAZ mechanobiology is regulated by filamentous actin (F-actin) dynamics through Rho guanosine triphosphatases (GTPases), which probe the physical

¹Steele Laboratories, Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA. ²Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA.

³Ludwig Center at Harvard, Harvard Medical School, Boston, MA 02115, USA.

*Corresponding author. Email: jain@steele.mgh.harvard.edu

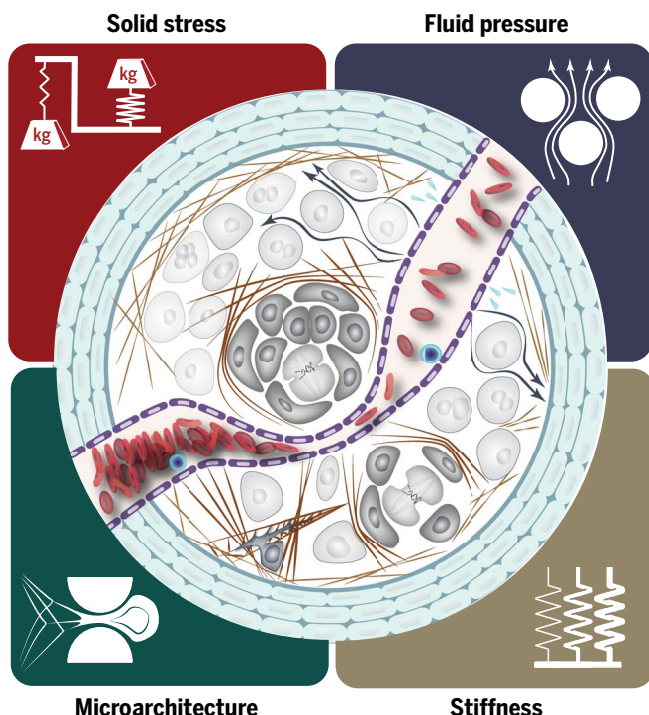


Fig. 1. Physical traits of cancer. On the basis of the advancements of the past few decades, we suggest that the physical traits of cancer can be categorized into four major groups: (i) elevated solid stress, (ii) elevated interstitial pressure, (iii) increased stiffness, and (iv) altered architecture and geometry. Solid stresses and fluid pressure are the mechanical stresses (force per unit area) contained in, and transmitted by, solid and fluid phases of the tumor, respectively. Solid stresses and fluid pressure are reported in pascals or millimeters of mercury ($1 \text{ mmHg} \approx 133.3 \text{ Pa}$). Stiffness (elasticity) is defined as the resistance of a material to deformation in response to an applied force, and elastic modulus is reported in pascals. Viscoelasticity defines the resistance of the material to deformation in response to a force applied at a given rate. Most soft tissues, including tumors, exhibit higher resistance to force (e.g., higher stiffness) when the force is applied at high rates. Solid stress, the latent or stored stress in a tissue, should not be confused with elasticity (stiffness) or viscoelasticity (time-dependent stiffness), which define how much or how fast, respectively, a tissue will deform if a force is applied. A tissue can be stiff (rigid) or soft (compliant), and, independently, it can be under compressive and/or tensile solid stresses (4) or, like most normal tissues, it can be unstressed. The proposed physical traits characterize most cancers, and their distinct origins and consequences make them indispensable to a comprehensive picture of cancer.

microenvironment via cell-matrix and cell-cell adhesion complexes at the cell surface (35). Thus, mechanosignaling by YAP and TAZ can be modulated by modifiers of actin and Rho GTPases such as cofilin, gelsolin, and F-actin-capping protein (CAP-Z). In a two-dimensional (2D) epithelial monolayer model of stretch, cells under tensile stresses showed activation of YAP/TAZ that led to cell proliferation (34) and induced cell cycle entry (22). Activation of the YAP/TAZ pathway contributes to tumor malignancy in many ways, including cell proliferation, cell cycle regulation, overcoming anoikis and mitochondria-induced apoptosis, inducing cancer stem cell functions, and accelerating fibrosis and desmoplasia by activating CAFs (36).

As cancer growth causes crowding of cells in the tissue, there is inevitably competition between cell populations for nutrients and free

space. Cell competition has recently received attention because of its relevance in development, where it controls organ size and eliminates suboptimal cells (37), and in tumorigenesis, where cancer cells expand into new space by damaging and killing the normal surrounding cells (38). Compressive and tensile stresses, generated from the differential growth of cell layers, has been suggested to be a mechanical driver of cell competition (39). How cancer cells are able to outcompete the surrounding normal cells, which experience similar solid stresses at the tumor-host interface, remains an open question.

Because solid stresses are harbored in matrix components, many of the resulting problems can be reversed by drugs that degrade matrix components and reduce fibrosis. For example, losartan, an angiotensin receptor 1 blocker, reduces both collagen I and HA by inhibiting

TGF- β (16). In preclinical models of PDAC, losartan alleviates solid stress and decompresses blood vessels, enhancing chemotherapy and increasing overall survival (16). This strategy is currently being tested in a randomized clinical trial (NCT01821729) based on promising results of a phase 2 trial (40). In another successful PDAC preclinical study, PEGPH20 (a pegylated recombinant human hyaluronidase) that reduced fibrosis in these tumors increased overall survival when combined with chemotherapy (41). Other approaches that have shown similar potential in preclinical models include inhibiting the vitamin D receptor (42), sonic hedgehog signaling (43), and C-X-C motif chemokine receptor 4 signaling (44). Targeting the vitamin D receptor is currently being tested in patients (NCT03472833). However, both PEGPH20 and sonic hedgehog targeting have been unsuccessful in clinical trials (45), highlighting the need for a deeper understanding of these pathways. Alleviating stress using these approaches may improve response to various treatments, such as immunotherapy (44, 46).

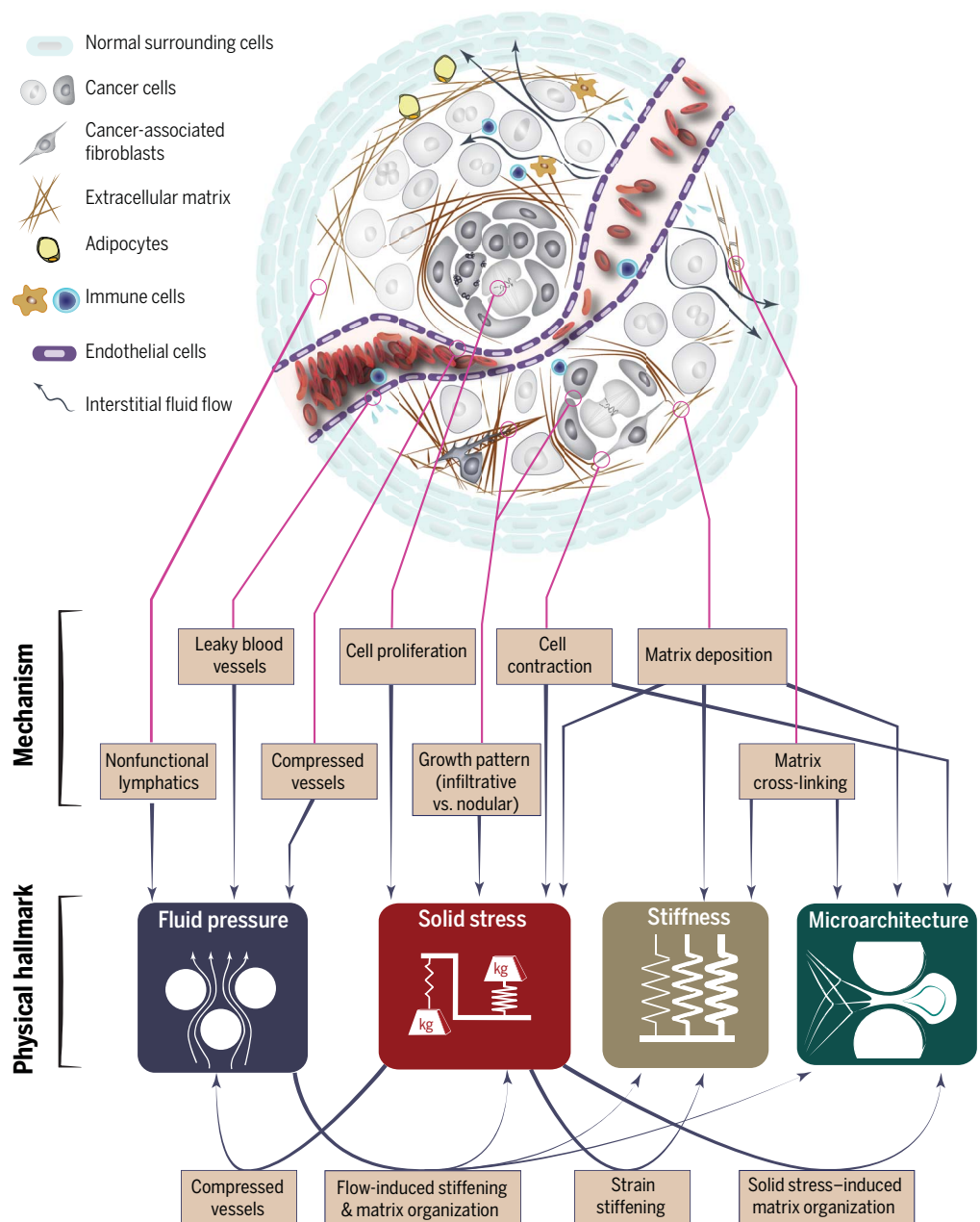
Interstitial fluid pressure

In most organs, the blood arrives via arteries and leaves via veins, and any excess tissue fluid is drained by lymphatic vessels. This maintains fluid homeostasis and results in near-zero IFP in most normal organs. This balance is disturbed by abnormalities in tumors, including hyperpermeable blood vessels and compression of blood and lymphatic vessels by solid stress. Leaky vessels, combined with a compromised drainage system, result in high IFP (Fig. 2), ranging from $<1 \text{ kPa}$ (7.5 mmHg) in brain tumors to 5 kPa (37 mmHg) in renal cell carcinomas. IFP is fairly uniform within a tumor and drops precipitously in the tumor margin, which generates a fluid flow toward lymphatic vessels in the surrounding normal tissue, where IFP is close to 0 mmHg . Note that IFP and solid stress are independent mechanical stresses with distinct origins and consequences (47).

High IFP in tumors was first reported in 1950 (48) and then later studied in detail through experiments and computational models (49, 50). Elevated fluid pressure drives interstitial flow in the tumor margin, exposing extravascular cells to shear stress. Because flow velocity and shear stresses depend strongly on the pore size between cells and matrix components, shear stresses likely vary widely, even along individual cell membranes. The shear stresses affect the biology of cancer and stromal cells in several ways (51) (Fig. 3), including activation of fibroblasts (51); modulation of endothelial sprouting (52), which affects angiogenesis and lymphangiogenesis (51); induction of matrix metalloproteinase (MMP) activity and cell motility (53); and activation of cancer cell migration (54) and invasion (55). Fluid flow has

Fig. 2. Origins of the physical traits of cancer.

Physical interactions of cancer cells with stroma give rise to physical traits of tumors through distinct and interconnected mechanisms. Leaky and compressed blood vessels and nonfunctional lymphatics lead to increased interstitial fluid pressure within the tumor and interstitial fluid flow in the tumor margin. Cellular proliferation, matrix deposition, cell contraction, and abnormal growth patterns lead to compressive and tensile solid stresses. Matrix deposition and cross-linking cause increased stiffness in tumors. Cell contraction, matrix deposition, and cross-linking also alter the architecture of the tissue. The physical traits also interact with each other; solid stresses compress blood and lymphatic vessels and contribute to increased fluid pressure in tumors. Tensile solid stresses result in stretched and aligned matrix, and through strain-stiffening, solid stresses also increase tumor stiffness. Fluid flow activates fibroblasts, which then contribute to increased solid stresses and stiffness values and alter ECM architecture.



also been shown to induce cell cycle arrest through integrin signaling (56). Because immune cells are also responsive to interstitial flow, these fluid forces are likely also involved in regulation of immunity (51). Mechanisms for mechanotransduction of flow signals include sensors within the focal adhesions (54, 57), the cell glycocalyx (55), cell-cell junctions (58, 59), ion channels (60), Notch receptor (61), and cilia (62). The resulting signals can up-regulate TGF- β expression and activate YAP/TAZ downstream pathways (51, 63, 64).

In addition to direct mechanotransduction mechanisms, fluid flow created by IFP gradients can affect tumor progression and treatment

response in multiple ways. High IFP hinders the convection of drugs from the vasculature into the bulk of the tumor (2, 49). Moreover, the steep IFP gradient at the tumor boundary drives the flow of interstitial fluid from the tumor toward the surrounding tissue. This flow can promote tumor invasion and growth by facilitating the transport of growth factors and cancer cells into the surrounding normal tissue and peritumor lymphatics (65). The outward fluid flow may also facilitate angiogenesis in the tumor margin (52) and remove therapeutic agents from the tumor, reducing drug retention times (65). IFP has also been proposed as a diagnostic marker differentiating malig-

nant from benign breast, head, and neck tumors (66, 67) and as a prognostic marker in some clinical studies (68).

Therapeutic strategies for correcting the fluid abnormalities in tumors have also been developed. One approach is to normalize the leaky and tortuous vasculature so that the intraluminal pressure operating within microvessels is not transmitted directly to the surrounding interstitium. Vascular normalization restores abnormal tumor vasculature to a more functional state closer to that of normal vessels. Using judicious doses of antiangiogenic therapy to normalize tumor vasculature (17, 65, 69), it is possible to increase pericyte coverage,

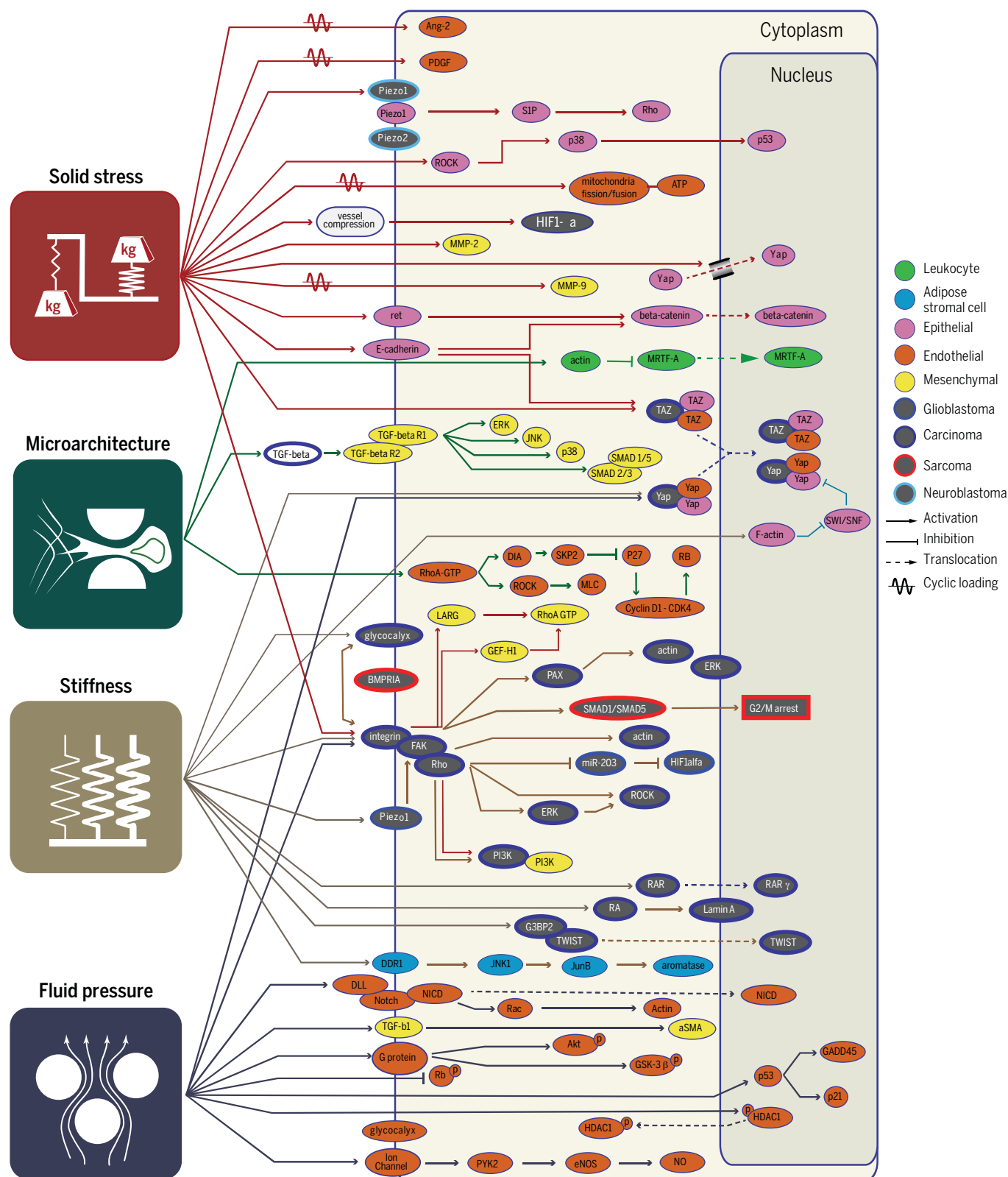


Fig. 3. Pathways associated with the physical traits of cancer. Physical traits of cancer activate a large cascade of mechanoresponsive pathways in cancer cells and stromal cells, including endothelial, epithelial, mesenchymal, and immune cells. Pathways such as integrin and YAP/TAZ are responsive to all four physical traits, whereas many other pathways appear to be more specific.

decrease vessel leakiness, increase tumor vascular perfusion, and decrease IFP. In the clinic, there are many agents with the ability to normalize vessels, including bevacizumab, an antibody that blocks vascular endothelial growth factor A (VEGF-A) and inhibitors of VEGF receptor tyrosine kinases, which have been approved for more than a dozen tumor types (70). As mentioned in the previous section, accumulation of solid stress also disrupts the vascular flow in tumors by compressing the more fragile outflow vessels (veins and lymphatics), which contributes to the elevated IFP. Therefore, alleviating solid stress can also decompress blood and lymphatic vessels, resulting in better perfusion and more normal levels of IFP (7).

Stiffness (elasticity)

Stiffness, also known as rigidity or elastic modulus, is defined as the resistance of a material to deformation in response to a force applied at a very slow rate (quasi-statically). Stiffness is an intrinsic material property of the tissue—unlike solid or fluid mechanical stresses, which describe forces exerted on a material—and ranges from 1 kPa in brain tumors to 70 kPa in cholangiocarcinomas (2).

Increased tissue stiffness is the most tangible and best-recognized mechanical abnormality in tumors. Stiffness has traditionally been used as a diagnostic marker (71) and more recently as a prognostic factor (72, 73). In multiple cancer types, including breast (74), pancreatic (75), liver (76), and prostate (77), malignant tumors have been shown to be considerably stiffer than benign tumors. In 2006, stiffness sensing was implicated in determining cell lineage (78). Today, there are numerous studies showing how the material properties—in particular, the stiffness of the microenvironment—are central to many traits of cancer (79), including proliferation (80, 81), angiogenesis (82), metabolism (83), invasion (84–86), and migration and metastasis (87–89) (Fig. 3).

Stiffening promotes tumor progression in many tumor types, including breast (80, 90), pancreatic (85, 91), colorectal (92), and brain (93). Increased stiffness can also promote an invasive phenotype in cancer cells (84–86), induce invasion and metastasis (87, 88, 94, 95), enhance immune cell infiltration (90), facilitate the epithelial-mesenchymal transition through TGF- β (96), promote stem cell differentiation (97), alter growth factor secretion and signaling, and increase angiogenesis and vessel permeability (82).

One of the primary causes of matrix stiffening is increased deposition and cross-linking of ECM (Fig. 2). In fibrotic tumors, CAFs are primarily responsible for collagen production, and these cells have more actin stress fibers, alpha smooth muscle actin, and focal adhesions than nonactivated fibroblasts (14, 98). Collagen

fibers can be cross-linked to different degrees by lysyl oxidase, and more cross-linking increases ECM stiffness. Transglutaminase 2, a calcium ion-dependent enzyme abundantly expressed by pancreatic cancer cells, also contributes to covalent collagen cross-linking and, consequently, activation of fibroblasts in pancreatic cancer (99). Increased ECM stiffness and TGF- β signaling activate fibroblasts to become CAFs, initiating a positive feedback loop that further enhances ECM stiffening. The profibrotic activation of cells in response to substrate stiffness can be perpetuated by mechanical memory, with microRNA 21 serving as one of the memory keepers (100).

Mechanical stresses can also alter the stiffness of the matrix through a phenomenon called strain-stiffening (101). Some collagen fibers are under tension, either because of cell contraction (102, 103) or because tumor growth causes local expansion, stretching the ECM (4). These tensile stresses increase the stiffness of the collagen network, which in turn further activates the focal adhesion contractility of the CAFs in their vicinity (104), leading to a vicious cycle of matrix deposition and stiffening. Matrix contraction by myofibroblasts is related to wound contraction, which resolves in wound healing but not in tumors (105, 106). In addition to stretching, the compressive stresses produced by tumor growth (5) can increase the stiffness of both normal and tumor tissue, as demonstrated with normal brain and glioma tissue (107). Strain-stiffening also happens at the subcellular level; mechanical stretch applied to nuclei can increase the stiffness of the nucleus through phosphorylation of emerin, one of the nuclear envelope proteins that provides structural stability to the nucleus (108).

There is considerable evidence that increased stiffness in breast tissue is associated with higher risk of breast cancer (72), and mammographic density (which is related to tissue stiffness and density) has been proposed as a predictor of poor survival (73, 109). In PDAC, increased stiffness negatively correlates with the response to chemotherapy (91). Consistently, in lung adenocarcinoma tumors, the stiffness of regions with dense ECM increases with tumor stage. Interestingly, however, the stiffness of the cells showed an inverse relationship with tumor stage (110), suggesting that cytoskeleton stiffness, in addition to ECM stiffness, can be used to stage lung tumors.

Two major pathways that are sensitive to changes in stiffness are focal adhesion kinase (FAK), which is induced through integrin ligation (111), and the Hippo pathway transcription factors YAP and TAZ (112, 113). YAP activated by increased ECM stiffness through Rho-associated protein kinase (ROCK), myosin, and Src activation further promotes CAF activity, which fuels a feed-forward self-enhancing loop to maintain the CAF phenotype (113).

Matrix stiffness and tumor cell metabolism are interdependent (83, 114). The cell metabolic rate increases when cells migrate on stiff substrates or through confined spaces (115). On the other hand, targeting abnormal tumor metabolism with metformin, an agonist of adenosine monophosphate-activated protein kinase, reduces fibrosis and stiffness (83, 116). Although increased tumor stiffness increases the malignancy of tumors, there may also be opportunities to take advantage of the increased stiffness by developing mechanosensitive treatments (117). Other targets for reversing fibrosis include the angiotensin system, which modulates CAF activity through the TGF- β and CTGF signaling pathways, contributing to fibrosis (16). ECM components can also be targeted directly, for example, by enzymatic depletion of HA and/or collagenase (2).

Matrix architecture and cell geometry

Organs are constructed of collections of cells and matrix components arranged with specific microarchitecture, which has evolved to optimize the stability, efficiency, and function of the tissue. For example, gut epithelium exists in a 2D sheet, with the basement membrane on one side and the luminal space on the other. This asymmetric arrangement, or polarization, allows the cell to respond to fluid forces on the luminal side, monitor the basement membrane and the underlying tissue through focal adhesions, and sense any changes in the neighboring cells through cadherin adhesions and gap junctions. Other cells require different microanatomy; for example, myocytes form aligned bundles in muscle tissue, and neurons exist as an interconnected linear network embedded in other tissues. These structures form during development, and in adult tissue, the immediate microenvironment (composition and geometry) of each cell serves as a cue for homeostasis or transformation and morphogenesis. For example, endothelial morphogenesis is triggered when blood flow stops and thrombosis fills the blood vessel lumen, resulting in vasculogenesis or angiogenesis. This process is, in part, induced by the lack of blood shear stresses and by the presence of the intravascular fibrin contacting the luminal side of the endothelial cells.

As tumors grow, both tumor and associated normal tissues are structurally disrupted in an ongoing, dynamic process that disturbs homeostasis. Cell overcrowding, protease activity, and changes in matrix production can all alter cell-matrix and cell-cell associations, signaling morphogenesis. Indeed, the much-studied epithelial-to-mesenchymal transition is an example of a morphogenic switch of cells from a 2D epithelial, surface-dependent geometry to a mesenchymal, infiltrative phenotype, where the cell is now comfortably surrounded by ECM.

The local tissue architecture plays a central role in cancer progression and treatment response, independent of the solid stress, fluid forces, and stiffness of the microenvironment (Fig. 3). A simple and familiar demonstration of the importance of architecture is surface-dependent growth. Normal and cancerous breast epithelial cells have drastically different morphology and growth rates when cultured in 3D matrices but are difficult to distinguish when cultured on 2D surfaces (118). This is an example of the dynamic reciprocity between tissue architecture, function, and neoplastic transformation, as cells not only create their environment but are also affected by it. Studies of 3D matrix architecture, mainly focused on collagen, show that collagen organization can be a prognostic biomarker (119) and that certain arrangements facilitate cancer cell migration, proliferation, and actomyosin contractility (120, 121). Studies designed to recapitulate various matrix architectures show that Rho/ROCK-mediated matrix alignment is a key step that promotes cancer cell migration in the early stages of invasion (122). Collagen alignment also modulates MMP-dependent mechanisms (123) and integrin $\beta 1$ expression (124), which affect cells' ability to migrate.

Many important discoveries regarding the influence of cell microenvironment on phenotype come from carefully designed *in vitro* studies. One of the earliest studies linking cell geometry to biological responses was reported by Folkman and Moscona in 1978 (125); they showed that DNA synthesis decreases as cell-substrate contact area is reduced. In another seminal study, Chen *et al.* were able to control growth and viability by confining cells to micropatterned islands, which controlled the extent of cell spreading while maintaining the total cell-substrate contact area (126). This model system has later been extended to micro- and nanopillar substrate systems (127) that allow specification of cell contact area in addition to substrate stiffness by varying pillar length, width, and spacing. Mechanistic studies using these model systems showed that cell proliferation is regulated by cell shape through two major mechanisms: (i) by regulating the cellular contractility through myosin light chain phosphorylation via ROCK (127, 128), and (ii) by phosphorylation of retinoblastoma protein (RB) (129). In addition to cell shape, there is recent evidence that cell volume can also affect cell response, including cell stiffness and stem cell fate, through water efflux (130). Surface features also affect the migration of cells, through a process called topotaxis (131).

Cell geometry also influences nuclear geometry, which can control gene expression (28, 29). Cells spread on a substrate have a more flattened nucleus than the same cells in 3D culture. Such shape dynamics and nuclear deforma-

tions affect perinuclear actin and microtubule networks (132), resulting in an altered arrangement of chromosomes, changes in gene expression, and YAP/TAZ nucleus translocation (34, 112). Cell and nuclear geometries are also altered during cell migration through constrictions. Migration of cancer cells (133–135), leukocytes (135), and primary mesenchymal stem cells (134) through pores smaller than their nucleus diameter results in severe compression and deformation of the nucleus, leading to loss of integrity of the nuclear envelope, herniation of chromatin via rupture through the nuclear membrane, and eventually DNA double-strand breaks (133–135), chromosomal aberration, and genomic instabilities (134).

The architecture of the local environment can also drastically affect migration (136). When cells are confined such that the plasma membrane experiences deformation, intercellular calcium ions are allowed to enter the cytoplasm through the stretch-activated ion channel Piezo1. This causes suppression of protein kinase A activity, which regulates migration of carcinoma cells via RhoA and Rac1 (137).

In addition to affecting cell migration, pore size and microarchitecture of the ECM determine the diffusion and convection of cytokines and therapeutic reagents. The key relevant parameters include pore size of vasculature for intravascular transport (138); pore size, charge, and orientation of ECM constituents (139, 140) for interstitial transport; and size, shape (e.g., spherical versus rod shape), and surface chemistry (e.g., cationic versus anionic) of the therapeutic reagents (141).

Outlook

The tumor microenvironment is aberrant both biologically and physically. The growing appreciation of the role of the physical microenvironment in cancer has led to several discoveries about the origins and consequences of the physical traits, which have resulted in new targets and treatment strategies in patients. Close collaboration between cancer biologists, clinicians, physical scientists, engineers, and data scientists will be required to ensure that research into the physical sciences of cancer—a highly multidisciplinary area—remains an active and progressive subfield of cancer research. Many of the concepts at play are nonintuitive and require rigorous and broad understanding of both the physical and biological aspects of cancer.

Continued growth of this subfield will require overcoming a number of challenges. As these proposed physical traits have received less research attention than their biological hallmark counterparts, the available tools for studying them are limited. Thus, more and improved *in vivo* and *in vitro* model systems are needed to recapitulate and study tumor physical abnormalities. Better model systems will aid in the discovery of solid stress-responsive

pathways and delineate the biological consequences of solid stress from other traits, for example, increased stiffness. Similarly, additional measurement tools are needed to distinguish different causes of solid stress. Delineating the contribution of different factors to the accumulation of solid stress, such as cell-cell and cell-matrix interactions, is an unmet need with potential for revealing additional therapeutic targets for reducing solid stress in tumors and normalizing the tumor physical microenvironment.

While malignant and benign tumors have been shown to differ in stiffness (74–77) and IFP (66), similar comparative evaluations of other physical traits, specifically solid stress and microarchitecture, are lacking. Furthermore, we know little about the origins of the differences in any of the physical traits of cancer in benign versus malignant tumors. Uncovering the potential mechanisms, such as infiltration of stroma, mutational loads, contractility of cancer and stromal cells, and collagen shell separating tumor stroma, should help researchers better understand the causes and consequences of physical cancer traits and their bidirectional links with the biological hallmarks of cancer.

Another promising area of research is the role that physical traits of cancer play in cancer cell biology at different stages of tumorigenesis, from early tumor formation to transformation, local invasion of basement membrane, and then dissemination and colonization at a distant site. The contribution and responsiveness of cancer cells to physical traits of cancer may vary at different stages of tumorigenesis and with different genetic aberrations. These factors may explain why certain transformed cells lose responsiveness to substrate stiffness (142), while other studies have shown that substrate stiffness promotes proliferation of cancer cells (80, 87).

Although it may appear that the physical traits of cancer discussed in this Review are specific to solid cancers, there is increasing evidence that they may also contribute to the progression and treatment response of hematological cancers. Phenomena such as swelling of lymph nodes, spleen, and even the liver subject both cancerous cells and normal immune cells to abnormal mechanical forces, which have yet to be studied in depth but may have important consequences on cell biology and anticancer immunity. Overcrowding of cells in the confined spaces of the vasculature and bone marrow results in a microenvironment with limited oxygen and nutrients and potentially important physical considerations that are currently unexplored. For example, in multiple myeloma (143), patients experience bone pain and spinal cord compression due to proliferation of cancer cells in confined spaces near nerves. The bone marrow niche—the origin

of most liquid and hematopoietic cancers—represents a distinctive mechanical environment consisting of viscoelastic tissue bathed in flowing fluid and surrounded by bone (144). The role of the physical properties of the bone marrow niche has recently gained attention: ECM stiffness was shown to alter the proliferation and treatment response of myeloid leukemia in an in vitro model (145). The physical properties of the bone marrow also determine drug delivery and the progression and invasion of liquid cancers (146). Finally, similar to carcinoma and sarcoma cells, hematologic cancer cells are also subjected to shear stress in systemic circulation.

Despite numerous studies on the role of physical cancer traits in the progression and initial treatment response of several tumor types, recurrence and secondary treatment resistance have not been studied in depth in association with the physical tumor microenvironment. However, there is early evidence linking the physical traits to the recurrence of cancer; in a mouse model of breast cancer, compliant tumors had higher rates of recurrence than stiffer counterparts (147). In a clinical study of 175 patients with hepatocellular carcinoma (HCC), the stiffness of the spleen was an independent predictor of tumor recurrence (148), and in another study, recurrence rates of HCC after thermal ablation correlated with tissue stiffness (149). These limited but promising studies highlight the need for more thorough investigations of the role of physical traits of cancer in recurrence and secondary treatment resistance. Finally, emerging data indicate that obesity increases the incidence of cancer, aids tumor progression, impairs treatment response, and facilitates tumor recurrence, but that physical exercise can ameliorate many of these adverse consequences of obesity. How obesity and physical exercise differentially affect the physical traits of cancer may reveal previously unexplored ways to slow tumor progression and improve treatment response (116, 150).

REFERENCES AND NOTES

1. D. Hanahan, R. A. Weinberg, Hallmarks of cancer: The next generation. *Cell* **144**, 646–674 (2011). doi: [10.1016/j.cell.2011.02.013](#); pmid: [21376230](#)
2. R. K. Jain, J. D. Martin, T. Stylianopoulos, The role of mechanical forces in tumor growth and therapy. *Annu. Rev. Biomed. Eng.* **16**, 321–346 (2014). doi: [10.1146/annurev-bioeng-071813-105259](#); pmid: [25014786](#)
3. G. Helminger, P. A. Netti, H. C. Lichtenfeld, R. J. Melder, R. K. Jain, Solid stress inhibits the growth of multicellular tumor spheroids. *Nat. Biotechnol.* **15**, 778–783 (1997). doi: [10.1038/nbt0897-778](#); pmid: [9255794](#)
4. H. T. Nia et al., Solid stress and elastic energy as measures of tumor mechanopathology. *Nat. Biomed. Eng.* **1**, 0004 (2016). doi: [10.1038/s41551-016-0004](#); pmid: [28966873](#)
5. G. Seano et al., Solid stress in brain tumours causes neuronal loss and neurological dysfunction and can be reversed by lithium. *Nat. Biomed. Eng.* **3**, 230–245 (2019). doi: [10.1038/s41551-018-0334-7](#); pmid: [30948807](#)
6. T. P. Padara et al., Cancer cells compress intratumour vessels. *Nature* **427**, 695 (2004). doi: [10.1038/427695a](#); pmid: [14973470](#)
7. G. Griffon-Etienne, Y. Boucher, C. Brekken, H. D. Suit, R. K. Jain, Taxane-induced apoptosis decompresses blood vessels and lowers interstitial fluid pressure in solid tumors: Clinical implications. *Cancer Res.* **59**, 3776–3782 (1999). pmid: [10446995](#)
8. C. P. Heisenberg, Y. Bellaiche, Forces in tissue morphogenesis and patterning. *Cell* **153**, 948–962 (2013). doi: [10.1016/j.cell.2013.05.008](#); pmid: [23706734](#)
9. K. D. Irvine, B. I. Shraiman, Mechanical control of growth: Ideas, facts and challenges. *Development* **144**, 4238–4248 (2017). doi: [10.1242/dev.151902](#); pmid: [29183937](#)
10. C. Fernández Moro, B. Bozóky, M. Gerling, Growth patterns of colorectal cancer liver metastases and their impact on prognosis: A systematic review. *BMJ Open Gastroenterol.* **5**, e000217 (2018). doi: [10.1136/bmjgast-2018-000217](#); pmid: [30073092](#)
11. C. Voutouri, C. Polydorou, P. Papageorgis, V. Gkretsi, T. Stylianopoulos, Hyaluronan-derived swelling of solid tumors, the contribution of collagen and cancer cells, and implications for cancer therapy. *Neoplasia* **18**, 732–741 (2016). doi: [10.1016/j.neo.2016.10.001](#); pmid: [27886639](#)
12. A. J. Grodzinsky, *Fields, Forces, and Flows in Biological Systems* (Garland Science, 2011), chap. 4, pp. 139–173.
13. D. D. Simon, C. O. Horgan, J. D. Humphrey, Mechanical restrictions on biological responses by adherent cells within collagen gels. *J. Mech. Behav. Biomed. Mater.* **14**, 216–226 (2012). doi: [10.1016/j.jmbbm.2012.05.009](#); pmid: [23022259](#)
14. E. Sahai et al., A framework for advancing our understanding of cancer-associated fibroblasts. *Nat. Rev. Cancer* **20**, 174–186 (2020). doi: [10.1038/s41568-019-0238-1](#); pmid: [31980749](#)
15. T. Stylianopoulos et al., Causes, consequences, and remedies for growth-induced solid stress in murine and human tumors. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 15101–15108 (2012). doi: [10.1073/pnas.1213353109](#); pmid: [22932871](#)
16. V. P. Chauhan et al., Angiotensin inhibition enhances drug delivery and potentiates chemotherapy by decompressing tumour blood vessels. *Nat. Commun.* **4**, 2516 (2013). doi: [10.1038/ncomms3516](#); pmid: [24084631](#)
17. R. K. Jain, Antiangiogenesis strategies revisited: From starving tumors to alleviating hypoxia. *Cancer Cell* **26**, 605–622 (2014). doi: [10.1016/j.ccr.2014.10.006](#); pmid: [25517747](#)
18. L. L. Munn, R. K. Jain, Vascular regulation of antitumor immunity. *Science* **365**, 544–545 (2019). doi: [10.1126/science.aaw7875](#); pmid: [31395771](#)
19. J. M. Tse et al., Mechanical compression drives cancer cells toward invasive phenotype. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 911–916 (2012). doi: [10.1073/pnas.1118910109](#); pmid: [22203958](#)
20. M. E. Fernández-Sánchez et al., Mechanical induction of the tumorigenic β -catenin pathway by tumour growth pressure. *Nature* **523**, 92–95 (2015). doi: [10.1038/nature14329](#); pmid: [25970250](#)
21. J. Z. Kechagia, J. Ivaska, P. Roca-Cusachs, Integrins as biomechanical sensors of the microenvironment. *Nat. Rev. Mol. Cell Biol.* **20**, 457–473 (2019). doi: [10.1038/s41580-019-0134-2](#); pmid: [31182865](#)
22. B. W. Benham-Pyle, B. L. Pruitt, W. J. Nelson, Mechanical strain induces E-cadherin-dependent Yap1 and β -catenin activation to drive cell cycle entry. *Science* **348**, 1024–1027 (2015). doi: [10.1126/science.aaa4559](#); pmid: [26023140](#)
23. B. Coste et al., Piezo1 and Piezo2 are essential components of distinct mechanically activated cation channels. *Science* **330**, 55–60 (2010). doi: [10.1126/science.1193270](#); pmid: [20813920](#)
24. P.-J. Wipff, D. B. Rifkin, J.-J. Meister, B. Hinz, Myofibroblast contraction activates latent TGF- β 1 from the extracellular matrix. *J. Cell Biol.* **179**, 1311–1323 (2007). doi: [10.1083/jcb.200704042](#); pmid: [18086923](#)
25. M. L. Smith et al., Force-induced unfolding of fibronectin in the extracellular matrix of living cells. *PLoS Biol.* **5**, e268 (2007). doi: [10.1371/journal.pbio.0050268](#); pmid: [17914904](#)
26. K. Saini, S. Cho, L. J. Dooling, D. E. Discher, Tension in fibrils suppresses their enzymatic degradation - A molecular mechanism for 'use it or lose it'. *Matrix Biol.* **85–86**, 34–46 (2020). doi: [10.1016/j.matbio.2019.06.001](#); pmid: [31201857](#)
27. K. E. Kubow et al., Mechanical forces regulate the interactions of fibronectin and collagen I in extracellular matrix. *Nat. Commun.* **6**, 8026 (2015). doi: [10.1038/ncomms9026](#); pmid: [26272817](#)
28. T. J. Kirby, J. Lammerding, Emerging views of the nucleus as a cellular mechanosensor. *Nat. Cell Biol.* **20**, 373–381 (2018). doi: [10.1038/s41556-018-0038-y](#); pmid: [29467443](#)
29. S. Cho, J. Irianto, D. E. Discher, Mechanosensing by the nucleus: From pathways to scaling relationships. *J. Cell Biol.* **216**, 305–315 (2017). doi: [10.1083/jcb.201610042](#); pmid: [28043971](#)
30. A. Elsegui-Artola et al., Force triggers YAP nuclear entry by regulating transport across nuclear pores. *Cell* **171**, 1397–1410.e14 (2017). doi: [10.1016/j.cell.2017.10.008](#); pmid: [29107331](#)
31. A. Tajik et al., Transcription upregulation via force-induced direct stretching of chromatin. *Nat. Mater.* **15**, 1287–1296 (2016). doi: [10.1038/nmat4729](#); pmid: [27548707](#)
32. A. Kumar et al., ATR mediates a checkpoint at the nuclear envelope in response to mechanical stress. *Cell* **158**, 633–646 (2014). doi: [10.1016/j.cell.2014.05.046](#); pmid: [25083873](#)
33. T. Panciera, L. Azzolin, M. Cordenonsi, S. Piccolo, Mechanobiology of YAP and TAZ in physiology and disease. *Nat. Rev. Mol. Cell Biol.* **18**, 758–770 (2017). doi: [10.1038/nrm.2017.87](#); pmid: [28951564](#)
34. M. Aragona et al., A mechanical checkpoint controls multicellular growth through YAP/TAZ regulation by actin-processing factors. *Cell* **154**, 1047–1059 (2013). doi: [10.1016/j.cell.2013.07.042](#); pmid: [23954413](#)
35. P. Gaspar, N. Tapon, Sensing the local environment: actin architecture and Hippo signalling. *Curr. Opin. Cell Biol.* **31**, 74–83 (2014). doi: [10.1016/j.cob.2014.09.003](#); pmid: [25259681](#)
36. F. Zanconato, M. Cordenonsi, S. Piccolo, YAP/TAZ at the roots of cancer. *Cancer Cell* **29**, 783–803 (2016). doi: [10.1016/j.ccr.2016.05.005](#); pmid: [27300434](#)
37. C. Clavería, G. Giovino, R. Sierra, M. Torres, Myc-driven endogenous cell competition in the early mammalian embryo. *Nature* **500**, 39–44 (2013). doi: [10.1038/nature12389](#); pmid: [23842495](#)
38. M. M. Merino, R. Levayer, E. Moreno, Survival of the fittest: Essential roles of cell competition in development, aging, and cancer. *Trends Cell Biol.* **26**, 776–788 (2016). doi: [10.1016/j.tcb.2016.05.009](#); pmid: [27319281](#)
39. B. I. Shraiman, Mechanical feedback as a possible regulator of tissue growth. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3318–3323 (2005). doi: [10.1073/pnas.0404782102](#); pmid: [15728365](#)
40. J. E. Murphy et al., Total neoadjuvant therapy with FOLFIRINOX in combination with losartan followed by chemoradiotherapy for locally advanced pancreatic cancer: A phase 2 clinical trial. *JAMA Oncol.* **5**, 1020–1027 (2019). doi: [10.1001/jamaoncol.2019.0892](#); pmid: [31145418](#)
41. P. P. Provenzano et al., Enzymatic targeting of the stroma ablates physical barriers to treatment of pancreatic ductal adenocarcinoma. *Cancer Cell* **21**, 418–429 (2012). doi: [10.1016/j.ccr.2012.01.007](#); pmid: [22439937](#)
42. M. H. Sherman et al., Vitamin D receptor-mediated stromal reprogramming suppresses pancreatitis and enhances pancreatic cancer therapy. *Cell* **159**, 80–93 (2014). doi: [10.1016/j.cell.2014.08.007](#); pmid: [25259922](#)
43. K. P. Olive et al., Inhibition of Hedgehog signaling enhances delivery of chemotherapy in a mouse model of pancreatic cancer. *Science* **324**, 1457–1461 (2009). doi: [10.1126/science.1171362](#); pmid: [19460966](#)
44. I. X. Chen et al., Blocking CXCR4 alleviates desmoplasia, increases T-lymphocyte infiltration, and improves immunotherapy in metastatic breast cancer. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 4558–4566 (2019). doi: [10.1073/pnas.1815515116](#); pmid: [30700545](#)
45. A. Wang-Gillam, Targeting stroma: A tale of caution. *J. Clin. Oncol.* **37**, 1041–1043 (2019). doi: [10.1200/JCO.19.00056](#); pmid: [30860950](#)
46. V. P. Chauhan et al., Reprogramming the microenvironment with tumor-selective angiotensin blockers enhances cancer immunotherapy. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 10674–10680 (2019). doi: [10.1073/pnas.1819889116](#); pmid: [31042028](#)
47. V. P. Chauhan et al., Compression of pancreatic tumor blood vessels by hyaluronan is caused by solid stress and not interstitial fluid pressure. *Cancer Cell* **26**, 14–15 (2014). doi: [10.1016/j.ccr.2014.06.003](#); pmid: [25026209](#)
48. J. S. Young, C. E. Lumsden, A. L. Stalker, The significance of the tissue pressure of normal testicular and of neoplastic (Brown-Pearce carcinoma) tissue in the rabbit. *J. Pathol. Bacteriol.* **62**, 313–333 (1950). doi: [10.1002/path.1700620303](#); pmid: [14784896](#)
49. R. K. Jain, L. T. Baxter, Mechanisms of heterogeneous distribution of monoclonal antibodies and other macromolecules in tumors: Significance of elevated interstitial pressure. *Cancer Res.* **48**, 7022–7032 (1988). pmid: [3191477](#)

50. Y. Boucher, L. T. Baxter, R. K. Jain, Interstitial pressure gradients in tissue-isolated and subcutaneous tumors: Implications for therapy. *Cancer Res.* **50**, 4478–4484 (1990). pmid: [2369726](#)
51. M. A. Swartz, A. W. Lund, Lymphatic and interstitial flow in the tumour microenvironment: Linking mechanobiology with immunity. *Nat. Rev. Cancer* **12**, 210–219 (2012). doi: [10.1038/nrc3186](#); pmid: [22362216](#)
52. J. W. Song, L. L. Munn, Fluid forces control endothelial sprouting. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15342–15347 (2011). doi: [10.1073/pnas.1105316108](#); pmid: [21876168](#)
53. H. Qazi, Z. D. Shi, J. M. Tarbell, Fluid shear stress regulates the invasive potential of glioma cells via modulation of migratory activity and matrix metalloproteinase expression. *PLOS ONE* **6**, e20348 (2011). doi: [10.1371/journal.pone.0020348](#); pmid: [21637818](#)
54. W. J. Polacheck, J. L. Charest, R. D. Kamm, Interstitial flow influences direction of tumor cell migration through competing mechanisms. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11115–11120 (2011). doi: [10.1073/pnas.1103581108](#); pmid: [21690404](#)
55. H. Qazi, R. Palomino, Z. D. Shi, L. L. Munn, J. M. Tarbell, Cancer cell glyocalyx mediates mechanotransduction and flow-regulated invasion. *Integr. Biol.* **5**, 1334–1343 (2013). doi: [10.1039/c3ib40057c](#); pmid: [24077103](#)
56. S. F. Chang et al., Tumor cell cycle arrest induced by shear stress: Roles of integrins and Smad. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 3927–3932 (2008). doi: [10.1073/pnas.0712353105](#); pmid: [18310319](#)
57. W. J. Polacheck, A. E. German, A. Mammoto, D. E. Ingber, R. D. Kamm, Mechanotransduction of fluid stresses governs 3D cell migration. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2447–2452 (2014). doi: [10.1073/pnas.1316848111](#); pmid: [24550267](#)
58. T. Lecuit, A. S. Yap, E-cadherin junctions as active mechanical integrators in tissue dynamics. *Nat. Cell Biol.* **17**, 533–539 (2015). doi: [10.1038/ncb3136](#); pmid: [25925582](#)
59. D. E. Leckband, J. de Rooij, Cadherin adhesion and mechanotransduction. *Annu. Rev. Cell Dev. Biol.* **30**, 291–315 (2014). doi: [10.1146/annurev-cellbio-100913-013212](#); pmid: [25062360](#)
60. K. Yamamoto et al., Impaired flow-dependent control of vascular tone and remodeling in P2X4-deficient mice. *Nat. Med.* **12**, 133–137 (2006). doi: [10.1038/nm1338](#); pmid: [16327800](#)
61. W. J. Polacheck et al., A non-canonical Notch complex regulates adherens junctions and vascular barrier function. *Nature* **552**, 258–262 (2017). doi: [10.1038/nature24998](#); pmid: [29160307](#)
62. R. Toftgård, Two sides to cilia in cancer. *Nat. Med.* **15**, 994–996 (2009). doi: [10.1038/nm0909-994](#); pmid: [19734870](#)
63. K. C. Wang et al., Flow-dependent YAP/TAZ activities regulate endothelial phenotypes and atherosclerosis. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 11525–11530 (2016). doi: [10.1073/pnas.1613121113](#); pmid: [27671657](#)
64. L. Wang et al., Integrin-YAP/TAZ-JNK cascade mediates atheroprotective effect of unidirectional shear flow. *Nature* **540**, 579–582 (2016). doi: [10.1038/nature20602](#); pmid: [27926730](#)
65. R. K. Jain, Normalizing tumor microenvironment to treat cancer: Bench to bedside to biomarkers. *J. Clin. Oncol.* **31**, 2205–2218 (2013). doi: [10.1200/JCO.2012.46.3653](#); pmid: [23669226](#)
66. S. D. Nathanson, L. Nelson, Interstitial fluid pressure in breast cancer, benign breast conditions, and breast parenchyma. *Ann. Surg. Oncol.* **1**, 333–338 (1994). doi: [10.1007/BF03187139](#); pmid: [7850532](#)
67. R. Gutmann et al., Interstitial hypertension in head and neck tumors in patients: Correlation with tumor size. *Cancer Res.* **52**, 1993–1995 (1992). pmid: [1551128](#)
68. A. Fyles et al., Long-term performance of interstitial fluid pressure and hypoxia as prognostic factors in cervix cancer. *Radiother. Oncol.* **80**, 132–137 (2006). doi: [10.1016/j.radonc.2006.07.014](#); pmid: [16920212](#)
69. R. K. Jain, Normalization of tumor vasculature: An emerging concept in antiangiogenic therapy. *Science* **307**, 58–62 (2005). doi: [10.1126/science.1104819](#); pmid: [15637262](#)
70. D. Fukumura, J. Kloepper, Z. Amoozgar, D. G. Duda, R. K. Jain, Enhancing cancer immunotherapy using antiangiogenics: Opportunities and challenges. *Nat. Rev. Clin. Oncol.* **15**, 325–340 (2018). doi: [10.1038/nrclinonc.2018.29](#); pmid: [29508855](#)
71. D. L. Cochlin, R. H. Ganatra, D. F. Griffiths, Elastography in the detection of prostatic cancer. *Clin. Radiol.* **57**, 1014–1020 (2002). doi: [10.1053/crad.2002.0989](#); pmid: [12409113](#)
72. N. F. Boyd et al., Evidence that breast tissue stiffness is associated with risk of breast cancer. *PLOS ONE* **9**, e100937 (2014). doi: [10.1371/journal.pone.0100937](#); pmid: [25010427](#)
73. G. Maskarinec et al., Mammographic density as a predictor of breast cancer survival: The Multiethnic Cohort. *Breast Cancer Res.* **15**, R7 (2013). doi: [10.1186/bcr3378](#); pmid: [23339436](#)
74. A. Evans et al., Differentiating benign from malignant solid breast masses: Value of shear wave elastography according to lesion stiffness combined with greyscale ultrasound according to BI-RADS classification. *Br. J. Cancer* **107**, 224–229 (2012). doi: [10.1038/bjc.2012.253](#); pmid: [22691969](#)
75. S. Carrara et al., EUS elastography (strain ratio) and fractal-based quantitative analysis for the diagnosis of solid pancreatic lesions. *Gastrointest. Endosc.* **87**, 1464–1473 (2018). doi: [10.1016/j.gie.2017.12.031](#); pmid: [29329992](#)
76. M. Shahryari et al., Tomoelastography distinguishes noninvasively between benign and malignant liver lesions. *Cancer Res.* **79**, 5704–5710 (2019). doi: [10.1158/0008-5472.CAN-19-2150](#); pmid: [31551364](#)
77. O. Rouvière et al., Stiffness of benign and malignant prostate tissue measured by shear-wave elastography: A preliminary study. *Eur. Radiol.* **27**, 1858–1866 (2017). doi: [10.1007/s00330-016-4534-9](#); pmid: [27553936](#)
78. A. J. Engler, S. Sen, H. L. Sweeney, D. E. Discher, Matrix elasticity directs stem cell lineage specification. *Cell* **126**, 677–689 (2006). doi: [10.1016/j.cell.2006.06.044](#); pmid: [16923388](#)
79. M. W. Pickup, J. K. Mouw, V. M. Weaver, The extracellular matrix modulates the hallmarks of cancer. *EMBO Rep.* **15**, 1243–1253 (2014). doi: [10.15252/embr.201439246](#); pmid: [25381661](#)
80. M. J. Paszek et al., Tensional homeostasis and the malignant phenotype. *Cancer Cell* **8**, 241–254 (2005). doi: [10.1016/j.ccr.2005.08.010](#); pmid: [16169468](#)
81. T. A. Ulrich, E. M. de Juan Pardo, S. Kumar, The mechanical rigidity of the extracellular matrix regulates the structure, motility, and proliferation of glioma cells. *Cancer Res.* **69**, 4167–4174 (2009). doi: [10.1158/0008-5472.CAN-08-4859](#); pmid: [19435897](#)
82. F. Bordeleau et al., Matrix stiffening promotes a tumor vasculature phenotype. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 492–497 (2017). doi: [10.1073/pnas.1613855114](#); pmid: [28034921](#)
83. J. C. Tung et al., Tumor mechanics and metabolic dysfunction. *Free Radic. Biol. Med.* **79**, 269–280 (2015). doi: [10.1016/j.freeradbiomed.2014.11.020](#); pmid: [25532934](#)
84. K. R. Levental et al., Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell* **139**, 891–906 (2009). doi: [10.1016/j.cell.2009.10.027](#); pmid: [19931152](#)
85. B. W. Miller et al., Targeting the LOX/hypoxia axis reverses many of the features that make pancreatic cancer deadly: Inhibition of LOX abrogates metastasis and enhances drug efficacy. *EMBO Mol. Med.* **7**, 1063–1076 (2015). doi: [10.15252/emmm.201404827](#); pmid: [26077591](#)
86. N. R. Lang et al., Biphasic response of cell invasion to matrix stiffness in three-dimensional biopolymer networks. *Acta Biomater.* **13**, 61–67 (2015). doi: [10.1016/j.actbio.2014.11.003](#); pmid: [25462839](#)
87. D. Wirtz, K. Konstantopoulos, P. C. Searson, The physics of cancer: The role of physical interactions and mechanical forces in metastasis. *Nat. Rev. Cancer* **11**, 512–522 (2011). doi: [10.1038/nrc3080](#); pmid: [21701513](#)
88. A. Pathak, S. Kumar, Independent regulation of tumor cell migration by matrix stiffness and confinement. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 10334–10339 (2012). doi: [10.1073/pnas.1118073109](#); pmid: [22689955](#)
89. G. Charras, E. Sahai, Physical influences of the extracellular environment on cell migration. *Nat. Rev. Mol. Cell Biol.* **15**, 813–824 (2014). doi: [10.1038/nrm3897](#); pmid: [25355506](#)
90. I. Acerbi et al., Human breast cancer invasion and aggression correlates with ECM stiffening and immune cell infiltration. *Integr. Biol.* **7**, 1120–1134 (2015). doi: [10.1039/c5ib00040h](#); pmid: [25959051](#)
91. R. Alvarez et al., Stromal disrupting effects of nab-paclitaxel in pancreatic cancer. *Br. J. Cancer* **109**, 926–933 (2013). doi: [10.1038/bjc.2013.415](#); pmid: [23907428](#)
92. N. N. Rahbari et al., Anti-VEGF therapy induces ECM remodeling and mechanical barriers to therapy in colorectal cancer liver metastases. *Sci. Transl. Med.* **8**, 360ra135 (2016). doi: [10.1126/scitranslmed.aaf5219](#); pmid: [27733559](#)
93. Y. A. Miroshnikova et al., Tissue mechanics promote IDH1-dependent HIF1 α -tenascin C feedback to regulate glioblastoma aggression. *Nat. Cell Biol.* **18**, 1336–1345 (2016). doi: [10.1038/ncb3429](#); pmid: [27820599](#)
94. B. C. Isenberg, P. A. Dimilla, M. Walker, S. Kim, J. Y. Wong, Vascular smooth muscle cell durotaxis depends on substrate stiffness gradient strength. *Biophys. J.* **97**, 1313–1322 (2009). doi: [10.1016/j.bpj.2009.06.021](#); pmid: [19720019](#)
95. D. M. Gilkes, G. L. Semenza, D. Wirtz, Hypoxia and the extracellular matrix: Drivers of tumour metastasis. *Nat. Rev. Cancer* **14**, 430–439 (2014). doi: [10.1038/nrc3726](#); pmid: [24827502](#)
96. J. L. Leight, M. A. Wozniak, S. Chen, M. L. Lynch, C. S. Chen, Matrix rigidity regulates a switch between TGF- β -induced apoptosis and epithelial-mesenchymal transition. *Mol. Biol. Cell* **23**, 781–791 (2012). doi: [10.1091/mbc.e11-06-0537](#); pmid: [22238361](#)
97. M.-F. Pang et al., Tissue stiffness and hypoxia modulate the integrin-linked kinase ILK to control breast cancer stem-like cells. *Cancer Res.* **76**, 5277–5287 (2016). doi: [10.1158/0008-5472.CAN-16-0579](#); pmid: [27503933](#)
98. D. Öhlund, E. Elyada, D. Tuveson, Fibroblast heterogeneity in the cancer wound. *J. Exp. Med.* **211**, 1503–1523 (2014). doi: [10.1084/jem.20140692](#); pmid: [25071162](#)
99. J. Lee et al., Tissue transglutaminase mediated tumor-stroma interaction promotes pancreatic cancer progression. *Clin. Cancer Res.* **21**, 4482–4493 (2015). doi: [10.1158/1078-0432.CCR-15-0226](#); pmid: [26041746](#)
100. C. X. Li et al., MicroRNA-21 preserves the fibrotic mechanical memory of mesenchymal stem cells. *Nat. Mater.* **16**, 379–389 (2017). doi: [10.1038/nmat4780](#); pmid: [27798620](#)
101. C. Storm, J. J. Pastore, F. C. MacKintosh, T. C. Lubensky, P. A. Janmey, Nonlinear elasticity in biological gels. *Nature* **435**, 191–194 (2005). doi: [10.1038/nature03521](#); pmid: [15889088](#)
102. M. S. Samuel et al., Actomyosin-mediated cellular tension drives increased tissue stiffness and β -catenin activation to induce epidermal hyperplasia and tumor growth. *Cancer Cell* **19**, 776–791 (2011). doi: [10.1016/j.ccr.2011.05.008](#); pmid: [21665151](#)
103. Y. L. Han et al., Cell contraction induces long-ranged stress stiffening in the extracellular matrix. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4075–4080 (2018). doi: [10.1073/pnas.1722619115](#); pmid: [29618614](#)
104. D. Choquet, D. P. Felsenfeld, M. P. Sheetz, Extracellular matrix rigidity causes strengthening of integrin-cytoskeleton linkages. *Cell* **88**, 39–48 (1997). doi: [10.1016/S0092-8674\(00\)81856-5](#); pmid: [9019403](#)
105. L. L. Munn, Cancer and inflammation. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **9**, e1370 (2017). doi: [10.1002/wsbm.1370](#); pmid: [27943646](#)
106. F. R. Giretti, S. I. Grivennikov, Inflammation and cancer: Triggers, mechanisms, and consequences. *Immunity* **51**, 27–41 (2019). doi: [10.1016/j.immuni.2019.06.025](#); pmid: [31315034](#)
107. K. Pogoda et al., Compression stiffening of brain and its effect on mechanosensing by glioma cells. *New J. Phys.* **16**, 075002 (2014). doi: [10.1088/1367-2630/16/7/075002](#); pmid: [25844043](#)
108. C. Guilluy et al., Isolated nuclei adapt to force and reveal a mechanotransduction pathway in the nucleus. *Nat. Cell Biol.* **16**, 376–381 (2014). doi: [10.1038/ncb2927](#); pmid: [24609268](#)
109. M. F. Bakker et al., Supplemental MRI screening for women with extremely dense breast tissue. *N. Engl. J. Med.* **381**, 2091–2102 (2019). doi: [10.1056/NEJMoa1903986](#); pmid: [31774954](#)
110. V. Panzetta et al., Mechanical phenotyping of cells and extracellular matrix as grade and stage markers of lung tumor tissues. *Acta Biomater.* **57**, 334–341 (2017). doi: [10.1016/j.actbio.2017.05.002](#); pmid: [28483699](#)
111. F. J. Sulzmaier, C. Jean, D. D. Schlaepfer, FAK in cancer: Mechanistic findings and clinical applications. *Nat. Rev. Cancer* **14**, 598–610 (2014). doi: [10.1038/nrc3792](#); pmid: [25098269](#)
112. S. Dupont et al., Role of YAP/TAZ in mechanotransduction. *Nature* **474**, 179–183 (2011). doi: [10.1038/nature10137](#); pmid: [21654799](#)
113. F. Calvo et al., Mechanotransduction and YAP-dependent matrix remodelling is required for the generation and

- maintenance of cancer-associated fibroblasts. *Nat. Cell Biol.* **15**, 637–646 (2013). doi: [10.1038/ncb2756](https://doi.org/10.1038/ncb2756); pmid: [23708000](https://pubmed.ncbi.nlm.nih.gov/23708000/)
114. P. DelNero, B. D. Hopkins, L. C. Cantley, C. Fischbach, Cancer metabolism gets physical. *Sci. Transl. Med.* **10**, eaag1011 (2018). doi: [10.1126/scitranslmed.aag1011](https://doi.org/10.1126/scitranslmed.aag1011); pmid: [29794058](https://pubmed.ncbi.nlm.nih.gov/29794058/)
 115. M. R. Zanotelli *et al.*, Energetic costs regulated by cell mechanics and confinement are predictive of migration path during decision-making. *Nat. Commun.* **10**, 4185 (2019). doi: [10.1038/s41467-019-12155-z](https://doi.org/10.1038/s41467-019-12155-z); pmid: [31519914](https://pubmed.ncbi.nlm.nih.gov/31519914/)
 116. J. Incio *et al.*, Metformin reduces desmoplasia in pancreatic cancer by reprogramming stellate cells and tumor-associated macrophages. *PLOS ONE* **10**, e0141392 (2015). doi: [10.1371/journal.pone.0141392](https://doi.org/10.1371/journal.pone.0141392); pmid: [26641266](https://pubmed.ncbi.nlm.nih.gov/26641266/)
 117. L. Liu *et al.*, Mechanoresponsive stem cells to target cancer metastases through biophysical cues. *Sci. Transl. Med.* **9**, ean2966 (2017). doi: [10.1126/scitranslmed.aan2966](https://doi.org/10.1126/scitranslmed.aan2966); pmid: [28747514](https://pubmed.ncbi.nlm.nih.gov/28747514/)
 118. B. M. Baker, C. S. Chen, Deconstructing the third dimension: How 3D culture microenvironments alter cellular cues. *J. Cell Sci.* **125**, 3015–3024 (2012). doi: [10.1242/jcs.079509](https://doi.org/10.1242/jcs.079509); pmid: [22797912](https://pubmed.ncbi.nlm.nih.gov/22797912/)
 119. M. W. Conklin *et al.*, Aligned collagen is a prognostic signature for survival in human breast carcinoma. *Am. J. Pathol.* **178**, 1221–1232 (2011). doi: [10.1016/j.ajpath.2010.11.076](https://doi.org/10.1016/j.ajpath.2010.11.076); pmid: [21356373](https://pubmed.ncbi.nlm.nih.gov/21356373/)
 120. P. P. Provenzano *et al.*, Collagen reorganization at the tumor-stromal interface facilitates local invasion. *BMC Med.* **4**, 38 (2006). doi: [10.1186/1741-7015-4-38](https://doi.org/10.1186/1741-7015-4-38); pmid: [17190588](https://pubmed.ncbi.nlm.nih.gov/17190588/)
 121. P. K. Chaudhuri, C. Q. Pan, B. C. Low, C. T. Lim, Topography induces differential sensitivity on cancer cell proliferation via Rho-ROCK-Myosin contractility. *Sci. Rep.* **6**, 19672 (2016). doi: [10.1038/srep19672](https://doi.org/10.1038/srep19672); pmid: [26795068](https://pubmed.ncbi.nlm.nih.gov/26795068/)
 122. P. P. Provenzano, D. R. Inman, K. W. Eliceiri, S. M. Trier, P. J. Keely, Contact guidance mediated three-dimensional cell migration is regulated by Rho/ROCK-dependent matrix reorganization. *Biophys. J.* **95**, 5374–5384 (2008). doi: [10.1529/biophysj.108.133116](https://doi.org/10.1529/biophysj.108.133116); pmid: [18775961](https://pubmed.ncbi.nlm.nih.gov/18775961/)
 123. S. I. Fraley *et al.*, Three-dimensional matrix fiber alignment modulates cell migration and MT1-MMP utility by spatially and temporally directing protrusions. *Sci. Rep.* **5**, 14580 (2015). doi: [10.1038/srep14580](https://doi.org/10.1038/srep14580); pmid: [26423227](https://pubmed.ncbi.nlm.nih.gov/26423227/)
 124. D. O. Velez *et al.*, 3D collagen architecture induces a conserved migratory and transcriptional response linked to vasculogenic mimicry. *Nat. Commun.* **8**, 1651 (2017). doi: [10.1038/s41467-017-01556-7](https://doi.org/10.1038/s41467-017-01556-7); pmid: [29162797](https://pubmed.ncbi.nlm.nih.gov/29162797/)
 125. J. Folkman, A. Moscona, Role of cell shape in growth control. *Nature* **273**, 345–349 (1978). doi: [10.1038/273345a0](https://doi.org/10.1038/273345a0); pmid: [661946](https://pubmed.ncbi.nlm.nih.gov/661946/)
 126. C. S. Chen, M. Mrksich, S. Huang, G. M. Whitesides, D. E. Ingber, Geometric control of cell life and death. *Science* **276**, 1425–1428 (1997). doi: [10.1126/science.276.5317.1425](https://doi.org/10.1126/science.276.5317.1425); pmid: [9162012](https://pubmed.ncbi.nlm.nih.gov/9162012/)
 127. J. L. Tan *et al.*, Cells lying on a bed of microneedles: An approach to isolate mechanical force. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1484–1489 (2003). doi: [10.1073/pnas.0235407100](https://doi.org/10.1073/pnas.0235407100); pmid: [12552122](https://pubmed.ncbi.nlm.nih.gov/12552122/)
 128. D. M. Pirone *et al.*, An inhibitory role for FAK in regulating proliferation: A link between limited adhesion and RhoA-ROCK signaling. *J. Cell Biol.* **174**, 277–288 (2006). doi: [10.1083/jcb.200510062](https://doi.org/10.1083/jcb.200510062); pmid: [16847103](https://pubmed.ncbi.nlm.nih.gov/16847103/)
 129. S. Huang, C. S. Chen, D. E. Ingber, Control of cyclin D1, p27^{Kip1}, and cell cycle progression in human capillary endothelial cells by cell shape and cytoskeletal tension. *Mol. Biol. Cell* **9**, 3179–3193 (1998). doi: [10.1091/mbc.9.11.3179](https://doi.org/10.1091/mbc.9.11.3179); pmid: [9802905](https://pubmed.ncbi.nlm.nih.gov/9802905/)
 130. M. Guo *et al.*, Cell volume change through water efflux impacts cell stiffness and stem cell fate. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E8618–E8627 (2017). doi: [10.1073/pnas.1705179114](https://doi.org/10.1073/pnas.1705179114); pmid: [28973866](https://pubmed.ncbi.nlm.nih.gov/28973866/)
 131. J. Park *et al.*, Directed migration of cancer cells guided by the graded texture of the underlying matrix. *Nat. Mater.* **15**, 792–801 (2016). doi: [10.1038/nmat4586](https://doi.org/10.1038/nmat4586); pmid: [26974411](https://pubmed.ncbi.nlm.nih.gov/26974411/)
 132. N. M. Ramdas, G. V. Shivashankar, Cytoskeletal control of nuclear morphology and chromatin organization. *J. Mol. Biol.* **427**, 695–706 (2015). doi: [10.1016/j.jmb.2014.09.008](https://doi.org/10.1016/j.jmb.2014.09.008); pmid: [25281900](https://pubmed.ncbi.nlm.nih.gov/25281900/)
 133. C. M. Denais *et al.*, Nuclear envelope rupture and repair during cancer cell migration. *Science* **352**, 353–358 (2016). doi: [10.1126/science.aad7297](https://doi.org/10.1126/science.aad7297); pmid: [27013428](https://pubmed.ncbi.nlm.nih.gov/27013428/)
 134. J. Irianto *et al.*, DNA damage follows repair factor depletion and portends genome variation in cancer cells after pore migration. *Curr. Biol.* **27**, 210–223 (2017). doi: [10.1016/j.cub.2016.11.049](https://doi.org/10.1016/j.cub.2016.11.049); pmid: [27989676](https://pubmed.ncbi.nlm.nih.gov/27989676/)
 135. M. Raab *et al.*, ESCRT III repairs nuclear envelope ruptures during cell migration to limit DNA damage and cell death. *Science* **352**, 359–362 (2016). doi: [10.1126/science.aad7611](https://doi.org/10.1126/science.aad7611); pmid: [27013426](https://pubmed.ncbi.nlm.nih.gov/27013426/)
 136. C. D. Paul, P. Mistriotis, K. Konstantopoulos, Cancer cell motility: Lessons from migration in confined spaces. *Nat. Rev. Cancer* **17**, 131–140 (2017). doi: [10.1038/nrc.2016.123](https://doi.org/10.1038/nrc.2016.123); pmid: [27909339](https://pubmed.ncbi.nlm.nih.gov/27909339/)
 137. W. C. Hung *et al.*, Confinement sensing and signal optimization via Piezo1/PKA and myosin II pathways. *Cell Rep.* **15**, 1430–1441 (2016). doi: [10.1016/j.celrep.2016.04.035](https://doi.org/10.1016/j.celrep.2016.04.035); pmid: [27160899](https://pubmed.ncbi.nlm.nih.gov/27160899/)
 138. V. P. Chauhan *et al.*, Normalization of tumour blood vessels improves the delivery of nanomedicines in a size-dependent manner. *Nat. Nanotechnol.* **7**, 383–388 (2012). doi: [10.1038/nnano.2012.45](https://doi.org/10.1038/nnano.2012.45); pmid: [22484912](https://pubmed.ncbi.nlm.nih.gov/22484912/)
 139. P. A. Netti, D. A. Berk, M. A. Swartz, A. J. Grodzinsky, R. K. Jain, Role of extracellular matrix assembly in interstitial transport in solid tumors. *Cancer Res.* **60**, 2497–2503 (2000). pmid: [10811131](https://pubmed.ncbi.nlm.nih.gov/10811131/)
 140. T. Stylianopoulos, B. Diop-Frimpong, L. L. Munn, R. K. Jain, Diffusion anisotropy in collagen gels and tumors: The effect of fiber network orientation. *Biophys. J.* **99**, 3119–3128 (2010). doi: [10.1016/j.bpj.2010.08.065](https://doi.org/10.1016/j.bpj.2010.08.065); pmid: [21081058](https://pubmed.ncbi.nlm.nih.gov/21081058/)
 141. V. P. Chauhan, R. K. Jain, Strategies for advancing cancer nanomedicine. *Nat. Mater.* **12**, 958–962 (2013). doi: [10.1038/nmat3792](https://doi.org/10.1038/nmat3792); pmid: [24150413](https://pubmed.ncbi.nlm.nih.gov/24150413/)
 142. H.-B. Wang, M. Dembo, Y.-L. Wang, Substrate flexibility regulates growth and apoptosis of normal but not transformed cells. *Am. J. Physiol. Cell Physiol.* **279**, C1345–C1350 (2000). doi: [10.1152/ajpcell.2000.279.5.C1345](https://doi.org/10.1152/ajpcell.2000.279.5.C1345); pmid: [11029281](https://pubmed.ncbi.nlm.nih.gov/11029281/)
 143. A. Baur *et al.*, Diffusion-weighted MR imaging of bone marrow: Differentiation of benign versus pathologic compression fractures. *Radiology* **207**, 349–356 (1998). doi: [10.1148/radiology.207.2.9577479](https://doi.org/10.1148/radiology.207.2.9577479); pmid: [9577479](https://pubmed.ncbi.nlm.nih.gov/9577479/)
 144. L. E. Jansen, N. P. Birch, J. D. Schiffman, A. J. Crosby, S. R. Peyton, Mechanics of intact bone marrow. *J. Mech. Behav. Biomed. Mater.* **50**, 299–307 (2015). doi: [10.1016/j.jmbbm.2015.06.023](https://doi.org/10.1016/j.jmbbm.2015.06.023); pmid: [26189198](https://pubmed.ncbi.nlm.nih.gov/26189198/)
 145. J.-W. Shin, D. J. Mooney, Extracellular matrix stiffness causes systematic variations in proliferation and chemosensitivity in myeloid leukemias. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12126–12131 (2016). doi: [10.1073/pnas.1611338113](https://doi.org/10.1073/pnas.1611338113); pmid: [27790998](https://pubmed.ncbi.nlm.nih.gov/27790998/)
 146. A. Chramiec, G. Vunjak-Novakovic, Tissue engineered models of healthy and malignant human bone marrow. *Adv. Drug Deliv. Rev.* **140**, 78–92 (2019). doi: [10.1016/j.addr.2019.04.003](https://doi.org/10.1016/j.addr.2019.04.003); pmid: [31002835](https://pubmed.ncbi.nlm.nih.gov/31002835/)
 147. J. Fenner *et al.*, Macroscopic stiffness of breast tumors predicts metastasis. *Sci. Rep.* **4**, 5512 (2014). doi: [10.1038/srep05512](https://doi.org/10.1038/srep05512); pmid: [24981707](https://pubmed.ncbi.nlm.nih.gov/24981707/)
 148. G. Marasco *et al.*, Role of liver and spleen stiffness in predicting the recurrence of hepatocellular carcinoma after resection. *J. Hepatol.* **70**, 440–448 (2019). doi: [10.1016/j.jhep.2018.10.022](https://doi.org/10.1016/j.jhep.2018.10.022); pmid: [30389551](https://pubmed.ncbi.nlm.nih.gov/30389551/)
 149. R. Zhang *et al.*, Increased matrix stiffness promotes tumor progression of residual hepatocellular carcinoma after insufficient heat treatment. *Cancer Sci.* **108**, 1778–1786 (2017). doi: [10.1111/cas.13322](https://doi.org/10.1111/cas.13322); pmid: [28699238](https://pubmed.ncbi.nlm.nih.gov/28699238/)
 150. J. Incio *et al.*, Obesity-induced inflammation and desmoplasia promote pancreatic cancer progression and resistance to chemotherapy. *Cancer Discov.* **6**, 852–869 (2016). doi: [10.1158/2159-8290.CD-15-1177](https://doi.org/10.1158/2159-8290.CD-15-1177); pmid: [27246539](https://pubmed.ncbi.nlm.nih.gov/27246539/)

ACKNOWLEDGMENTS

Funding: NIH grants R01-CA208205 and U01-CA 224348, Outstanding Investigator Award R35-CA197743, and grants from the National Foundation for Cancer Research, Jane's Trust Foundation, American Medical Research Foundation, and Ludwig Cancer Center at Harvard (R.K.J.); NIH R01-CA2044949 (L.L.M.); F32-CA216944, Johnson and Johnson Cancer Initiative, Boston University Dean's Catalyst Award, and American Cancer Society Pilot Grant (H.T.N.). **Competing interests:** R.K.J. received an honorarium from Amgen; consultant fees from Merck, Ophthotech, Pfizer, SPARC, and SynDevRx; owns equity in Enlight, Ophthotech, and SynDevRx; and serves on the board of trustees of Tekla Healthcare Investors, Tekla Life Sciences Investors, Tekla Healthcare Opportunities Fund, and Tekla World Healthcare Fund. L.L.M. owns equity in Bayer AG. Neither any reagent nor any funding from these organizations was used in this study.

10.1126/science.aaz0868

RESEARCH ARTICLE SUMMARY

STRUCTURAL BIOLOGY

The coupling mechanism of mammalian respiratory complex I

Domen Kampjut and Leonid A. Sazanov*

INTRODUCTION: Complex I is the first and, with 45 subunits and a total mass of ~1 MDa, the most elaborate of the mitochondrial electron transfer chain enzymes. Complex I converts energy stored in chemical bonds into a proton gradient across the membrane that drives the synthesis of adenosine triphosphate (ATP), the universal energy currency of the cell. In each catalytic cycle, the transfer of two electrons from nicotinamide adenine dinucleotide (NADH) to a hydrophobic electron carrier quinone, which happens in the peripheral arm of the enzyme, is coupled to the translocation of four protons across the inner mitochondrial membrane in the membrane arm. The exact mechanism of this energy conversion currently presents an enigma because of complex I's size and the spatial separation between the two reactions.

RATIONALE: To understand the coupling mechanism of complex I, we solved its cryo-electron microscopy (cryo-EM) structures in five differ-

ent conditions, including the substrate- and inhibitor-bound states and during active turnover, unlocking the various conformations that the enzyme goes through during the catalytic cycle. We also improved the resolution to up to 2.3 to 2.5 Å, allowing us to directly observe water molecules critical for proton pumping.

RESULTS: We showed that opening and closing movements of the peripheral and membrane arms of complex I are critical for catalysis. Opening and closing is accompanied by coordinated conformational changes at the junction between the two arms, around the quinone binding cavity. These changes involve five conserved protein loops and are initiated by the reduction of quinone, the resulting negative charge in its cavity, and decylubiquinone (DQ) movement between the deep and the shallow binding sites. The bulky inhibitor rotenone also binds at these two sites and, unexpectedly, also within ND4—one of the three antiporter-like subunits. The deactive state is defined by a

notable relocation of the entire ND6 transmembrane (TM) helix 4, arresting the enzyme in the open conformation.

The PSST and 49-kDa subunit loops need to be ordered in the retracted state to enable quinone reduction as observed in the turnover closed class. Upon enzyme reduction with NADH in the open state, the 49-kDa loop extends into the cavity and the ND1 subunit loop flips upward, thus ejecting the reduced quinol. Conformational changes of the ND1 and ND3 loops also transmit the conformational changes in the quinone cavity to the rest of the enzyme (labeled “coupling” in the figure) by influencing the open-closed transition in the E-channel (the proton channel nearest to the quinone site). Entire TM helices of the ND1 subunit tilt upon opening, leading to a notable rotation of the ND6 TM3 helix, accompanied by the formation of the π -bulge.

Crucially, the rotation of this helix controls the formation of a critical water wire, which delivers protons from the conserved glutamates in subunit ND4L to the quinone site. This key feature brings the “charge action” of the quinone reaction directly next to ND2, the first out of the three homologous antiporter-like subunits, initiating a “wave” of electrostatic interactions propagating to the distal antiporter ND5. Analysis of water networks and charge distribution in the closed and open states of complex I under turnover explains how the protons are translocated in these waves within the antiporters and how this is coordinated between the four separate proton pumps and quinone reduction. A key role in this process is played by electrostatic interactions between the conserved charged residues, forming the highly hydrated “central axis” of the membrane arm. The distribution of the observed water molecules also suggests that links to the matrix and intermembrane space (IMS) sides in the distal subunit ND5 are much more hydrated than in other antiporters, and we propose the possibility that all four protons per cycle are ejected into the IMS via this subunit, rather than one per each antiporter (dashed arrows in the figure).

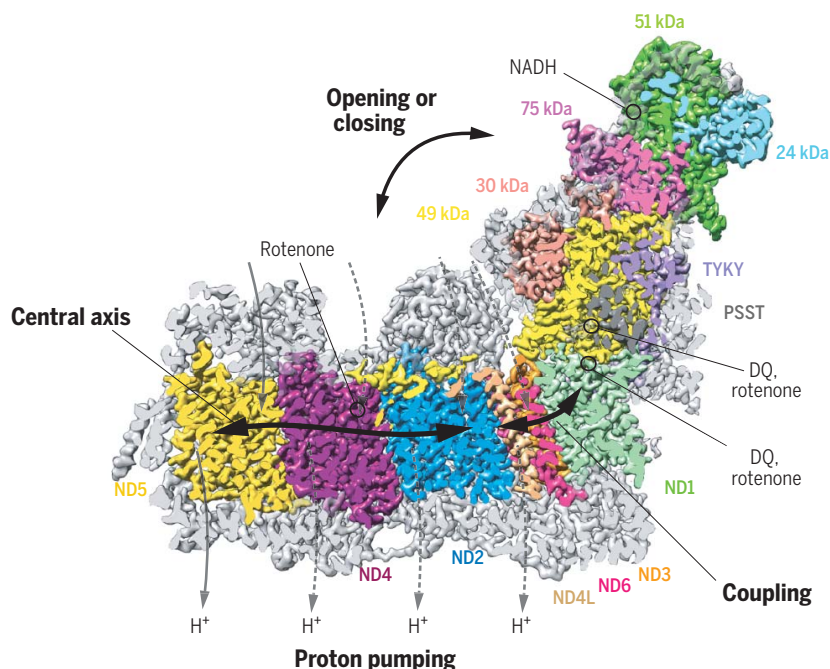
CONCLUSION: A comparison of conformational changes induced by substrate binding, turnover, inhibition, and deactivation allowed us to propose a detailed mechanistic model of the entire catalytic cycle in mammalian complex I, which combines elements of conformational (quinone site–E-channel) and electrostatic (antiporters) coupling. ■

The list of author affiliations is available in the full article online.
IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria.

*Corresponding author. Email: sazanov@ist.ac.at

Cite this article as D. Kampjut and L. A. Sazanov, *Science* **370**, eabc4209 (2020). DOI: 10.1126/science.abc4209

S READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.abc4209>



A roadmap to the binding sites, proton pathways, and coupling mechanism of complex I. A cross section of the cryo-EM density of complex I during turnover reveals an intricate machinery involved in catalysis. The solid black arrows show the propagation of the conformational changes and electrostatic interactions during the catalytic cycle; the gray arrows show the proton translocation pathways, with dashed arrows indicating the less likely paths. Core subunits are colored, and supernumerary subunits are in gray.

RESEARCH ARTICLE

STRUCTURAL BIOLOGY

The coupling mechanism of mammalian respiratory complex I

Domen Kampjut and Leonid A. Sazanov*

Mitochondrial complex I couples NADH:ubiquinone oxidoreduction to proton pumping by an unknown mechanism. Here, we present cryo-electron microscopy structures of ovine complex I in five different conditions, including turnover, at resolutions up to 2.3 to 2.5 angstroms. Resolved water molecules allowed us to experimentally define the proton translocation pathways. Quinone binds at three positions along the quinone cavity, as does the inhibitor rotenone that also binds within subunit ND4. Dramatic conformational changes around the quinone cavity couple the redox reaction to proton translocation during open-to-closed state transitions of the enzyme. In the induced deactive state, the open conformation is arrested by the ND6 subunit. We propose a detailed molecular coupling mechanism of complex I, which is an unexpected combination of conformational changes and electrostatic interactions.

Complex I is the largest of the respiratory complexes and, in mammals, is composed of 45 subunits with a total mass of about 1 MDa (Fig. 1A). It catalyzes the transfer of two electrons from NADH [reduced form of nicotinamide adenine dinucleotide (NAD)] to ubiquinone, coupled to the translocation of four protons across the bacterial or inner mitochondrial membrane, generating proton-motive force (PMF). It can also work in reverse, using PMF to reduce NAD⁺ (1, 2). Fourteen core subunits that are conserved in all species are necessary for the reaction. About 30 eukaryotic supernumerary subunits contribute to the regulation and stability but are nevertheless crucial for the catalytic activity of the mammalian complex I (1, 3). Initial crystal structures of the bacterial enzyme (4, 5) were followed by cryo-electron microscopy (cryo-EM) structures of mammalian mitochondrial complexes (6–9).

The catalytic mechanism of complex I, in particular the coupling between NADH:ubiquinone oxidoreduction and proton pumping over a distance of more than 200 Å, remains elusive. In the peripheral arm (PA) of the enzyme, electrons from NADH are accepted by flavin mononucleotide (FMN) and passed along a chain of iron-sulfur clusters to cluster N2 and to ubiquinone, which binds in a narrow tunnel at the interface between the PA and the membrane domain (MD) (Fig. 1A). The MD comprises four separate proton pumps connected by a string of conserved charged residues along the middle of the entire MD length, forming the central hydrophilic axis. The closest to the quinone cavity is the E-channel composed of ND1, ND6, and

ND4L subunits, followed by three homologous antiporter-like subunits—ND2, ND4, and ND5 (Fig. 1A). Antiporters are composed of two symmetry-related half-channels of five transmembrane helices (TMH) each, with the N-terminal half thought to be open to the matrix and the C-terminal half to the intermembrane space (IMS). Half-channels contain conserved key residues, lysine or glutamate, sitting on broken TMH7 and TMH12 and are connected into a full channel by a central lysine from broken TMH8. The N-terminal key TMH7 lysine forms a pair with the conserved TMH5 glutamate, which is thought to modulate the pK_a of lysine, where K_a is the acid dissociation constant (1, 4, 5).

The crucial energy-releasing step in the reaction is likely quinone reduction or its release out of the cavity because there is no drop in the midpoint redox potential of electron carriers until the electron reaches the final cluster N2 and quinol is released into the bilayer (10, 11). Molecular dynamics simulations suggested that diffusion of quinol out of the cavity is accompanied by the rearrangement of the quinone cavity loops (12). Quinol was also predicted to bind at two broad sites, one in the deep part of the cavity next to N2 (Q_d, or molecular dynamics sites 1 and 2) and one in the shallow part (Q_s, or molecular dynamics sites 4 and 5) close to the exit from the tunnel (13, 14).

We proposed that the negative charge within the quinone cavity initiates conformational changes in the E-channel, propagating further into the antiporters via the central hydrophilic axis, facilitated by flexible broken TMH (1, 4). These movements would influence interactions between the key lysine and glutamate residues within and between antiporters, leading to pK_a changes and proton translocation (5). In the alternative electrostatic spring

proposal, coupling in the membrane domain is coordinated by a wave of electrostatic interactions (15). The most detailed mechanism includes quinol-induced changes of the charge state of conserved residues, resulting in forward and backward electrostatic waves and breaking and formation of water wires in the antiporters and proton pumping (16, 17). However, all proposals so far lack experimental grounding and thus specifics of any conformational changes.

Complex I either isolated or within supercomplexes has so far been observed in two states: the closed and the open conformations, which differ in the PA-MD angle (6, 7, 9) (Fig. 1B). The opening coincides with the unfolding of several loops, which can be divided into the quinone cavity-forming loops and the interface-forming loops (Fig. 1C). The former include the 49-kDa β1-β2 loop (yellow) and the PSST loop (gray) that form the deep part of the quinone binding pocket and the ND1 TMH5-6 loop (green) that forms the shallow part of the pocket. The interface-forming loops consist of the ND3 TMH2-3 loop (orange) and the TMH3-4 loop of ND6 (blue). Closed-to-open transition is also associated with a notable rotation of the C-terminal half of ND6 TMH3 (blue) and the appearance of the π-bulge in the middle of the helix (8, 9).

It has been argued that the open conformations of mouse and bovine complex I correspond to the deactive state, which is a catalytically inert state of complex I that occurs at increased temperatures (30° to 37°C) in the absence of substrates and can be reversed by slow turnover (8, 18, 19). However, open and closed conformations of complex I were observed within active preparations of ovine respiratory supercomplexes (9, 20), which suggests that the open conformation per se is not a deactive state but a bona fide catalytic intermediate of complex I.

To investigate the catalytic mechanism of complex I in detail, we compared high-resolution cryo-EM structures of ovine complex I in five different conditions: native (no additions), NADH (enzyme reduced in the presence of NADH), rotenone (enzyme inhibited by rotenone in the presence of NADH), deactive (enzyme converted into deactive state), and turnover (enzyme flash-frozen while actively catalyzing NADH:decyl-ubiquinone oxidoreduction).

Comparison of open and closed conformations of complex I

An active and stable complex I preparation in a mild lauryl maltose neopentyl glycol (LMNG) detergent enabled us to achieve resolutions of up to 2.3 Å for the PA, 2.5 Å for the MD, and 2.6 Å overall (Fig. 1B and figs. S1 to S5). We built an almost complete model of complex I (96.9% residues modeled), which has better-defined geometry in previously ambiguous

IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria.

*Corresponding author. Email: sazanov@ist.ac.at

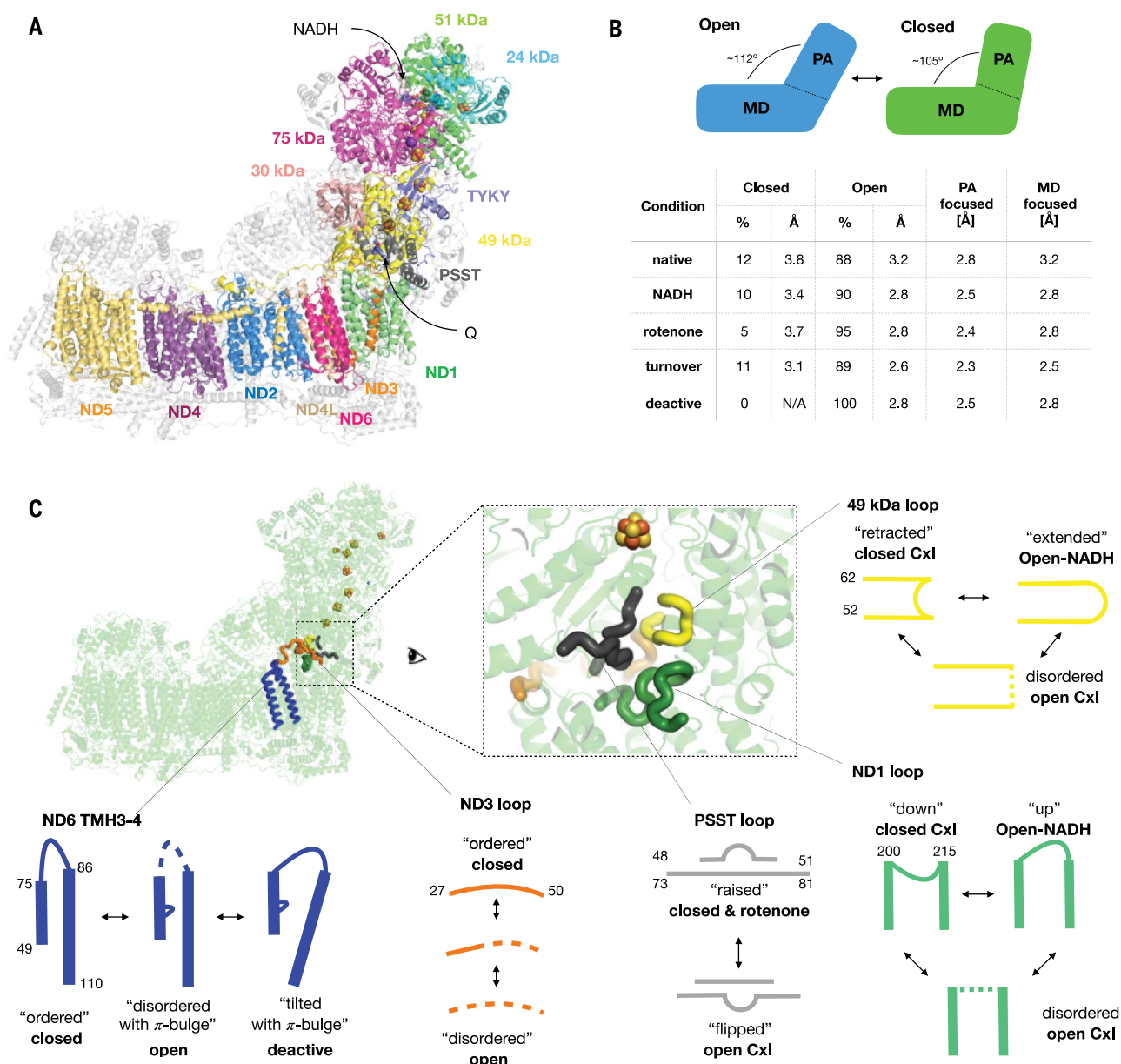


Fig. 1. Complex I catalyzes redox-coupled proton pumping. (A) Overall structure of ovine complex I. Core subunits necessary for the reaction of complex I are labeled with corresponding colors, and mammalian supernumerary subunits are shown in gray. NADH and quinone binding sites are indicated. The membrane arm contains four separate proton-pumping channels: three in the antiporter-like subunits ND2, ND4, and ND5 and one in the E-channel, composed of subunits ND1, ND6, and ND4L. Q, quinone. (B) Ovine complex I exists in open and closed conformations, the ratios of which differ between the different conditions investigated. The resolutions

achieved for each state overall and for PA- or MD-focused refinements are shown. N/A, not applicable. (C) Conformational dynamics of the quinone binding site loops. Loops at the interface between the peripheral and membrane arms, ND3 (orange) and ND6 (blue), adopt different states, as shown schematically. Quinone binding cavity loops include the 49-kDa loop (yellow), PSST loop (gray), and ND1 loop (green), which also undergo radical changes. Residue numbers indicate the start and end of the loops. The eight iron-sulfur clusters that transfer electrons from NADH to quinone are shown as orange and yellow spheres. CxI, complex I.

regions and includes experimentally defined water molecules. We also resolved other high-resolution features, including posttranslational modifications (figs. S6 and S7 and supplementary text).

Our datasets contained a mixture of open and closed conformations of complex I that arose from the combined side- and frontward tilting of the PA by up to 7°, as described

previously (6, 7, 9) (fig. S9D). The relative ratios between open and closed conformations were roughly the same in most datasets, with the exception of the deactive dataset, which completely lacked a closed conformation, and the rotenone-inhibited dataset, which had a smaller proportion of the closed state (Fig. 1B).

In each dataset, several open classes could be classified. The most-open classes (largest

PA-MD angle) of the native, deactive, and rotenone datasets showed partial disordering of the ND6 TMH4, 49-kDa N terminus, anchor helix of ND5, and B14.7, which could represent a partially disassembled or deactivated enzyme. In other open classes, these regions were ordered and did not differ from each other, particularly at the catalytically important sites. The different degrees of openness in open classes

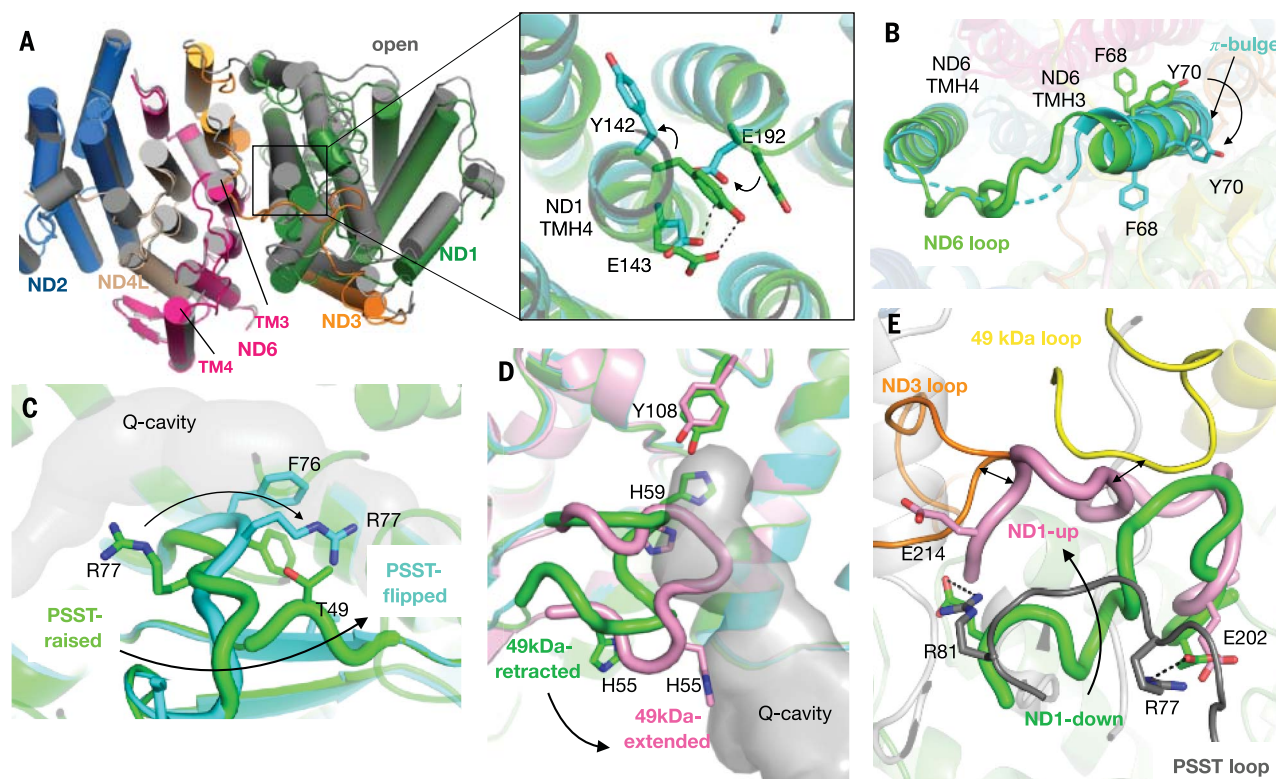


Fig. 2. Opening and reduction of the complex induce long-range conformational changes. (A) Reorganization of the ND1 subunit upon opening, with a detailed view of the rearrangements of the conserved glutamates that are part of the E-channel. On the left, the open state is gray and the closed state is colored by subunit, whereas on the right, the open state is cyan and the closed state is green. (B) TMH3 of ND6 rotates by almost 180° in the open state (cyan) and acquires a π -bulge near Tyr⁶⁰, and the TMH3-4 loop gets disordered. (C) The PSST loops and Arg⁷⁷ undergo large conformational changes between open (cyan) and closed

(green) states. (D) The 49-kDa loop is ordered in the closed state (green), gets disordered upon opening (cyan), and reorders in the extended position in the reduced open state (purple). (E) The ND1 loop, which is in the down conformation in the closed state (green), gets disordered upon opening and switches to the up conformation in the reduced open state (purple). In the up position, ND1 clashes with the ND3 (orange) and 49-kDa (yellow) loops, which are also reordered. Reorganization of the ND1 loop results in the breaking of two salt bridges with PSST_Arg⁷⁷ and Arg⁸¹ (gray). E, Glu; F, Phe; H, His; R, Arg; T, Thr; Y, Tyr.

thus likely reflect the loose PA-MD interface rather than functionally different states of complex I. We therefore treat open complex I as a single, albeit heterogeneous, state.

The opening of complex I coincided with disordering of several PA-MD interface-forming loops (Fig. 1C and fig. S6, C to E). In all of the datasets, we observed disordering of the central part of the ND3 loop, the ND6 loop, and the amphipathic helix (residues 259 to 272) of the supernumerary 39-kDa subunit. In the entire ND1 subunit, dramatic tilts of most helices resulted in up to 5-Å displacements at the matrix side, which propagate the conformational change all the way to the E-channel and also lead to the observed tilting of the PA (fig. S9, D and E). Notably, as the ND1 TMH4 bends near Tyr¹⁴² upon such a transition, this tyrosine swings out by almost 180°, freeing the space for the conserved Glu¹⁹² to move within hydrogen-bonding distance of the conserved Glu¹⁴³ (Fig. 2A). Changes in the density appearance indicate that at least one of the glutamates becomes protonated in the open state, because the estimated pK_a increases by ~2 units for

both [ProPKa software (27)]. The formation of the π -bulge in ND6 TMH3, resulting in a notable 120° rotation of half of the helix, accompanies this closed-to-open transition (Fig. 2B).

Quinone binding loops

In contrast to the interface-forming loops, the conformations of the quinone-site loops (49 kDa, PSST, and ND1) depended not only on the openness but also on the ligands bound and the redox state of the whole complex. The 49-kDa loop was ordered in the retracted (out of quinone cavity) conformation in the closed classes and disordered in the open classes of the native, turnover, and deactive datasets. When rotenone was bound in the Q_d site in open complex I, this loop was ordered in the retracted conformation, suggesting that binding of a ligand in the Q_d site can order the loop. Notably, in open classes of NADH-reduced enzyme (but not closed classes), the 49-kDa loop was ordered in the previously not observed extended (into the quinone cavity) conformation, shortening the cavity considerably (Fig. 2D). Because the extended conformation

was not observed in the rotenone and turnover datasets, which also contained NADH, the quinone-like ligand probably must be fully ejected from the cavity before the loop can extend.

The ND1 loop was similarly ordered in the “down” conformation in the closed classes and disordered in the open classes. Only in the NADH open classes (but not closed), however, did the ND1 loop rearrange into an “up” position, which displaced the ND3 loop as well as the other two ND1 matrix side loops (Fig. 2E). Disorder of the ND1 loop breaks several salt bridges, among which are highly conserved ND1_Glu²¹⁴-PSST_Arg⁸¹ and ND1_Glu²⁰²-PSST_Arg⁷⁷. Mutagenesis showed inhibitory effects of PSST_Arg⁷⁷ and PSST_Arg⁸¹ mutations, and molecular dynamics simulations suggested that quinone can only diffuse within the cavity when the ND1 loop is disordered (22). In the up conformation, ND1_Glu²¹⁴ formed an alternative salt bridge with ND1_Arg⁶², whereas ND1_Glu²⁰² was exposed into the cavity (Fig. 2E). The ND1 loop did not switch into the up conformation in the rotenone and turnover datasets, which also contained NADH,

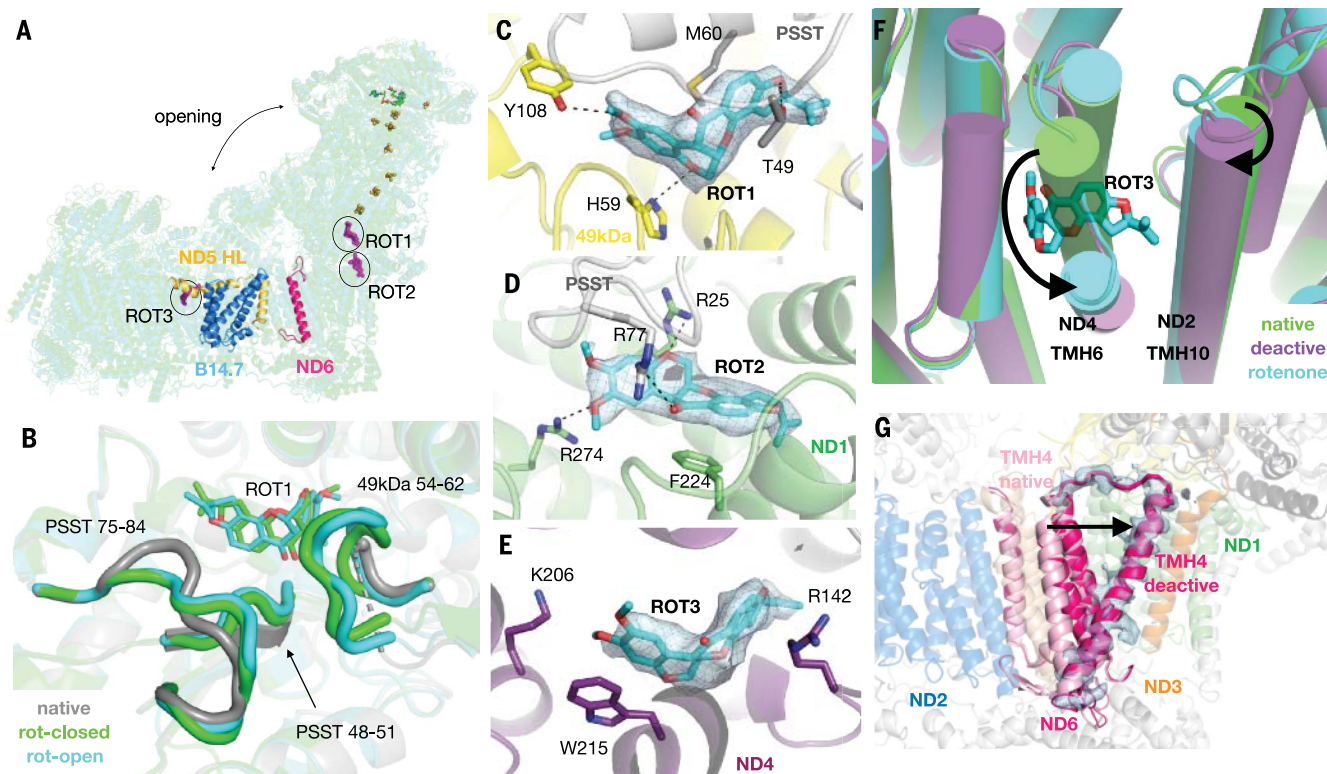


Fig. 3. Deactivation and rotenone inhibition elicit alternative conformations of complex I. (A) The most-open classes of rotenone-inhibited and deactive complex I exhibit disorder in the ND5 horizontal and anchor helix as well as in the B14.7 subunit (colored). TMH4 of ND6, which undergoes rearrangements in the deactive state, is labeled in magenta. The three rotenone binding sites are indicated as well; ROT2 binds to open classes only, and ROT3 binds to open2 and open3 classes. (B) The PSST and 49-kDa loops do not change between open (cyan) and closed (green) rotenone states. The native open conformation is shown in gray. (C) The deep rotenone binding site: Rotenone is within hydrogen-bonding distance to 49kDa_His⁵⁹ and Tyr¹⁰⁸ and mimics the conformation of bound quinone during reduction by the N2 cluster. Rotenone density is shown at

the contour of 2σ [(C) and (E)] or 3σ (D). (D) The shallow rotenone binding site: Rotenone is stabilized by polar interactions with conserved arginines and stacking interactions with ND1_Phe²²⁴. (E) ND4 rotenone binding site: Rotenone is stabilized by positively charged residues, including Arg¹⁴² and proton-pumping Lys²⁰⁶, and by stacking interactions with Trp²¹⁶. (F) To open the ND4 subunit binding site, the outer TMH6 needs to undergo a slight tilting and rearrangement, which is observed in the rotenone open2 and open3 classes (cyan) and also in the most-open class of the deactive dataset (purple). The native structure is overlaid in green. (G) In deactive complex I, ND6 TMH4 gets tilted (density in blue) toward the matrix-membrane arm interface and the TMH3-4 loop inserts between the two arms, stabilizing them in the open conformation. K, Lys; M, Met; W, Trp.

suggesting that binding of the ligand in the quinone cavity reverses the NADH-induced changes and could be an important step during the catalytic cycle.

The PSST loop itself also changed the conformation in most datasets. In the closed classes, it adopted the “raised” conformation, in which residues 48 to 51 bulge out into the quinone cavity, whereas in the open classes they form part of the β sheet. Furthermore, residues 75 to 80 flip in the open classes, resulting in a 180° rotation of PSST_Arg⁷⁷ (Fig. 2C). Only in the rotenone dataset, however, did the PSST loop always adopt the raised conformation and PSST_Arg⁷⁷ contribute to rotenone binding.

Rotenone binds in three sites, including a site in the antiporter subunit

Rotenone is one of the strongest specific inhibitors [half-maximal inhibitory concentration (IC₅₀) ~ 1 nM] of mammalian complex I (23). Because rotenone is much bulkier than typical

Q-like inhibitors (such as piericidin A), it was hypothesized that it cannot penetrate into the Q cavity and might instead block it by binding outside (24). However, in the rotenone dataset (enzyme incubated with 5 mM NADH and 300 μ M rotenone), we observed rotenone density in the Q_d site in all 3D classes, coordinated by the key 49kDa_Tyr¹⁰⁸ and 49kDa_His⁵⁹, mimicking decylubiquinone (DQ) and piericidin A binding in *Thermus thermophilus* (ROT1; Fig. 3, A to C) (4). Additionally, in the open classes, a second rotenone molecule was found in the Q_s site close to ND1_Arg²⁵ [molecular dynamics site 4 (13, 14)] (ROT2; Fig. 3, A and D). Finally, in the open2 and open3 classes (84% of all particles), a third rotenone density was unexpectedly observed in the ND4 subunit, far away from the quinone cavity (ROT3; Fig. 3, A and E). This binding site is created by the tilting out of TMH6 and is deep between TMH5-7, directly contacting the key ND4_Lys²⁰⁶ in the proton channel (Fig. 3F). The ND4

binding site could correspond to the lower-affinity (80 nM) binding site reported for the deactive conformation (23), which would also be consistent with changes that we observe in the deactive conformation, exposing this pocket (Fig. 3F). Binding in the ND4 site could also explain why rotenone has been observed to inhibit Na/H⁺ antiporter activity of deactive complex I (25). ND4-bound rotenone, curiously, did not influence the conformations of the proton-pumping residues, despite being bound by ND4_Lys²⁰⁶. It would be interesting to see if other inhibitors bind in this region, particularly biguanides, which bind more strongly to the deactive enzyme (19). Inhibitors specific for antiporters could be a useful tool for studies on proton pumping.

As noted above, in the open enzyme, binding of two rotenone molecules in the quinone cavity ordered the 49-kDa and PSST loops (Fig. 3B) but not the ND1 or ND3 loops, which suggests that ligand binding to the

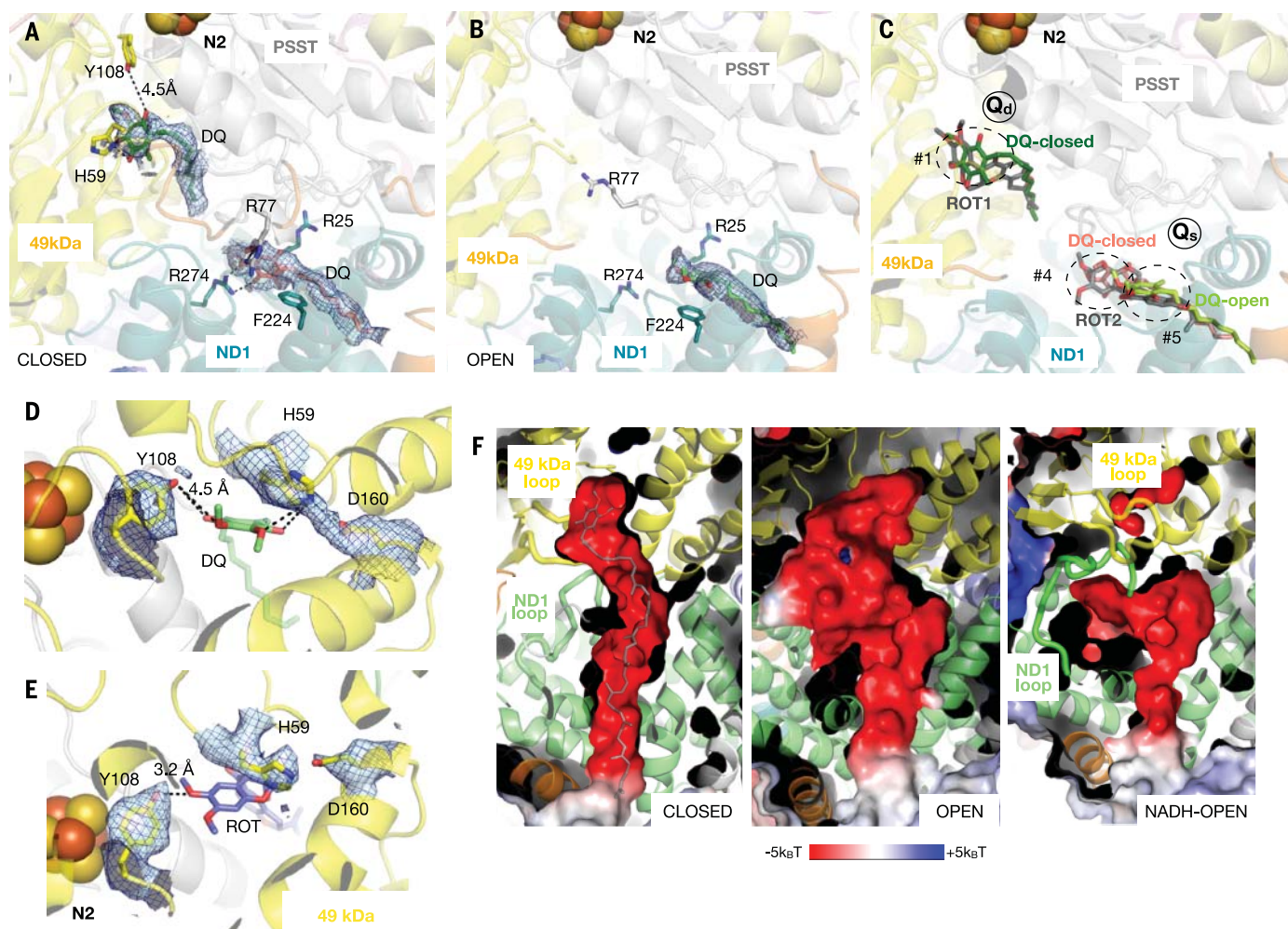


Fig. 4. Quinone density is observed at different binding sites in open and closed complex I during turnover. (A) In closed complex I, one quinone molecule is bound in the deep pocket between 49kDa_Tyr¹⁰⁸ and His⁵⁹ and another in the shallow pocket close to ND1_Arg²⁵. Density is shown at the σ contour of 1.3 (A), 2 (B), 1.8 (D), and 2 (E). (B) In open complex I, a single quinone is bound at the entry to the quinone binding cavity close to ND1_Phe²²⁴. (C) Comparison of binding positions of quinone and rotenone in different classes. (D) Detailed view of the binding of quinone next to the N2 cluster and the density

showing the salt bridge between 49kDa_His⁵⁹ and Asp¹⁶⁰. (E) 49kDa_Asp¹⁶⁰ displays weak density and no connection to His⁵⁹ in all of the classes apart from the closed turnover class; the rotenone open2 structure is shown here. (F) Cross sections through the quinone binding cavity in the closed, open, and NADH-open states colored by electrostatic potential reveal that opening and reduction result in large changes of the quinone binding cavity shape. In the closed state, ubiquinone-10 (UQ₁₀) is modeled (gray) based on decyl-ubiquinone binding, illustrating tight fit of the native substrate. $k_B T$, the product of the Boltzmann constant and the temperature; D, Asp.

open complex I directly affects the conformations of the PSST and 49-kDa loops, whereas the ordering of the ND3 and ND1 loops can only happen upon reduction or closing of the complex. Overall, rotenone appears to be distinctive among complex I inhibitors by having multiple binding sites, in contrast to Q-like inhibitors that bind only in the Q_d site in bacterial complex I (4) or close to it in the *Yarrowia lipolytica* enzyme (26).

Deactivation leads to tilting of the ND6 TMH4 helix, arresting the complex

To better understand the structural rearrangements leading to complex I deactivation, we prepared deactive ovine complex I by incubation without substrates at an increased tem-

perature (fig. S9, A to C) (27) and collected a cryo-EM dataset. The deactive sample did not contain any closed conformation classes, but we could classify four open classes, which were overall more open than the native open classes. They also had a large portion of the B14.7 subunit disordered, and classes 3 and 4 also showed some disorder around the horizontal helix HL and its transmembrane anchor, the N terminus of the 49-kDa subunit and the TMH5 of ND4 (Fig. 3, A and F).

The most notable feature of all deactive classes, however, was the almost complete relocation (by tilting) of the ND6 TMH4 helix, accompanied by the insertion of the TMH3-4 loop between the PA and MD, as well as unfolding of the residues 112 to 118 of the ND6

β sheet (Fig. 3G). The TMH3-4 loop gets disordered in the open state and TMH3 acquires a π -bulge, which is likely the “waiting” state of the catalytic cycle when quinone cavity loops get partially and transiently disordered. During a prolonged absence of turnover, unfolding proceeds to encompass wider regions of the complex further away from the quinone cavity and results in TMH4 tilt and a stable insertion of the ND6 loop between the two arms, preventing transition into the closed conformation and thus deactivating the enzyme. Reversal of this process is consequently slow and requires multiple turnovers. The stable decoupling of matrix and membrane arms of complex I and perhaps also the rearrangements within the ND4 subunit are consistent with the observation

that deactive bovine complex I is a Na/H⁺ antiporter (25).

A distinct deactive state structure provides strong evidence that the open state of complex I is a part of the catalytic cycle and not a deactive form, as proposed earlier (8, 18). The open enzyme is thus a quinone-binding—i.e., waiting—intermediate. The high kinetic barrier of active-to-deactive transition in mammals is thus explained by the large-scale reorganization of ND6. This is consistent with previous mouse and bovine “deactive” conformations that showed weak ND6 TMH4 density, indicating that they were partly, but not fully, deactivated (or the fully deactive state class was not separated in 3D classification) (8, 18). They are also consistent with *Y. lipolytica* structures in which ND6 TMH4 is much less likely to shift owing to its close association with ND5 TMH16 and the NUJM (homolog of B14.7) subunit (28). Hence, the *Y. lipolytica* complex I quasi-deactive state is actually a waiting open state, in which the ND3 loop is disordered and thus is susceptible to alkylation, but the enzyme reverts into the processive state without lag.

Under turnover, quinone binds both in the deep and shallow sites

We collected a dataset of complex I incubated with NADH and decylubiquinone and flash-frozen before the substrates were exhausted. Under such active turnover conditions, one may expect to identify quinone binding modes and any conformational changes associated with the catalytic cycle, provided that relevant conformations can be separated by classification and that the enzyme explores the entire conformational space in the absence of membrane and PMF. Such an approach with the *Y. lipolytica* enzyme suggested an absence of large conformational changes, although medium resolution (4.5 Å) prevented further mechanistic insights (29).

Both open and closed classes of our turnover dataset contained a clear density for bound NADH, showing that this is the most populated state of the NADH-binding site during turnover, even though iron-sulfur clusters and FMN are likely to be mostly reduced because quinol release is a limiting step (11, 30).

Closed complex I contained two bound decylubiquinones. Quinone in the deep binding pocket (Q_d) was bound between 49kDa_His⁵⁹ and 49kDa_Tyr¹⁰⁸, exactly where rotenone binds (Fig. 4A). We also observed a strong density near the entry to the quinone cavity, overlapping with the second rotenone site, the molecular dynamics site 4, which we modeled as another molecule of decylubiquinone (Fig. 4, A and C). Notably, residues 49kDa_Asp¹⁶⁰ and 49kDa_His⁵⁹, which coordinate quinone in the Q_d site, had a continuous density between them, indicating the existence of a salt bridge, probably as a result of proton donation to

quinone (Fig. 4D). This density was not present in any of the other conditions (Fig. 4E and fig. S10H) and 49kDa_Asp¹⁶⁰ density was weak, as is usual for unbound carboxylates in cryo-EM (31). Therefore, it is likely that 49kDa_His⁵⁹ first donates a proton to reduce quinone and then accepts a proton from 49kDa_Asp¹⁶⁰, so that negative charge resides on 49kDa_Asp¹⁶⁰ and the salt bridge is formed. This is the first experimental demonstration that the 49kDa_His⁵⁹-49kDa_Asp¹⁶⁰ pair acts as a proton donor for the quinone. It also indicates that the species bound in the Q_d site in the turnover dataset is quinol rather than quinone.

In the open turnover classes, a quinone density was present only in the shallow pocket (Q_s), but it was located further outward compared with that in the closed class and was overlapping with molecular dynamics site 5 (Fig. 4, B and C). The mobile loops of the quinone binding cavity were disordered to approximately the same extent as in the native open classes, and we did not see any density in the deep binding pocket. This suggests that enzyme opening affects quinone binding and dynamics in the cavity and likely facilitates quinol ejection or quinone uptake by the complex. The presence of two molecules of quinone in the closed class is probably an artifact of using decylubiquinone with its short carbon tail. If native ubiquinone-10 were bound in Q_d, its isoprenoid tail would clearly prevent quinone binding in Q_s in the tight quinone cavity of the closed enzyme (Fig. 4F). The opening of complex I widens the negatively charged quinone binding cavity, in large part owing to PSST rearrangements (Fig. 4F), which will facilitate quinone movements in and out of the cavity before and after reduction. In the NADH open conformation, the cavity becomes much shorter, mostly owing to 49-kDa loop insertion, which will facilitate quinol ejection.

The role of water molecules in quinone protonation and proton pumping

The abundant water molecules that we observed in the PA cryo-EM density were mostly consistent across the different conditions; hence, they are likely to represent tightly bound stationary molecules (fig. S10C). Several of them are close to the iron-sulfur clusters, but none lie directly between them; thus, they are not likely to participate in electron tunneling, as proposed before from simulations (32). Unlike the other iron-sulfur clusters, N2 contains two water molecules at hydrogen-bonding distance, which may help to increase its redox potential (fig. S10D).

High resolution of the turnover data allowed us to analyze the mechanistically important water distribution in the MD. At 2.5-Å resolution in cryo-EM, most water molecules are resolved, as confirmed by our analyses—we observed considerably more waters (1293) in

the open class MD at 2.5 Å than predicted by Dowser software (742). The MD was less hydrated than the PA, with a clear pattern of hydrated matrix and IMS surfaces and an abundantly hydrated central axis with relatively dry regions in between (Fig. 5A). About 90 waters were identified on the central axis, within 6 Å from key residues, whereas only about 60 waters were predicted by Dowser, indicating that we likely have a complete experimental picture of water distribution in the MD core. Especially hydrated were the cavities around key TMH7 and TMH12 lysines and glutamate, containing about 10 waters each. A similar overall pattern was obtained for the turnover closed state MD, although fewer waters were identified (fig. S10B) owing to lower resolution (2.9 Å).

All the key protonatable residues along the entire hydrophilic axis from the ND5 tip to the Q cavity are clearly interconnected by many water molecules, with the exception of a long break in the E-channel between ND4L_Glu³⁴ and ND3_Asp⁶⁶ in the open state (green arrow in Fig. 5A). This is caused by the ND6 TMH3 π -bulge with bulky hydrophobic residues such as Phe⁶⁸ and Met⁶⁴ blocking the path. Notably, when TMH3 rotates and the π -bulge disappears in the closed state, these residues are rotated away and are replaced by the conserved Gly⁶² and Gly⁶³, creating a water-filled cavity between ND3_Asp⁶⁶ and ND4L_Glu³⁴. This creates a water chain extending from ND1 all the way to ND2 only in the closed state (Fig. 5, B and C). TMH3 of ND6 is the most conserved TMH in the MD, and even conservative mutations in the π -bulge area are highly detrimental to activity (5), indicating that breaking and reforming of the water wire is probably essential for the mechanism. Furthermore, three critical glutamates of the E-channel, ND1_Glu¹⁴³, ND4L_Glu³⁴, and ND4L_Glu⁷⁰ become charged in the closed enzyme, as suggested by their density appearance and pattern of hydrogen-bond interactions (fig. S9I). ND4L_Glu³⁴ also rotates away from ND4L_Glu⁷⁰ as they become charged and surrounded by waters.

Beyond the E-channel, we did not observe any conformational changes in any dataset, including turnover, either in the broken antiporter helices or in the traverse HL helix or connecting β hairpins. The HL helix thus probably serves mainly a structural role and is not involved in the coupling mechanism, as advocated previously (4, 5, 16, 17). Analysis of the proton transfer pathways connecting Grothuss-competent residues (Lys, Glu, Asp, His, Tyr, Thr, and Ser) and experimentally identified waters within hydrogen-bonding distances (with bond slack up to ~4 Å) suggested that E-channel and ND2 and ND4 antiporters lack any connections to the IMS (Fig. 5A, indicated by dashed pathways). These parts of the structure are particularly dry and hydrophobic. The matrix-facing

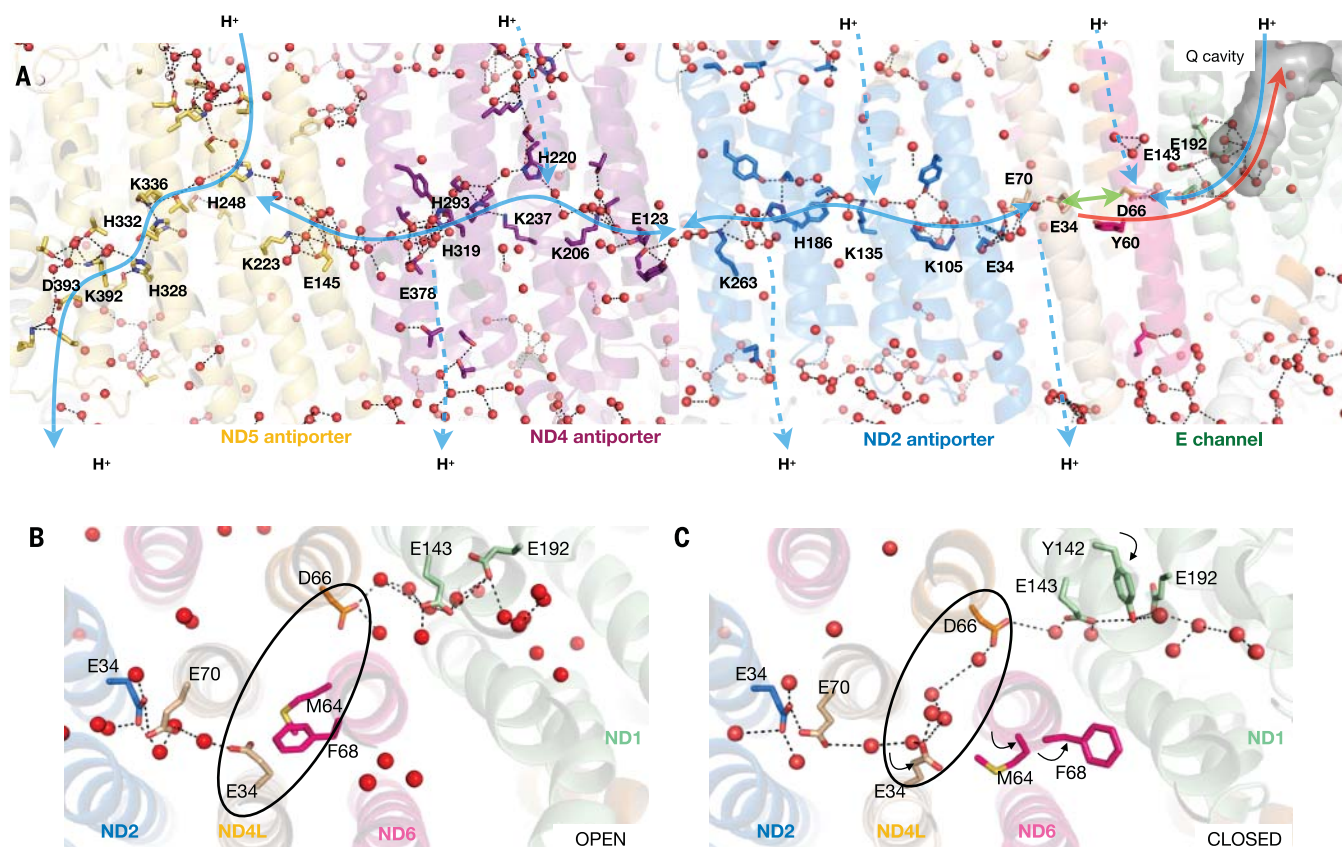


Fig. 5. Quinone protonation and proton pumping require coordinated water molecules. (A) The central hydrophilic axis connects the three antiporters in the open state of complex I. Both N- and C-terminal Lys and Glu are abundantly hydrated. Water wires toward the IMS and matrix side are visible only in ND5. The E-channel has a dry region between ND3_Asp⁶⁶ and ND4L_Glu³⁴ and no connection to the IMS or matrix side. It is connected to the quinone binding cavity only in the closed state (green arrow), allowing for protons from ND4L_Glu³⁴ and Glu⁷⁰ to be delivered to the Q site (red arrow). (B) Top view of the E-channel in the open state of complex I shows that the E-channel dry region (indicated by the black oval) is caused by the ND6 π -bulge formation and ND4L_Glu³⁴ rotation. (C) Continuous water wire in the E-channel of the closed state of complex I. The black oval indicates the same area as in (B). (D) Proposed proton-pumping mechanism. Charge distribution of the proton-pumping residues differs between the open and the closed states. The colored circles correspond to glutamate or aspartate (red), lysine (blue), or histidine (cyan) and are shown in the same orientation as in Fig. 1C. Full circles represent charged residues, and empty circles are neutral (the state of the histidines is not clear). Arrows show proton movements to achieve the charge distribution as depicted in each state.

part of ND4 is less dry and can allow connection to the matrix through conserved ND4_His²²⁰ upon conformational changes in broken TMH7, where ND4_His²²⁰ sits. This histidine is not conserved in ND2, which lacks any clear connections to the matrix. Notably, the distal antiporter ND5 is radically different—

the key TMH12 Lys³⁹² is clearly connected to the IMS through the conserved ND5_Asp³⁹³ and the abundant surrounding waters. Connection to the matrix side through the conserved ND5_His²⁴⁸ also exists. The side chain of this histidine is roughly in the same position as ND4_His²²⁰, but it sits on TMH8, replacing

LysTMH8 in ND5. Thus, both ND4_His²²⁰ and ND5_His²⁴⁸ sit on broken helices, allowing for their flexibility as needed to control the connectivity to the matrix, in tandem with key LysTMH8 in ND4 and the additional conserved essential Lys³³⁶ in ND5. The breaks of about 6 Å prevent links of ND4_His²²⁰ to the

matrix and ND5_His²⁴⁸ to the IMS (Fig. 5A, red-colored dashes), consistent with out-of-sync operation of ND4 and ND5 (as described below) and probably preventing the back-leak of protons.

The observed high connectivity along the central axis, the lack of links to the IMS side in antiporters other than ND5, and the absence of any conformational changes in antiporters even under turnover lead us to propose that all four protons may be pumped out to the IMS via ND5. This is a radical departure from the traditional view that protons are ejected one by one by each antiporter individually, as originally proposed by us (4, 5). However, the experimental structures argue that an ND5-only model should be considered as a more likely option. All four protons could hop over to ND5 via the abundant central axis links using conserved connecting histidine residues (cyan in Fig. 5D) as a temporary storage. The ND5 subunit, despite being distal, is the most conserved of the antiporters and has distinct features supporting this proposal (supplementary text). The input of protons from the matrix occurs in ND5 and possibly also ND4, with ND2 or E-channel inputs less likely. In all scenarios, protons taken up from the matrix are probably quickly redistributed along the fully interconnected central axis (Fig. 5A).

Proton-pumping mechanism

The absence of conformational changes in antiporter-like subunits under turnover suggests that proton pumping in them is mainly electrostatically driven. This is supported by a clear pattern of changes in charge distribution of key residues between open and closed states revealed by the analysis of cryo-EM densities for hydrogen bonds and carboxylate side chains (fig. S9, F to I, and table S7). There is an asymmetric distribution of charges between the antiporters, with ND4 being out of sync with ND2 and ND5 (Fig. 5D). This may help to prevent excessive electrostatic imbalance in the MD if all antiporters were to transfer charges in sync. Asymmetry is probably achieved by the replacement of the key TMH12 lysine with glutamate (Glu³⁷⁸) in ND4. The switch in the orientation or openness of the proton channels is likely associated with the global change in the structure, i.e., between open and closed states. Because the open state is energetically more favorable, it is likely that in this relaxed state, most channels are open to the matrix side, so that protein can reprotonate the central axis from the matrix.

We propose the following proton-pumping mechanism. Our results indicate that quinone can bind and enter the cavity only in the open state of the complex. The proton-pumping cycle is initiated by quinone binding, accompanied by the transition to the closed state with the ND1 and 49-kDa loops reordering into the

down and retracted conformations, allowing quinone movement into the deep site, where it gets reduced. These changes lead to ND6 TMH3 rotation (movie S1), which in turn creates a water wire (isolated from bulk solvent) linking the Q cavity to ND4L_Glu³⁴ and Glu⁷⁰ (Fig. 5C). These residues donate two protons for reprotonation of the quinol-coordinating residues 49kDa_Asp¹⁶⁰, His⁵⁹, and Tyr¹⁰⁸, which creates a strong negative charge in ND4L. This charge is compensated by the nearby ND2_GluTMH5 taking up the proton from its partner LysTMH7. (We use the same notation for key residues for all antiporters, even though mammalian ND2 lacks three N-terminal TMH.) Absence of positive charge on LysTMH7 removes the electrostatic barrier and allows the connecting LysTMH8 to lose a proton to LysTMH12. This positive charge is compensated by the nearby ND4_GluTMH5, which donates a proton to ND4_LysTMH7. The connecting ND4_LysTMH8 is protonated from the matrix, but the positive charge on LysTMH7 prevents the movement of its proton further on to GluTMH12. This glutamate loses a proton to the IMS owing to electrostatic interactions with charged LysTMH7 and LysTMH8, and its negative charge is compensated by ND5_GluTMH5 taking up the proton from ND5_LysTMH7, similarly to the situation at the ND4L-ND2 interface. As in ND2, this leads to protonation of ND5_LysTMH12 by Lys³³⁶ (or His²⁴⁸), completing a transition from the open to closed state.

Subsequent diffusion of quinol toward the shallow site causes disordering of the 49-kDa, ND1, and ND3 loops as well as reformation of the π -bulge and the opening of the complex. As the enzyme gets reduced, the 49-kDa loop gets extended and the ND1 loop flips upward, helping to eject the quinol (movie S2). Upon complex I opening, ND4L_Glu³⁴ and Glu⁷⁰ will be reprotonated, via ND2 and the central axis, because the water wire to the Q site will be now broken. At this stage, one proton is likely to be ejected from the E-channel into the IMS. Arrival of protons to ND4L_Glu³⁴ and Glu⁷⁰ will cause the nearby ND2_GluTMH5 to donate a proton to LysTMH7. As the connecting LysTMH8 is protonated from the matrix, this additional positive charge will help to drive the LysTMH12 proton into the IMS. As positive charge on ND2_LysTMH12 is lost, its neighbor ND4_GluTMH5 acquires the proton back from LysTMH7. Loss of charge on LysTMH7 allows the connecting LysTMH8 to lose its proton to GluTMH12, and owing to electrostatic interactions, the nearby ND5_GluTMH5 will lose its proton to LysTMH7. As in ND2, the connecting ND5 residue is protonated from the matrix and the double positive charge on it and LysTMH7 will drive the LysTMH12 proton out, resetting the system to the open state. Thus, simple electrostatic interactions

are sufficient to drive proton translocation in a consistent manner across each antiporter, a feature that is allowed by out-of-sync operation of ND4.

In our scenario, the N-terminal Glu-Lys pairs serve mainly as polar or charge switches, controlling proton translocation between TMH8 and TMH12 residues. A similar proposal was put forward before on the basis of molecular dynamics simulations, with a key role for opening and closing of Glu-Lys pairs (17). We do not observe such opening and closing on a predicted scale, but rather a change in charge state. Proton transfer within these pairs, as suggested by our observations (fig. S9, F to I), is probably a more robust way to achieve consistent switching of channels. Furthermore, instead of forward and reverse electrostatic waves (16), proton translocation in our mechanism proceeds through two forward electrostatic waves driven by protonation and deprotonation of ND4L glutamates, in turn driven by quinone oxidoreduction through defined conformational changes in the ND1 and E-channel. Our mechanism is applicable both to traditional one-per-antiporter or new ND5-only modes of pumping, because the only difference in the case of the ND5-only model would be a redistribution of protons toward the ND5 subunit (instead of directly into the IMS) via the central axis.

Concluding remarks

On the basis of the above model of proton pumping and the various conformations of complex I observed here, we propose a detailed model of the entire catalytic cycle (supplementary text and fig. S11). Our mechanism explains the directionality and tight coupling of the reaction because (i) NADH-induced changes in the quinone cavity can only happen in the open conformation of the complex, (ii) reduction of quinone happens only in the closed state, and (iii) proton pumping is induced during the transitions between the two states. The mechanism is reversible: High PMF would promote the reverse reaction by driving charge transitions in antiporters in reverse to those in Fig. 5D. Translocation of protons into the matrix would lead to transfer of protons from the 49kDa_Tyr¹⁰⁸-His⁵⁹-Asp¹⁶⁰ triad to ND4L glutamates, creating a negative charge on the triad. It would promote quinol binding and oxidation, as well as lower the N2 redox potential, enabling reverse electron transfer from N2 to FMN and NAD⁺.

The key features of our mechanism are as follows: first, the delivery of negative charge from the redox reaction site toward antiporters through defined concerted conformational changes propagating from the Q site toward the E-channel; second, the description of how this negative charge drives the proton pumping in the antiporters by electrostatic wave

propagation; and third, the optional, but probable, alternative pathway of proton ejection into the IMS only via the ND5 subunit. Although we cannot exclude that additional intricacies might be revealed by studies of complex I in energized liposomes or by using a time-resolved cryo-EM approach, the proposed mixed conformational-electrostatic model is strongly supported by our data, is fully consistent with available mutagenesis data, and is sufficient to explain the previously enigmatic mechanism.

Materials and methods

Purification of complex I

Purification procedures were done at 4°C, as described before (20). Briefly, mitochondria were purified from fresh ovine heart tissue as described before and stored as pellets at -80°C. The sheep hearts were cooled on ice within an hour of slaughter, whereas transport and mitochondrial isolation took an additional ~10 hours. On the day of complex I purification, a 10-g aliquot of mitochondrial pellet was thawed, and inner mitochondrial membranes were isolated by rupturing mitochondria by homogenization in 100 ml of Milli-Q water, addition of KCl to 150 mM, and centrifugation at 50,000g for 45 min. Afterward, membranes were resuspended in 100 ml of buffer M [20 mM HEPES, pH 7.4, 40 mM NaCl, 1 mM EDTA, 10% v/v glycerol, 2 mM dithiothreitol (DTT), and 0.002% phenylmethylsulfonyl fluoride (PMSF)], centrifuged again, and resuspended in 50 ml of buffer M. To solubilize the membrane proteins, 10% LMNG was added dropwise to 1% final concentration to the membrane suspension and stirred for 45 min. Upon centrifugation (50,000g, 45 min) the supernatant was filtered with a 0.22-μm filter and loaded onto a 45-ml Q-sepharose HP anion exchange column equilibrated in buffer Q-A (20 mM HEPES, pH 7.4, 40 mM NaCl, 2 mM EDTA, 10% v/v glycerol, 1 mM DTT, and 0.05% LMNG). The column was washed with 75 ml of Q-A buffer, 50 ml of 10% Q-B buffer (Q-A with 1 M NaCl), and 125 ml of 20% Q-B buffer and finally complex I was eluted by a 200-ml linear gradient with 20 to 27% Q-B buffer. Complex I-containing fractions were pooled, concentrated to ~15 mg/ml using a Millipore 100-kDa cut-off filter, and stored, with 30% glycerol added, under liquid nitrogen. A 250-μl sample was thawed on ice when cryo-EM grids were prepared and loaded onto a Superose 6 10/300 gel filtration column equilibrated in buffer GF (20 mM HEPES, pH 7.4, 50 mM NaCl, 1 mM EDTA, and 0.002% LMNG) to remove excess detergent and remaining protein contaminants. The purest and the most concentrated fractions of complex I eluted at ~11.5 ml and were concentrated to 3.5 mg/ml and used immediately for cryo-EM grid preparation. Deactive complex I was prepared by incubating a concentrated aliquot of complex I after anion

exchange at 32°C for 30 min. To prevent delipidation and proteolysis, 0.5 mg/ml of 1:4 mixture of cardiolipin and dioleoylphosphatidylcholine and 0.0075% PMSF were added to the sample. Afterward, the sample was cooled on ice and injected to the gel filtration column, as described above.

Cryo-EM

CHAPS was added (final concentration 0.2%) to the protein sample before grid preparation to improve ice quality and particle distribution. The native and the deactive complex I samples were frozen without further substrate additions. For the NADH dataset, 5 mM NADH was added to the protein 20 min before the grid preparation, and the samples were incubated on ice. For the rotenone dataset, NADH (5 mM) was added to the complex I sample, which was then incubated for 1 min on ice and followed by the addition of 300 μM rotenone [dissolved in dimethyl sulfoxide (DMSO); final DMSO concentration in the sample ~1%] and 20 min incubation on ice. For the turnover dataset, a mixture of substrates and CHAPS was prepared at 2× working concentration (2 mM DQ, 2 mM NADH, and 0.4% CHAPS) and mixed in 1:1 ratio with 6 mg/ml complex I immediately before applying to the grid, blotting, and freezing. With this simple procedure, complex I was active for 20 s before being frozen, and less than 10% of substrates would have been consumed in this time. The final concentration of complex I used for grid preparation was 3 mg/ml for all samples.

A 2.7-μl protein sample was applied to a freshly glow-discharged (0.7 mbar and 30 mA for 2 min in the ELMO Glow Discharge unit, Agar Scientific, Stansted, UK) Quantifoil 0.6/1 copper grid and blotted using the Whatman filter paper no. 1 for 6 to 8 s using a blotting force of 25 at 4°C and 100% humidity in an FEI Vitrobot Mark IV. Grids were flash-frozen in liquid ethane and stored in liquid nitrogen until data collection.

Native, NADH, rotenone, and turnover grids were imaged using a 300-kV Titan Krios electron microscope equipped with a Falcon III camera operating in an integrating mode at the Cryo-Electron Microscopy and Tomography Core Facility in CEITEC, Brno, Czech Republic. Turnover, NADH, and rotenone datasets were imaged at a nominal magnification of 130,000× with a physical pixel size of 1.061 Å. Total electron exposure of 89 e/Å² was fractionated into 40 frames of 21 ms each. The native dataset was imaged at a nominal magnification of 165,000× with a physical pixel size of 0.83 Å. Total electron exposure of 98 e/Å² was fractionated into 40 frames of 25 ms each. Images were collected from 3-by-3 patches of holes using image shift—the resulting beam tilt error was small and refined during processing and this mode of acquisition allowed for a high speed of col-

lection. Deactive complex I was imaged using a 300-kV Titan Krios electron microscope equipped with a K3 camera operating in a superresolution mode at the Institute of Science and Technology Austria. Grids were imaged at a nominal magnification of 81,000× and physical pixel size of 1.055 Å. Total electron exposure of 78.5 e/Å² was fractionated into 50 frames of 88 ms each. The deactive dataset was collected using SerialEM with an active beam tilt compensation; hence, beam tilt error was not refined separately for each of the nine holes imaged using the image shift method (33).

Image processing

We started with 2284 movies for the native dataset, 1443 movies for the NADH dataset, 4239 movies for the turnover dataset, 2323 movies for the rotenone dataset, and 2065 movies for the deactive dataset. The micrographs were included in the further processing steps if the Thon rings extended to 4 Å (turnover dataset) or 5 Å (all other datasets) according to the CTFFIND maximum resolution estimate. Processing was done in RELION 3.0 and 3.1 (34) unless otherwise stated. Movie frames were aligned using MotionCor2 (35), and initial CTF parameters were estimated from averaged images using CTFFIND 4.1.13 (36). Autopicking with 2D class averages as references was done in RELION and resulted in 295,630 particles for the native dataset; 364,499 particles for the NADH dataset; 938,197 particles for the turnover dataset; 354,690 particles for the rotenone dataset, and 740,026 particles for the deactive dataset.

For the sake of consistency, processing steps were done as similarly as possible for all the datasets. Particles were extracted at 3× the physical pixel size and processed using one round of 2D classification and one round of 3D classification (figs. S1 to S5). Good classes were then extracted at full pixel size and autorefined into a single “consensus” structure. Contrast transfer function (CTF) parameters and per-particle trajectories were then refined in an iterative manner (one round of CTF refinement followed by Bayesian polishing and another CTF refinement). This led to a consensus refined complex I structure of all particles, reaching nominally high resolutions (up to 2.5 Å in the turnover dataset) but with blurring around the edges and some weaker subunits (42 kDa and B14.7), suggesting a large degree of conformational heterogeneity and presence of partially disassembled particles. Particles were therefore 3D classified without searches to remove damaged particles and then focus-reverse-classified to sort out the heterogeneity, as done earlier (9). Briefly, all the intact particles were focus-refined using a peripheral arm mask and then classified without searches using a loose membrane mask to finely classify states with slightly different angles between the two arms.

This procedure resulted in a single closed class and 3 to 4 open classes from each dataset, based on which atomic models were built, analyzed, and compared. In most datasets, finer subclassification of classes with different masks, in particular around the quinone binding cavity, was tried but did not result in finding additional conformational states of complex I. Further classifications of consensus refinements of the MD or PA of all particles also did not result in finding additional conformational classes. Although some datasets have shown preferential orientation of particles, this did not influence the quality of final maps owing to the presence of sufficient amounts of particles in other orientations.

In the native dataset, the first round of 3D classification was done with $k = 6$ classes and regularization parameter $T = 6$, which resulted in 148,307 good particles. Upon duplicate removal, consensus refinement, and two rounds of CTF refinement and Bayesian polishing, 128,964 particles were refined around the matrix arm and subjected to final classification without searches around the membrane arm with $k = 6$ and $T = 20$. This resulted in four good classes with 91,979 particles in total and two partially broken classes. The good classes were refined and postprocessed separately to give final maps for model building.

In the NADH dataset, the first round of 3D classification was done with $k = 6$ and $T = 4$, which resulted in 276,169 good particles. After duplicate removal, consensus refinement, two rounds of CTF parameter refinement, and Bayesian polishing, another 3D classification with $k = 6$ and $T = 8$ without searches was performed to remove damaged particles. The remaining 182,053 particles were refined around the PA and subjected to classification around the membrane arm with $k = 6$ and $T = 8$. This resulted in four good classes with a total of 140,630 particles and two classes with bad density and partially disassembled complex. Particles from the four good classes were refined and postprocessed and the maps were used for model building.

In the deactive dataset, the first round of 3D classification was done with $k = 6$ and $T = 4$ and resulted in 458,556 good particles. After a consensus refinement, another 3D classification step without searches and with $k = 6$ and $T = 10$ was performed to remove further damaged particles. This resulted in 272,990 particles, which were refined around the PA and classified around the membrane arm using 3D classification without searches and $k = 4$ and $T = 10$. This resulted in four classes that differed from each other by the degree of openness. They were subjected to two rounds of CTF parameter refinement, Bayesian polishing, refinement, and postprocessing to obtain the final density maps used for model building.

In the rotenone dataset, the first round of 3D classification was done with $k = 6$ and $T = 4$ and resulted in 175,658 good particles after duplicate removal. After two rounds of CTF parameter refinement, Bayesian polishing, and autorefinement, another 3D classification without searches and with $k = 6$ and $T = 8$ was performed to remove broken particles, resulting in 142,565 good particles, which were focus-refined around the membrane domain. Because of the low abundance of the closed class in this dataset, the closed class was separated by 3D classification around the membrane arm with $k = 6$ and $T = 10$. The rest of the particles were 3D classified separately with $k = 6$ and $T = 6$ to better separate different open classes. The resulting particles were refined, postprocessed, and used for model building.

In the turnover dataset, the initial 3D classification was done with $k = 6$ and $T = 4$ and resulted in 518,117 good particles after duplicate removal. After two rounds of CTF parameter refinement, Bayesian polishing, and autorefinement, another 3D classification without searches and $k = 6$ and $T = 8$ was performed to remove damaged particles, resulting in 344,541 good particles retained. These were then 3D classified around the membrane arm without searches and with $k = 6$ and $T = 8$. Four of the resulting classes contained good particles with a total of 257,669 particles, which were refined and postprocessed for model building. The density within the quinone cavity of the closed class appeared broken; hence, this class was further classified without searches and with $k = 3$ and $T = 50$. This resulted in one major class with about 55% particles that had better defined density in the quinone cavity and elsewhere and another minor class with about 30% of particles that had worse density overall but otherwise no noticeable differences to the major class. Only the noticeable closed class was used for model building. For the closed MD model, all the particles of the turnover closed class were focus-refined around the MD, however.

Model building and refinement

The models presented in this paper are based on the initial 3.9-Å structure of ovine complex I determined previously [Protein Data Bank (PDB) 5LNK] (6). MD and PA were built on the basis of 2.5- and 2.3-Å focus-refinements of all the open particles from the turnover dataset, which allowed us to improve the geometry, fit, and completeness of the models considerably (supplementary text). We were also able to model additional lipid molecules, as well as several known posttranslational modifications, including N-terminal acylations of subunits B18, B14.5b, and B14.7; N-formyl methionines in ND6, ND5, and ND4L; symmetric dimethylation of 49 kDa Arg⁸⁵; and phosphorylation of 42 kDa Ser³⁶ (37) (fig. S7). For accurate modeling of water molecules, particularly to avoid

false positives, we filtered the maps by local resolution and resampled them at 0.5 Å per pixel in Relion. After this procedure, water molecules displayed strong signal ($>2\sigma$), had nearly spherical densities, were not clashing with other atoms, and participated in hydrogen bonds, which are all strongly indicative of real water molecules. This allowed automatic placement of water molecules in Coot, which were then all checked and corrected manually, to leave only waters with clear density and fulfilling geometry criteria. Using this procedure, we could reliably place water molecules not only into 2.3- to 2.8-Å structures but also into the closed MD focus refined map (overall resolution 2.9 Å, higher in the core). The two highest-resolution open PA and open MD models were then rigid-body fit into all the classes, corrected in Coot (38) and refined using PHENIX software (39). We used two rounds of a single cycle of group atomic displacement parameters (ADP) refinement followed by three cycles of global minimization in real space (with secondary structure, Ramachandran, and c-beta restraints enabled) to optimize B factors so that electron radiation-damaged carboxylate side chains acquire high B factors and do not lead to main-chain distortions as described previously (9). To maintain consistency, we performed this procedure first with rotamer fit option enabled, checked, and corrected the model in Coot and then finally repeated the procedure without rotamer fit.

Structure analysis and preparation of figures

Geometry and density fit of the models was analyzed using MolProbity and EMRinger (40, 41). pK_a was estimated using PROPKA 3.1 software (21). Visualization and analysis of the models and corresponding density maps was done in PyMOL 2.2.3 and UCSF Chimera. The highest-resolution maps from PA and MD focused refinements were approaching Nyquist resolution and were resampled at 0.5-Å pixel size in RELION for depiction purposes, which resulted in a smoother density (figs. S6 and S7).

Activity measurements

Complex I NADH:DQ oxidoreduction activity was measured at 30°C spectrophotometrically by following the NADH ($\epsilon = 6.1 \text{ mM}^{-1} \text{ cm}^{-1}$) oxidation at 340 nm using a Shimadzu UV-2600 ultraviolet-visible spectrophotometer. Reaction buffer was optimized for the ovine CxI purified in LMNG before (20 mM HEPES, pH 7.4, 50 mM NaCl, 1 mM EDTA, 0.2% CHAPS, 0.05% LMNG, 0.25 mg/ml DOPC:CL lipids, 3.5 mg/ml bovine serum albumin, 200 μM NADH, and 200 μM DQ) (20). Complex I was incubated in the above buffer with stirring for 5 min before the measurement was started by NADH addition, and typical activity was about 5 to 6 U/mg. Labeling with *N*-ethylmaleimide (NEM) was performed by incubating a sample

of 1 mg/ml of purified complex I with 1 mM NEM on ice for 5 min, after which NEM was neutralized by the addition of 1 mM DTT. Sub-mitochondrial particles (SMPs) were prepared by resuspending 1 g of mitochondrial pellet in 8 ml of 10 mM HEPES, pH 7.0, and 250 mM sucrose, frozen at -80°C and thawed again. Mitochondria were then pelleted in a bench-top centrifuge, washed with 4 ml of the same buffer, and resuspended in a total of 4 ml of the above buffer with 10 mM MgSO_4 added. Mitochondria were then sonicated on ice (10 15-s bursts at 30 W), centrifuged (27,100g; 20 min) to remove remaining mitochondria, and the submitochondrial particles were finally collected by ultracentrifugation (82,000g; 30 min) and resuspended in 2 ml of the sucrose buffer. NADH:DQ oxidoreduction activity of SMPs was measured in the same buffer as for isolated complex I but without lipids and detergents and with additional 10 μM carbonyl cyanide *m*-chlorophenyl hydrazine (CCCP).

REFERENCES AND NOTES

- L. A. Sazanov, A giant molecular proton pump: Structure and mechanism of respiratory complex I. *Nat. Rev. Mol. Cell Biol.* **16**, 375–388 (2015). doi: [10.1038/nrm3997](https://doi.org/10.1038/nrm3997); pmid: [25991374](https://pubmed.ncbi.nlm.nih.gov/25991374/)
- J. Hirst, Mitochondrial complex I. *Annu. Rev. Biochem.* **82**, 551–575 (2013). doi: [10.1146/annurev-biochem-070511-103700](https://doi.org/10.1146/annurev-biochem-070511-103700); pmid: [23527692](https://pubmed.ncbi.nlm.nih.gov/23527692/)
- D. A. Stroud *et al.*, Accessory subunits are integral for assembly and function of human mitochondrial complex I. *Nature* **538**, 123–126 (2016). doi: [10.1038/nature19754](https://doi.org/10.1038/nature19754); pmid: [27626371](https://pubmed.ncbi.nlm.nih.gov/27626371/)
- R. Baradaran, J. M. Berrisford, G. S. Minhas, L. A. Sazanov, Crystal structure of the entire respiratory complex I. *Nature* **494**, 443–448 (2013). doi: [10.1038/nature11871](https://doi.org/10.1038/nature11871); pmid: [23417064](https://pubmed.ncbi.nlm.nih.gov/23417064/)
- R. G. Efremov, L. A. Sazanov, Structure of the membrane domain of respiratory complex I. *Nature* **476**, 414–420 (2011). doi: [10.1038/nature10330](https://doi.org/10.1038/nature10330); pmid: [21822288](https://pubmed.ncbi.nlm.nih.gov/21822288/)
- K. Fiedorczuk *et al.*, Atomic structure of the entire mammalian mitochondrial complex I. *Nature* **538**, 406–410 (2016). doi: [10.1038/nature19794](https://doi.org/10.1038/nature19794); pmid: [27595392](https://pubmed.ncbi.nlm.nih.gov/27595392/)
- J. Zhu, K. R. Vinothkumar, J. Hirst, Structure of mammalian respiratory complex I. *Nature* **536**, 354–358 (2016). doi: [10.1038/nature19095](https://doi.org/10.1038/nature19095); pmid: [27509854](https://pubmed.ncbi.nlm.nih.gov/27509854/)
- A. A. Agip *et al.*, Cryo-EM structures of complex I from mouse heart mitochondria in two biochemically defined states. *Nat. Struct. Mol. Biol.* **25**, 548–556 (2018). doi: [10.1038/s41594-018-0073-1](https://doi.org/10.1038/s41594-018-0073-1); pmid: [29915388](https://pubmed.ncbi.nlm.nih.gov/29915388/)
- J. A. Letts, K. Fiedorczuk, G. Degliesposti, M. Skehel, L. A. Sazanov, Structures of respiratory supercomplex I-III₂ reveal functional and conformational crosstalk. *Mol. Cell* **75**, 1131–1146.e6 (2019). doi: [10.1016/j.molcel.2019.07.022](https://doi.org/10.1016/j.molcel.2019.07.022); pmid: [31492636](https://pubmed.ncbi.nlm.nih.gov/31492636/)
- L. Euro, D. A. Bloch, M. Wikström, M. I. Verkhovskaya, M. Verkhovskaya, Electrostatic interactions between FeS clusters in NADH:ubiquinone oxidoreductase (complex I) from *Escherichia coli*. *Biochemistry* **47**, 3185–3193 (2008). doi: [10.1021/bi702063t](https://doi.org/10.1021/bi702063t); pmid: [18269245](https://pubmed.ncbi.nlm.nih.gov/18269245/)
- M. Wikström, V. Sharma, V. R. I. Kaila, J. P. Hosler, G. Hummer, New perspectives on proton pumping in cellular respiration. *Chem. Rev.* **115**, 2196–2221 (2015). doi: [10.1021/cr500448t](https://doi.org/10.1021/cr500448t); pmid: [25694135](https://pubmed.ncbi.nlm.nih.gov/25694135/)
- V. Sharma *et al.*, Redox-induced activation of the proton pump in the respiratory complex I. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11571–11576 (2015). doi: [10.1073/pnas.1503761112](https://doi.org/10.1073/pnas.1503761112); pmid: [26330610](https://pubmed.ncbi.nlm.nih.gov/26330610/)
- J. Wamau *et al.*, Redox-coupled quinone dynamics in the respiratory complex I. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E8413–E8420 (2018). doi: [10.1073/pnas.1805468115](https://doi.org/10.1073/pnas.1805468115); pmid: [30120126](https://pubmed.ncbi.nlm.nih.gov/30120126/)
- O. Haapanen, A. Djurabekova, V. Sharma, Role of second quinone binding site in proton pumping by respiratory complex I. *Front. Chem.* **7**, 221 (2019). doi: [10.3389/fchem.2019.00221](https://doi.org/10.3389/fchem.2019.00221); pmid: [31024903](https://pubmed.ncbi.nlm.nih.gov/31024903/)
- M. Verkhovskaya, D. A. Bloch, Energy-converting respiratory complex I: On the way to the molecular mechanism of the proton pump. *Int. J. Biochem. Cell Biol.* **45**, 491–511 (2013). doi: [10.1016/j.biocel.2012.08.024](https://doi.org/10.1016/j.biocel.2012.08.024); pmid: [22982742](https://pubmed.ncbi.nlm.nih.gov/22982742/)
- V. R. I. Kaila, Long-range proton-coupled electron transfer in biological energy conversion: Towards mechanistic understanding of respiratory complex I. *J. R. Soc. Interface* **15**, 20170916 (2018). doi: [10.1098/rsif.2017.0916](https://doi.org/10.1098/rsif.2017.0916); pmid: [29643224](https://pubmed.ncbi.nlm.nih.gov/29643224/)
- A. Di Luca, A. P. Gamiz-Hernandez, V. R. I. Kaila, Symmetry-related proton transfer pathways in respiratory complex I. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E6314–E6321 (2017). doi: [10.1073/pnas.1706278114](https://doi.org/10.1073/pnas.1706278114); pmid: [28716925](https://pubmed.ncbi.nlm.nih.gov/28716925/)
- J. N. Blaza, K. R. Vinothkumar, J. Hirst, Structure of the deactive state of mammalian respiratory complex I. *Structure* **26**, 312–319.e3 (2018). doi: [10.1016/j.str.2017.12.014](https://doi.org/10.1016/j.str.2017.12.014); pmid: [29395787](https://pubmed.ncbi.nlm.nih.gov/29395787/)
- S. Matsuzaki, K. M. Humphries, Selective inhibition of deactivated mitochondrial complex I by biguanides. *Biochemistry* **54**, 2011–2021 (2015). doi: [10.1021/bi501473h](https://doi.org/10.1021/bi501473h); pmid: [25719498](https://pubmed.ncbi.nlm.nih.gov/25719498/)
- J. A. Letts, G. Degliesposti, K. Fiedorczuk, M. Skehel, L. A. Sazanov, Purification of ovine respiratory complex I results in a highly active and stable preparation. *J. Biol. Chem.* **291**, 24657–24675 (2016). doi: [10.1074/jbc.M116.735142](https://doi.org/10.1074/jbc.M116.735142); pmid: [27672209](https://pubmed.ncbi.nlm.nih.gov/27672209/)
- M. H. Olsson, C. R. Søndergaard, M. Rostkowski, J. H. Jensen, PROPKA3: Consistent treatment of internal and surface residues in empirical pK_a predictions. *J. Chem. Theory Comput.* **7**, 525–537 (2011). doi: [10.1021/ct100578z](https://doi.org/10.1021/ct100578z); pmid: [26596171](https://pubmed.ncbi.nlm.nih.gov/26596171/)
- E. Galemou Yoga *et al.*, Mutations in a conserved loop in the PSST subunit of respiratory complex I affect ubiquinone binding and dynamics. *Biochim. Biophys. Acta* **1860**, 573–581 (2019). doi: [10.1016/j.bbabi.2019.06.006](https://doi.org/10.1016/j.bbabi.2019.06.006); pmid: [31226318](https://pubmed.ncbi.nlm.nih.gov/31226318/)
- V. G. Grivennikova, E. O. Maklashina, E. V. Gavrikova, A. D. Vinogradov, Interaction of the mitochondrial NADH-ubiquinone reductase with rotenone as related to the enzyme active/inactive transition. *Biochim. Biophys. Acta* **1319**, 223–232 (1997). doi: [10.1016/S0005-2728\(96\)00163-6](https://doi.org/10.1016/S0005-2728(96)00163-6); pmid: [9131045](https://pubmed.ncbi.nlm.nih.gov/9131045/)
- O. Haapanen, V. Sharma, A modeling and simulation perspective on the mechanism and function of respiratory complex I. *Biochim. Biophys. Acta* **1859**, 510–523 (2018). doi: [10.1016/j.bbabi.2018.04.001](https://doi.org/10.1016/j.bbabi.2018.04.001); pmid: [29660310](https://pubmed.ncbi.nlm.nih.gov/29660310/)
- P. G. Roberts, J. Hirst, The deactive form of respiratory complex I from mammalian mitochondria is a Na^+/H^+ antiporter. *J. Biol. Chem.* **287**, 34743–34751 (2012). doi: [10.1074/jbc.M112.384560](https://doi.org/10.1074/jbc.M112.384560); pmid: [22854968](https://pubmed.ncbi.nlm.nih.gov/22854968/)
- V. Zickermann *et al.*, Structural biology. Mechanistic insight from the crystal structure of mitochondrial complex I. *Science* **347**, 44–49 (2015). doi: [10.1126/science.1259859](https://doi.org/10.1126/science.1259859); pmid: [25554780](https://pubmed.ncbi.nlm.nih.gov/25554780/)
- A. B. Kotlyar, A. D. Vinogradov, Slow active/inactive transition of the mitochondrial NADH-ubiquinone reductase. *Biochim. Biophys. Acta* **1019**, 151–158 (1990). doi: [10.1016/0005-2728\(90\)90137-3](https://doi.org/10.1016/0005-2728(90)90137-3); pmid: [2119805](https://pubmed.ncbi.nlm.nih.gov/2119805/)
- K. Parey *et al.*, High-resolution cryo-EM structures of respiratory complex I: Mechanism, assembly, and disease. *Sci. Adv.* **5**, eaax9484 (2019). doi: [10.1126/sciadv.aax9484](https://doi.org/10.1126/sciadv.aax9484); pmid: [31844670](https://pubmed.ncbi.nlm.nih.gov/31844670/)
- K. Parey *et al.*, Cryo-EM structure of respiratory complex I at work. *eLife* **7**, e39213 (2018). doi: [10.7554/eLife.39213](https://doi.org/10.7554/eLife.39213); pmid: [30277212](https://pubmed.ncbi.nlm.nih.gov/30277212/)
- A. D. Vinogradov, V. G. Grivennikova, The mitochondrial complex I: Progress in understanding of catalytic properties. *IUBMB Life* **52**, 129–134 (2001). doi: [10.1080/15216540152845920](https://doi.org/10.1080/15216540152845920); pmid: [11798024](https://pubmed.ncbi.nlm.nih.gov/11798024/)
- J. Wang, On the appearance of carboxylates in electrostatic potential maps. *Protein Sci.* **26**, 396–402 (2017). doi: [10.1002/pro.3093](https://doi.org/10.1002/pro.3093); pmid: [27977901](https://pubmed.ncbi.nlm.nih.gov/27977901/)
- T. Hayashi, A. A. Stuchebrukhov, Quantum electron tunneling in respiratory complex I. *J. Phys. Chem. B* **115**, 5354–5364 (2011). doi: [10.1021/jp109410j](https://doi.org/10.1021/jp109410j); pmid: [21495666](https://pubmed.ncbi.nlm.nih.gov/21495666/)
- M. Schorb, I. Haberbosch, W. J. H. Hagen, Y. Schwab, D. N. Mastrorade, Software tools for automated transmission electron microscopy. *Nat. Methods* **16**, 471–477 (2019). doi: [10.1038/s41592-019-0396-9](https://doi.org/10.1038/s41592-019-0396-9); pmid: [31086343](https://pubmed.ncbi.nlm.nih.gov/31086343/)
- J. Zivanov *et al.*, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018). doi: [10.7554/eLife.42166](https://doi.org/10.7554/eLife.42166); pmid: [30412051](https://pubmed.ncbi.nlm.nih.gov/30412051/)
- S. Q. Zheng *et al.*, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017). doi: [10.1038/nmeth.4193](https://doi.org/10.1038/nmeth.4193); pmid: [28250466](https://pubmed.ncbi.nlm.nih.gov/28250466/)
- A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015). doi: [10.1016/j.jsb.2015.08.008](https://doi.org/10.1016/j.jsb.2015.08.008); pmid: [26278980](https://pubmed.ncbi.nlm.nih.gov/26278980/)
- J. Carroll, I. M. Fearnley, J. E. Walker, Definition of the mitochondrial proteome by measurement of molecular masses of membrane proteins. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 16170–16175 (2006). doi: [10.1073/pnas.0607719103](https://doi.org/10.1073/pnas.0607719103); pmid: [17060615](https://pubmed.ncbi.nlm.nih.gov/17060615/)
- P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010). doi: [10.1107/S0907444910007493](https://doi.org/10.1107/S0907444910007493); pmid: [20383002](https://pubmed.ncbi.nlm.nih.gov/20383002/)
- P. D. Adams *et al.*, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010). doi: [10.1107/S0907444909052925](https://doi.org/10.1107/S0907444909052925); pmid: [20124702](https://pubmed.ncbi.nlm.nih.gov/20124702/)
- B. A. Barad *et al.*, EMRinger: Side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat. Methods* **12**, 943–946 (2015). doi: [10.1038/nmeth.3541](https://doi.org/10.1038/nmeth.3541); pmid: [26280328](https://pubmed.ncbi.nlm.nih.gov/26280328/)
- V. B. Chen *et al.*, MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010). doi: [10.1107/S0907444909042073](https://doi.org/10.1107/S0907444909042073); pmid: [20057044](https://pubmed.ncbi.nlm.nih.gov/20057044/)

ACKNOWLEDGMENTS

We thank J. Novacek (CEITEC Brno) and V.-V. Hodirnau (IST Austria) for their help with collecting cryo-EM datasets. We thank the IST Life Science and Electron Microscopy Facilities for providing equipment. **Funding:** This work has been supported by iNEXT, project number 653706, funded by the Horizon 2020 program of the European Union. This article reflects only the authors' view, and the European Commission is not responsible for any use that may be made of the information it contains. CIISB research infrastructure project LM2015043 funded by MEYS CR is gratefully acknowledged for the financial support of the measurements at the CF Cryo-electron Microscopy and Tomography CEITEC MU. This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement no. 665385. **Author contributions:** D.K. purified complex I, performed biochemical assays, prepared cryo-EM grids, acquired and processed EM data, built and analyzed the atomic models, and wrote the initial draft of the manuscript. L.A.S. designed and supervised the project, acquired funding, analyzed data, and revised the manuscript. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** Atomic models resulting from this study were deposited in the PDB under the following accession codes: 6ZKO (native closed), 6ZKP (native open1), 6ZKQ (native open2), 6ZKR (native open3), 6ZKH (NADH closed), 6ZKH (NADH open1), 6ZKI (NADH open2), 6ZKJ (NADH open3), 6ZKS (deactive open1), 6ZKT (deactive open2), 6ZKU (deactive open3), 6ZKV (deactive open4), 6ZKK (rotenone closed), 6ZKL (rotenone open1), 6ZKM (rotenone open2), 6ZKN (rotenone open3), 6ZKC (turnover closed), 6ZKD (turnover open1), 6ZKE (turnover open2), 6ZKF (turnover open3), 6ZK9 (turnover PA-open), 6ZKA (turnover MD-open), and 6ZKB (turnover MD-closed). Cryo-EM maps used to generate these models were deposited in the EMD: EMD-11256 (native closed), EMD-11257 (native open1), EMD-11258 (native open2), EMD-11259 (native open3), EMD-11248 (NADH closed), EMD-11249 (NADH open1), EMD-11250 (NADH open2), EMD-11251 (NADH open3), EMD-11260 (deactive open1), EMD-11261 (deactive open2), EMD-11262 (deactive open3), EMD-11263 (deactive open4), EMD-11252 (rotenone closed), EMD-11253 (rotenone open1), EMD-11254 (rotenone open2), EMD-11255 (rotenone open3), EMD-11244 (turnover closed), EMD-11245 (turnover open1), EMD-11246 (turnover open2), EMD-11247 (turnover open3), EMD-11241 (turnover PA-open), EMD-11242 (turnover MD-open), and EMD-11243 (turnover MD-closed).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/eabc4209/suppl/DC1
Supplementary Text
Figs. S1 to S11
Tables S1 to S7
References (42–60)
MDAR Reproducibility Checklist
Movies S1 and S2

[View/request a protocol for this paper from Bio-protocol.](#)

23 April 2020; accepted 8 September 2020
Published online 24 September 2020
[10.1126/science.abc4209](https://doi.org/10.1126/science.abc4209)

RESEARCH ARTICLE SUMMARY

HYBRID PEROVSKITES

Atomic-scale microstructure of metal halide perovskite

Mathias Uller Rothmann, Judy S. Kim, Juliane Borchert, Kilian B. Lohmann, Colum M. O'Leary, Alex A. Sheader, Laura Clark, Henry J. Snaith, Michael B. Johnston, Peter D. Nellist*, Laura M. Herz*

INTRODUCTION: Hybrid metal halide perovskites are highly favorable materials for efficient photovoltaic and optoelectronic applications. The mechanisms behind their impressive performance have yet to be fully understood, but they likely depend on atomic-level properties that may be unique to these perovskites. Atomic-resolution transmission electron microscopy is well suited to provide new insights but is challenging because of the highly beam-sensitive nature of hybrid perovskites.

RATIONALE: We used low-dose scanning transmission electron microscopy (STEM) imaging to determine the microstructure of thin hybrid perovskite films. Thermally evaporated thin films of formamidinium and methylammonium lead triiodide (FAPbI₃ and MAPbI₃, respectively) were examined on ultrathin carbon-coated copper TEM grids to reveal the nature of boundaries, defects, and decomposition pathways.

RESULTS: Using low-dose low-angle annular dark field (LAADF) STEM imaging, we obtained atomic-resolution micrographs of FAPbI₃ films in the cubic phase. We found that prolonged electron irradiation leads to a loss of FA⁺ ions, which initially causes the perovskite structure to change to a partially FA⁺-depleted

but ordered perovskite lattice, apparent as light-and-dark checkered patterns in STEM images. Further electron beam exposure leads to the expected deterioration to PbI₂ as the final decomposition product. We propose that the observed intermediate checkered pattern is triggered by an initially random, beam-induced loss of FA⁺, followed by subsequent reordering of FA⁺ ions. The discovery of this intermediate structure explains why the perovskite structure can sustain significant deviations from stoichiometry and recovers remarkably well from damage.

We further revealed the atomic arrangement at interfaces within the hybrid perovskite films. We found that PbI₂ precursor remnants commonly encountered in hybrid perovskite films readily and seamlessly intergrow with the FAPbI₃ and MAPbI₃ lattice and can distort from their bulk hexagonal structure to form a surprisingly coherent transition boundary, exhibiting low lattice misfit and strain. We observed PbI₂ domains that nearly perfectly follow the surrounding perovskite structure and orientation, which suggests that PbI₂ may seed perovskite growth. These observations help to explain why the presence of excess PbI₂ tends not to impede solar cell performance.

Images of FAPbI₃ grain boundaries further revealed that the long-range perovskite struc-

ture is preserved up to the grain boundaries, where sharp interfaces are generally present, without any obvious preferred orientation. Near-120° triple boundaries are most commonly observed at the intersection of three grains, which we generally found to be crystallographically continuous and associated with minimal lattice distortion.

Finally, we identified the nature of defects, dislocations, and stacking faults in the FAPbI₃ lattice. We discovered dislocations that are dissociated in a direction perpendicular to their glide plane (climb-dissociated), aligned point defects in the form of vacancies on the Pb-I sublattice, and stacking faults corresponding to a shift of half a unit cell, connecting Pb-I columns with I⁻ columns rather than with FA⁺ columns.

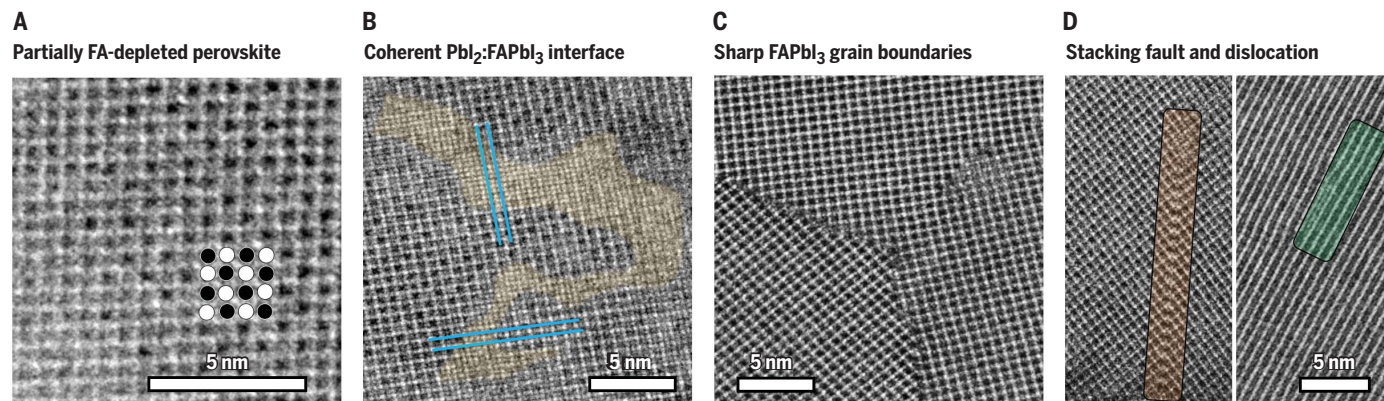
CONCLUSION: Our findings provide an atomic-level understanding of the technologically important class of hybrid lead halide perovskites, revealing several mechanisms that underpin their remarkable performance. The highly adaptive nature of the perovskite structure upon organic cation loss yields exceptional regenerative properties of partly degraded material. The observation of coherent perovskite interfaces with PbI₂ explains the barely diminished optoelectronic performance upon such precursor inclusions, while sharp interfaces between perovskite grains grant a benign nature. Such atomically localized information enables the targeted design of methods to eliminate defects and optimize interfaces in these materials. ■

The list of author affiliations is available in the full article online.

*Corresponding author. Email: laura.herz@physics.ox.ac.uk (L.M.H.); peter.nellist@materials.ox.ac.uk (P.D.N.)

Cite this article as M. U. Rothmann *et al.*, *Science* **370**, eabb5940 (2020). DOI: 10.1126/science.abb5940

READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.abb5940>



Atomic-resolution STEM images of FAPbI₃. (A) Checkered intensity pattern formed upon electron beam exposure with ~200 e Å⁻². (B) Native intergrowth between PbI₂ (shaded yellow) and FAPbI₃ within the vapor-deposited film. (C) Abrupt grain boundaries meeting at a near-120° triple junction, showing no amorphous or intergranular phase. (D) Intrinsic stacking fault (orange rectangle, left) and climb-dissociated dislocation (green rectangle, right). Scale bars, 5 nm.

RESEARCH ARTICLE

HYBRID PEROVSKITES

Atomic-scale microstructure of metal halide perovskite

Mathias Uller Rothmann¹, Judy S. Kim^{2,3,4}, Juliane Borchert¹, Kilian B. Lohmann¹, Colum M. O'Leary², Alex A. Sheader², Laura Clark², Henry J. Snaith¹, Michael B. Johnston¹, Peter D. Nellist^{2*}, Laura M. Herz^{1*}

Hybrid organic-inorganic perovskites have high potential as materials for solar energy applications, but their microscopic properties are still not well understood. Atomic-resolution scanning transmission electron microscopy has provided invaluable insights for many crystalline solar cell materials, and we used this method to successfully image formamidinium lead triiodide $[\text{CH}(\text{NH}_2)_2\text{PbI}_3]$ thin films with a low dose of electron irradiation. Such images reveal a highly ordered atomic arrangement of sharp grain boundaries and coherent perovskite/ PbI_2 interfaces, with a striking absence of long-range disorder in the crystal. We found that beam-induced degradation of the perovskite leads to an initial loss of formamidinium $[\text{CH}(\text{NH}_2)_2^+]$ ions, leaving behind a partially unoccupied perovskite lattice, which explains the unusual regenerative properties of these materials. We further observed aligned point defects and climb-dissociated dislocations. Our findings thus provide an atomic-level understanding of technologically important lead halide perovskites.

Many of the advances in the performance of hybrid metal halide perovskite (MHP) solar cells have resulted from improvements in device architecture and material composition studied on a mesoscopic to macroscopic scale (1–7), and changes in material properties have usually been inferred from relatively “large area” probes (8–11). Although these probes can resolve properties on the scale of a few hundred nanometers to a few millimeters, the mechanisms behind the impressive performance of hybrid perovskites have yet to be fully understood. However, the atomic-level nature of their properties can be inferred from the gradual rather than abrupt decrease in photovoltaic performance of MHPs over time (12), their ability to readily regenerate when exposed to gaseous iodo-organic precursors (13), and the counterintuitively beneficial effects of adding a small excess of PbI_2 [2% (14, 15) to 7.5% (16)] to the MHP precursor solution. Similarly, substantial efforts have been made to describe the properties of grain boundaries in MHP thin films by means of light and scanning probes (17–19), with highly inconsistent results (11, 17). Recent work has shown that a smaller spread of grain boundary orientations increases the photoluminescence intensity locally (20), but the exact atomic nature of grain boundaries has still not been described. Even the crystal phase

of the technologically important $\text{CH}(\text{NH}_2)_2\text{PbI}_3$ (formamidinium lead triiodide, FAPbI₃) MHP is still uncertain (21). Therefore, truly atomic-resolution imaging of MHPs has great potential to provide urgently needed understanding for a plethora of puzzling observations already established for perovskite devices, which in turn would lead to a further increase in their efficiency. In addition, grain boundaries have been postulated to have a substantial impact on the long-term stability of MHPs (22, 23), yet their exact nature is still mostly unknown.

Atomic-resolution transmission electron microscopy (TEM) may be of use in clarifying some of these aspects of MHPs, as it has for other crystalline solar cell materials (24, 25), but the highly beam-sensitive nature of hybrid perovskites makes high-resolution electron microscopy extremely challenging. Although fully inorganic photoactive perovskites have been successfully imaged (26), the few studies observing the pristine hybrid perovskite phase are based on nanostructures (27, 28), and no studies of pristine, native hybrid perovskite thin films have been published to date. Unstable under electron beams, MHPs rapidly change into a more stable PbI_2 phase under even mild electron irradiation (29, 30), which has resulted in striking differences in the crystallography reported by electron microscope and x-ray studies (11).

Imaging the perovskite structure

We used low-dose scanning transmission electron microscopy (STEM) to obtain images of pristine FAPbI₃, because it has been shown for other materials to produce less damage with the same total electron fluence (dose)

(31). To resolve both “near-stoichiometric” polycrystalline FAPbI₃ films and those formed with substantial excess PbI_2 precursor (“ PbI_2 -rich”), we prepared two different types of thermally evaporated materials on ultrathin carbon-coated copper TEM grids: thinner 30-nm FAPbI₃ films close to stoichiometry [FA:Pb:I ratios close to 1:1:3, containing a small excess of PbI_2 (as used to optimize for solar cell applications) (32)] and thicker 90-nm films containing substantial excess PbI_2 . To examine whether similar atomic-scale microstructures were also observable for methylammonium lead triiodide (MAPbI₃), we also deposited 30-nm MAPbI₃ films containing substantial excess PbI_2 . Full details of the sample preparation procedures and characterization are provided in the supplementary materials and fig. S1. We have reported thin films of FAPbI₃ and MAPbI₃ deposited by us from such dual-source vapor-deposition protocols to achieve photovoltaic power conversion efficiencies of 14.2 to 18.2% when incorporated into suitable device architectures (32, 33). STEM samples were annealed according to protocols previously developed to optimize photovoltaic performance.

In Fig. 1, we show a low-angle annular dark field (LAADF, inner radius of 33 mrad; see materials and methods for full low-dose image acquisition conditions) STEM micrograph of a 90-nm FAPbI₃ film containing excess PbI_2 . LAADF STEM is similar to high-angle annular dark field (HAADF) STEM in that it detects scattered electrons, making it sensitive to atomic number (Z) variations, but has the advantage of collecting a higher fraction of the scattered electrons, leading to a higher atomic detection efficiency. In addition, LAADF STEM allows efficient imaging of lower-Z elements, which scatter at lower angles (34). To improve the contrast of this image of a pristine crystal, we applied a Bragg filter, as described in detail in materials and methods and in fig. S2.

The lower-magnification image in Fig. 1A reveals an atomic lattice pattern consistent with a cubic structure observed along the $\langle 001 \rangle$ orientation. The magnified image in Fig. 1B clearly shows the atomic-level crystal structure corresponding to the cubic structure of room-temperature FAPbI₃, displayed schematically in Fig. 1C (21). We note that the preferred alignment of growth along the $[100]$ axis (“face-up”) is generally observed by us in these dual-source vapor-deposited films and is common for lead halide perovskite deposited across a variety of alternative processing techniques (35). We also note that an important advantage of LAADF imaging in this case is that, similar to HAADF imaging, the column intensities retain a monotonic dependence on atomic number. For completeness, a comparison with a HAADF image, acquired as described in materials and methods, is shown in fig. S3.

¹Department of Physics, Clarendon Laboratory, University of Oxford, Oxford OX1 3PU, UK. ²Department of Materials, University of Oxford, Oxford OX1 3PH, UK. ³ePSIC, Diamond Light Source, Harwell, Didcot OX11 0DE, UK. ⁴Rosalind Franklin Institute, Harwell, Didcot OX11 0QS, UK.

*Corresponding author. Email: laura.herz@physics.ox.ac.uk (L.M.H.); peter.nellist@materials.ox.ac.uk (P.D.N.)

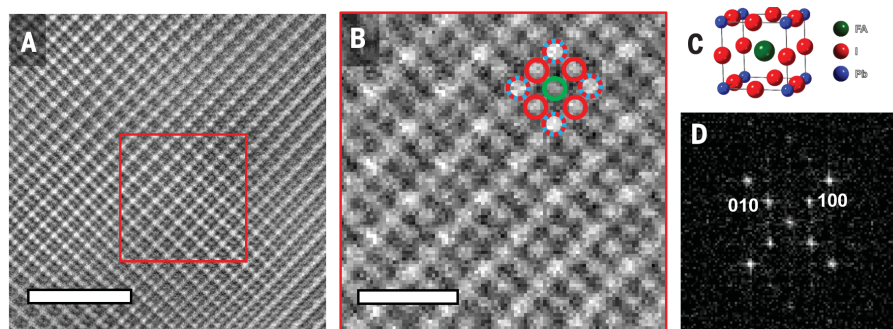


Fig. 1. Atomic-resolution imaging of the metal halide perovskite structure. (A) Bragg-filtered atomic-resolution LAADF STEM micrograph of a pristine 90-nm-thick, (001)-oriented FAPbI₃ film fabricated with excess PbI₂. A distinct pattern in the contrast of the individual atomic columns is clearly visible. (B) Magnification of the area within the red square in (A). Here, the variations in composition appear as differences in contrast; the dashed blue and red circles show the Pb/I columns, the red circles show the pure I columns, and the green circle shows the FA⁺ columns. (C) The atomic unit cell represented by the circles in (B). (D) The Fourier transformation (FT) confirms the presence of the {010} and {100} planes, which are often missing from atomic-resolution studies of beam-damaged FAPbI₃. Scale bars, 5 nm (A), 1 nm (B). The total dose used for image acquisition in (A) and (B) is ~100 e Å⁻². See fig. S2 for a comparison of the full image for raw (unfiltered), Butterworth-filtered, and Bragg-filtered data.

Three levels of intensity are observable in the LAADF images. The brightest corresponds to the corners of the cubic unit cell (dashed blue and red circles in Fig. 1B), containing alternating lead and iodine, which are both strongly scattering elements. The second brightest intensity corresponds to the edges of the unit cell containing only iodine, illustrated by the red circles in Fig. 1B. In the *z* viewing direction, the corners make up a total of one lead atom and one iodine per unit cell, whereas the edge position only contains one iodine, resulting in less scattering and a less bright image intensity. The central, lowest-intensity position corresponds to the FA⁺ ions, which scatter the least because of their organic nature and singular presence in the center of the unit cell (34). Intensity line profiles and a multislice simulation verifying the differences in contrast are provided in fig. S4.

To support structure identification, we performed a Fourier transformation (FT; Fig. 1D) on the area marked in Fig. 1A, which clearly identifies the presence of the {100} and {010} planes along with the {200} and {020} planes. Of the three proposed room-temperature crystal phases of FAPbI₃ [cubic, trigonal, and hexagonal (2)], the spots corresponding to the {100} and {010} planes are unique to the cubic perovskite structure and cannot correspond to either of the other two. This strongly indicates that the room-temperature phase of FAPbI₃ is cubic in the *Pm3m* space group.

We note that the {100} and {010} FT spots evident in Fig. 1D have been absent in previous atomic-resolution STEM studies of both FAPbI₃ and MAPbI₃, with only the supposed {200} and {020} or other higher-index spots still visible (36–39). Broad-beam diffraction

studies carried out on MAPbI₃ have explained this apparent absence of lower-order diffraction signatures as the result of structural damage caused by high electron irradiation dosage (29,30). Thus, the spots attributed to the second-order diffraction off the perovskite lattice may instead derive from the PbI₂ formed as a decomposition product under intense electron irradiation (40,41).

Electron-induced decomposition

Using a suitably chosen irradiation dosage, we were able to observe the gradual decomposition of FAPbI₃ with electron beam exposure. Figure 2A shows a lower-magnification Butterworth-filtered LAADF image of a pristine grain boundary in a 30-nm-thick FAPbI₃ film close to stoichiometry on a carbon-coated TEM grid after a single scan of the electron beam, with a total dose of ~66 e Å⁻². The subsequent scans were also recorded (Fig. 2, B to D). After three scans (~200 e Å⁻²), the contrast within the perovskite structure has visibly changed, and after nine scans (~600 e Å⁻²), the perovskite structure is completely depleted and replaced by a crystal structure with uniform column intensities (Fig. 2C). Further extensive exposure (~1333 e Å⁻²) slightly alters this phase so that it is very similar to 2H-hexagonal PbI₂ (Fig. 2D). This final structure is highly stable, does not change substantially after further exposure, and corresponds well to the PbI₂ phase observed by diffraction in broad-beam TEM studies (29). A full crystallographic determination of the phase seen in Fig. 2C is beyond the scope of this study, but fig. S10 shows how a small distortion of hexagonal PbI₂ can form a structure that projects as a square lattice and forms a coherent epitaxial

interface with FAPbI₃ as observed in the image data. The FTs in Fig. 2, E to H, show the degradation process on a crystallographic level: For the pristine perovskite, diffraction spots corresponding to the {100} plane family are initially clearly visible (Fig. 2E). Additional spots appear after brief exposure (Fig. 2F, orange circles), likely corresponding to the ordered loss of formamidinium ions discussed below. Continued exposure fully destroys the perovskite structure, transforming it into electron irradiation-stable PbI₂ concomitant with the disappearance of the {100} plane family of the original perovskite (Fig. 2G) and finally into a structure similar to that of 2H-hexagonal PbI₂ (Fig. 2H).

Nanoscale properties of metal halide perovskites

The implementation of a low electron dose, <200 e Å⁻², inevitably results in a low signal-to-noise ratio when images are recorded. We find that combining low-dose LAADF imaging with simple Butterworth and Bragg filters allows us to investigate pristine and slightly damaged FAPbI₃ thin films with atomic resolution, observing phenomena invisible to other techniques. The details of each filtering technique are provided in materials and methods and in fig. S5. Our findings are based on hundreds of STEM images taken across seven or eight thin-film samples prepared for each stoichiometry; the images we present were carefully chosen to be representative of each material's microstructural features.

In Fig. 3, A to C, we show LAADF images of polycrystalline perovskite films fabricated with excess PbI₂. We observe a surprisingly coherent transition boundary between remnant precursor PbI₂ (highlighted in light orange) and the FAPbI₃ grains, with undetectable lattice misfit. In addition, a light and dark checkered pattern becomes clearly visible in images of FAPbI₃ after mild radiation damage, as highlighted schematically in the lower right of Fig. 3A through white and black circles and seen surrounding the PbI₂ region in Fig. 3D. We propose that the observed checkered pattern in the central position of the FAPbI₃ unit cell is triggered by an initially random, beam-induced loss of FA⁺, likely due to dissociation of the FA⁺ by radiolysis (42,43), potentially complemented by loss of I⁻ through knock-on (displacement) effects (41). We postulate that the remaining FA⁺ ions then migrate to form an ordered structure with an associated energy reduction that helps to stabilize the FA⁺-deficient structure. Such migration of FA⁺ ions within the perovskite structure after electron irradiation then results in the observed checkered pattern as an intermediate equilibrium (29). A degradation mechanism assuming solely the ordered loss of I⁻ was previously proposed on the basis of analysis of diffraction

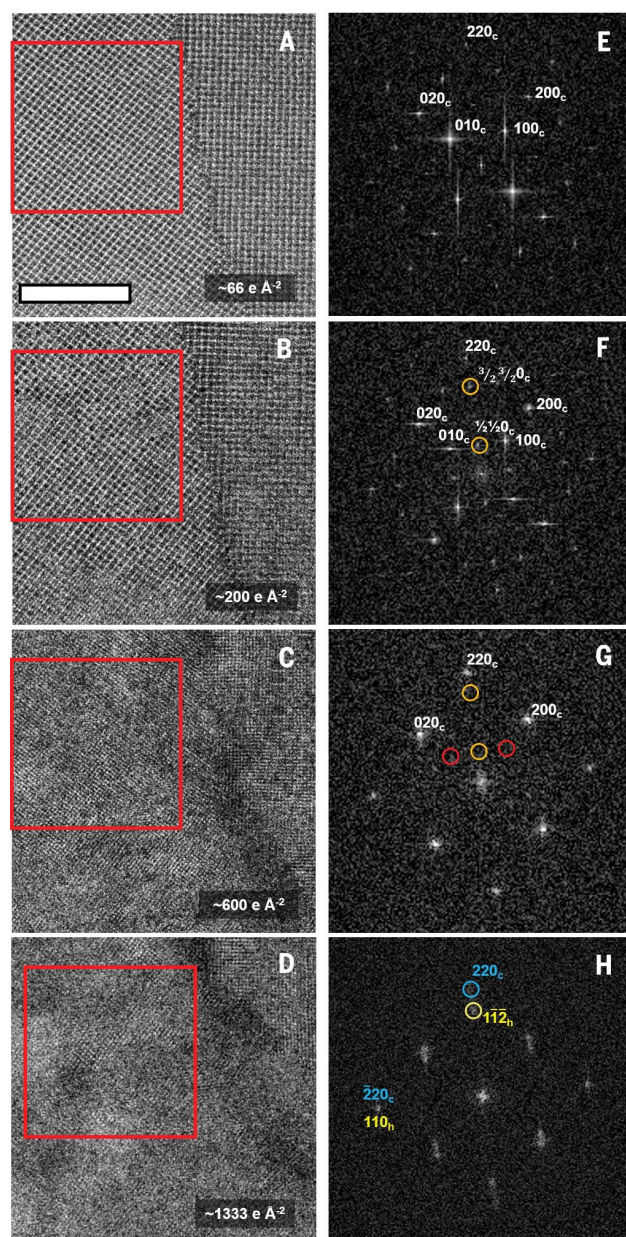


Fig. 2. Damage mechanism of FAPbI₃ under a scanning electron beam observed in a 30-nm-thick film prepared near the correct stoichiometry. (A) The first scan ($\sim 66 \text{ e } \text{\AA}^{-2}$) shows the pristine perovskite structure and a grain boundary. (B) After three scans ($\sim 200 \text{ e } \text{\AA}^{-2}$), changes in the crystal structure are apparent in the LAADF image but the grain boundary is still distinguishable. (C) After nine scans ($\sim 600 \text{ e } \text{\AA}^{-2}$), the perovskite structure has fully disappeared and a new crystalline phase has replaced it. Different crystal orientations are visible, but there is no clear grain boundary. (D) After 20 scans ($\sim 1333 \text{ e } \text{\AA}^{-2}$), the angles of the crystal planes change slightly, indicating the complete transformation into 2H-PbI₂. (E to H) The Fourier transformations corresponding to the areas in the red squares in (A) to (D), respectively, confirm the dynamic radiation-induced transformation. Initially, the FT in (E) reveals a pristine cubic structure, but additional spots appear after brief exposure, circled in orange in (F). After further exposure, the additional spots disappear again, along with the $\{100\}$ spots, whose expected location is marked by red circles in (G). Finally, after extended exposure, the FT contains two sets of spots: one set that can be indexed with the degraded cubic phase and another corresponding to the hexagonal PbI₂ phase, marked with a blue and a yellow circle in (H), respectively. The $\bar{2}20_c$ (2.2493 Å) and 110_h (2.2775 Å) spots overlap, indicating an intergrowth of the two phases. Indices marked with a subscript “c” are from the epitaxial FAPbI₃ phase; those marked with a subscript “h” are from the hexagonal PbI₂ phase. Note that the material is no longer in its pristine perovskite phase after the initial scan and the “c” subscript only serves to illustrate the corresponding plane in the pristine phase. Scale bar, 10 nm [(A) to (D)].

patterns alone (30); however, we note that this approach struggles to distinguish between ordered loss of FA⁺ and I[−]. Although vacancy ordering is clearly apparent on the FA⁺ sub-lattice, there is no equivalent effect evident for the I or Pb/I columns (see fig. S8A for intensity line profiles corresponding to Fig. 3D). Our direct interpretation of real-space images is thus able to reveal that the degradation is strongly driven by the loss, and likely subsequent reordering, of FA⁺ ions.

As further elaborated below, these microstructural features help to explain important phenomena pertinent to perovskite solar cells. Specifically, small amounts of PbI₂ are not detrimental to the solar cell performance (14–16) because perfect crystallographic alignment between the PbI₂ and FAPbI₃ perovskite occurs at the material interface, and no crystal defects are apparent in this region. The observation of the checkered pattern further explains why the perovskite structure can sustain substantial deviations from stoichiometry and recover remarkably well from damage.

We note that at cryogenic temperatures, MAPbI₃ has been shown to amorphize under electron irradiation rather than inducing the gradual, ordered restructuring of the crystal observed at room temperature (29). Rothmann *et al.* argued that such amorphization at low temperatures is the result of impeded ion mobility (29), further supporting our notion here that ion migration is a key factor in mitigating the beam-induced damage through atomic rearrangement.

We further examined the nature of the native interfaces formed between FAPbI₃ and its precursor PbI₂ (Fig. 3). Small amounts of PbI₂ surprisingly do not impede solar cell performance (14–16), but the reasons for such benign influence are still a matter of debate. We were able to reveal the presence of a low-misfit, low-lattice strain interface between FAPbI₃ and PbI₂ by examining the intergrowth between FAPbI₃ and PbI₂ in a film grown specifically under PbI₂-rich conditions (see materials and methods). In Fig. 3B, a nanoscopic domain of PbI₂ with a continuous transition into FAPbI₃ can be seen in the upper right corner, with the continuous transition likely arising because of the inclination of the interface to the beam direction. The PbI₂ domain perfectly follows the surrounding perovskite structure and orientation. The lack of evidence for misfit dislocations (see fig. S9 for more examples) shows that the interface is highly coherent and epitaxial. We note that the known 2H hexagonal phase of PbI₂ could not be coherent with the cubic FAPbI₃ in all directions in the interface, but that a modified 2H structure described in fig. S10 can be coherent with the perovskite and is possibly stabilized by its interface with FAPbI₃ (see fig. S10 for an atomic model), as evidenced by the

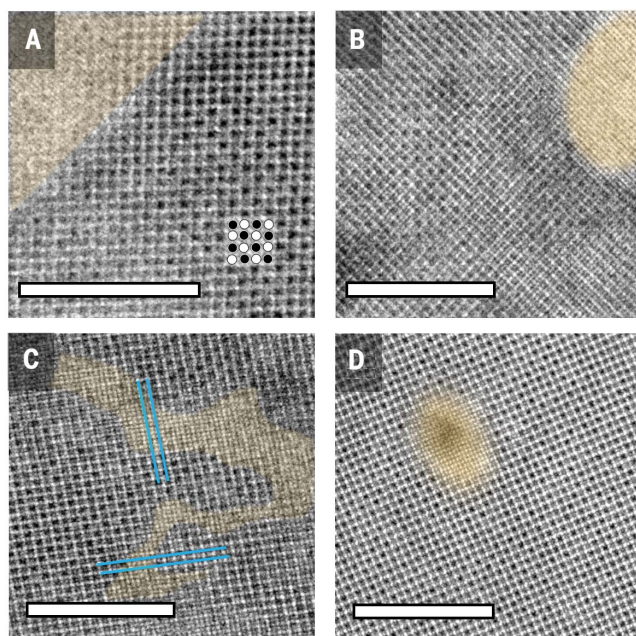


Fig. 3. Atomic-resolution LAADF micrographs of FAPbI₃ thin films showing undistorted transitions between PbI₂ and FAPbI₃, as well as a checkered intensity after mild radiation damage. PbI₂ is highlighted in light orange. (A to C) Native intergrowth between PbI₂ and FAPbI₃ formed during the deposition process in 90-nm-thick FAPbI₃ films fabricated with excess PbI₂. (D) A local area of PbI₂ formation induced by focusing an electron beam on a near-stoichiometric FAPbI₃ thin film. The highlighted area in (D) corresponds to the damage phase displayed in Fig. 2C; the surrounding area corresponds to that shown in Fig. 2B. The transitions do not induce any measurable phase boundary misfit distortions in the crystal lattice. The nanoscopic PbI₂ domain surrounded by FAPbI₃ in (C) indicates that PbI₂ can act as a seed for growth of FAPbI₃ crystals. The parallel blue lines in (C) begin in a Pb/I lattice site and end in an FA⁺ lattice site after crossing a PbI₂ domain, indicating that areas on either side of the PbI₂ are crystallographically separated and that the presence of PbI₂ is not beam-induced. A checkered pattern in the intensity of the FA⁺ lattice site is visible in (A), (C), and (D), schematically shown as black and white circles in (A). This pattern is indicative of a highly ordered restructuring of the moieties in the crystal after mild electron radiation damage. The Fourier transform of (A) is shown in fig. S6. All micrographs were Butterworth-filtered. Scale bars, 10 nm. The same figure is shown without the overlaid color graphics and with Fourier transforms in fig. S7.

square nature of the observed lattice. The weak van der Waals binding between the strongly ionically bound PbI₂ sheets enables the existence of many different PbI₂ polytypes (44), and a small shift between the layers of the 2H structure does allow coherent interfaces to form with FAPbI₃ (fig. S10) (44). This coherence also indicates that the PbI₂ has either acted as a seed for the surrounding perovskite or has grown out of the perovskite because of a local lack of FAI.

Further evidence for PbI₂ seeding is revealed in Fig. 3C, where irregularly shaped domains of PbI₂ can be seen to interweave with the perovskite lattice. Here, we observe a shift of half a unit cell in the perovskite crystal across the PbI₂ domain, as evidenced by the two sets of parallel blue lines. These lines, oriented in the (100) direction, begin in an FA⁺ lattice site on one side of the PbI₂ domain and end in a Pb/I lattice site on the other side. These columns are not found together on the same {100} plane in a perfect

crystal, so there must be a relative shift of half a unit cell on either side of the PbI₂ region. The LAADF signal intensity line profile displayed in fig. S8B further illustrates this shift across the PbI₂ region. It is unlikely for degradation of FAPbI₃ to have resulted in the growth of such an epitaxial PbI₂ phase that exhibits a half-unit cell distortion of the FAPbI₃ matrix, because the low misfit of the interface would not give a driving force for such a distortion of the FAPbI₃ structure. We can therefore conclude that the perovskite has grown out of the PbI₂ and not vice versa. Despite this shift, geometric phase analysis strain mapping (fig. S11) shows an absence of substantial strain associated with this perovskite/PbI₂ interface. Shown in fig. S12, A and B, are examples of the half-unit cell shift and the PbI₂ intergrowth, respectively, similarly occurring in pristine perovskite that has not suffered electron beam damage. These observations provide further evidence that the presence of perovskite/PbI₂ intergrowth is not caused by radiation-

induced damage, but is an intrinsic property of the material, present from the point of crystal growth. To verify that this approach is also applicable to other hybrid perovskites, we also imaged PbI₂-rich MAPbI₃ and found similar crystallographic properties (fig. S13).

Grain boundaries and crystal defects

We further revealed the nature of the boundaries between individual FAPbI₃ grains through high-resolution LAADF imaging. Such grain boundaries are of particular interest for perovskite solar cells because they have been proposed to act both as charge recombination centers (18, 45, 46) and as conduction pathways (16). In addition, grain boundaries have been shown to be intimately linked with the long-term stability of MHPs such as FAPbI₃ (22) and the photostability of mixed-halide perovskites toward halide segregation (23). Therefore, much effort has been devoted to passivating their negative impact and promoting their positive influence through appropriate treatments and film growth (11, 16, 19, 47–50). Although many techniques have been used to infer properties of the grain boundaries in photoactive perovskites (11, 17), reliable imaging with atomic resolution has remained elusive, which has led to some uncertainty about the inherent nature of the boundaries (17, 50). A full understanding of the crystallographic nature of grain boundaries is likely to explain a plethora of observations already well established in the field of perovskite solar cell devices, leading to further improvements in performance.

Butterworth-filtered LAADF images (Fig. 4, A to C) illustrate a representative set of grain boundaries in an evaporated thin film of FAPbI₃, with several striking features. First, the grains are highly crystalline adjacent to the boundaries, showing that the presence of the boundaries does not disrupt the long-range crystal quality of the surrounding perovskite. Second, the grain boundaries themselves are clearly defined, with the perovskite structure remaining unchanged to the interface, with no obvious preferred interfacial orientation and typically no amorphous material visible. These two points are illustrated by the strain maps displayed in fig. S16. Although an exhaustive description of the strain associated with all the potential boundaries is beyond the scope of this study, we have mapped the strain associated with the low-angle grain boundary seen in the lower left of Fig. 4B. We find that the strain is negligible to within one or two unit cells away from the boundary. Along the boundary itself, the strain exhibits a periodic behavior with alternating sign, which arises from the periodic array of edge dislocations forming the grain boundary. The *yy* strain (lying close to perpendicular to the dislocation Burgers vectors) is much lower at the boundary.

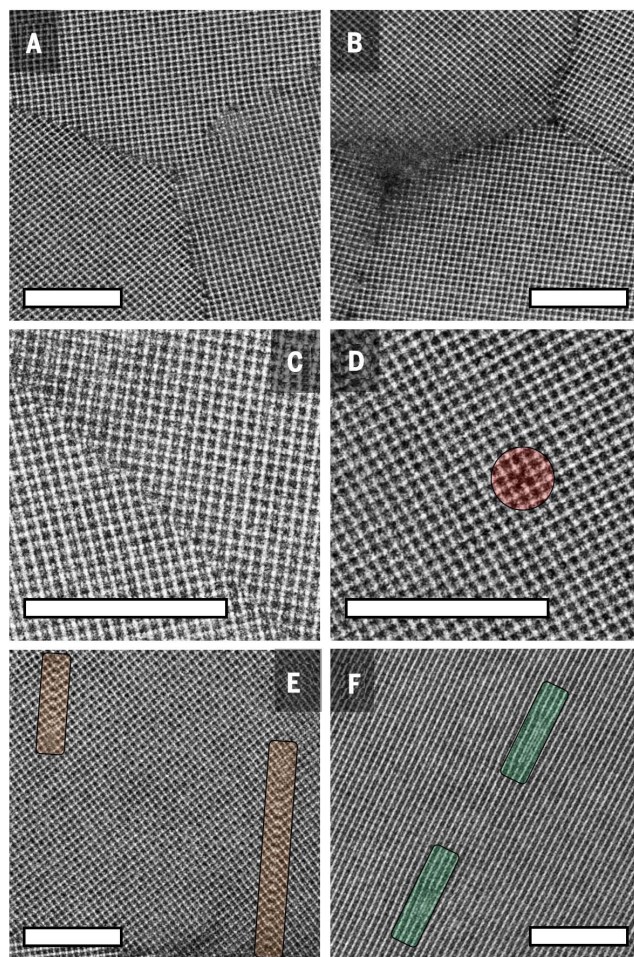


Fig. 4. Butterworth-filtered atomic-resolution LAADF micrographs of representative FAPbI₃ grain boundaries and crystal defects as seen in a 30-nm-thick film prepared close to the correct stoichiometry. (A) Abrupt grain boundaries meeting at a triple junction showing no amorphous or intergranular phase. (B) A triple junction showing signs of amorphization in the intersection at left and a line of vacancies at the intersection at right. The boundaries appear randomly oriented, with the presence of a curved boundary visible in the upper right corner. The low-angle grain boundary at the lower left is formed by a series of $\langle 100 \rangle$ edge dislocations. (C) A straight section of a grain boundary with virtually no empty lattice sites and very little distortion across the boundary. (D) Aligned vacancy defects on the I/Pb lattice site are seen in the red circle. (E) Intrinsic stacking faults are seen in the orange rectangles with a stacking fault vector of $\frac{1}{2}\langle 110 \rangle$. (F) Edge dislocations are seen in the green rectangles with a Burgers vector of $\langle 100 \rangle$. The dislocations can be seen to have dissociated in the climb direction, forming two partial dislocations with Burgers vector $\frac{1}{2}\langle 100 \rangle$ when projected into the image plane, resulting in a fault lying perpendicular to the glide plane. Scale bars, 10 nm. The figure is shown without the colored highlights in fig. S15.

Near-120° triple boundaries are commonly observed at the intersection of three grains. Figure 4A shows a typical image of a triple boundary in a FAPbI₃ film, which we generally found to be crystallographically continuous, with minimal lattice distortion and no defects extending away from the grain boundary plane associated with them. Across the STEM micrographs examined, we found this type of “clean” triple boundary to be the most common in FAPbI₃, with fig. S14 providing further examples as well as an illustration of the wide range of grain boundary angles found within the films.

Although most grain boundaries in these evaporated FAPbI₃ films appear to be highly ordered and sharp, some exceptions also occur. For example, the triple boundary displayed in Fig. 4B contains amorphous material and aligned point defects, as seen at the left and right, respectively. Such variation in the nature of grain boundaries observed here for FAPbI₃ may thus also explain the disparity in the grain boundary properties reported in MAPbI₃, ranging from detrimental (18, 45, 46) to beneficial (16), or even both (11, 17). If a particular preparation method favors a particu-

larly beneficial or detrimental type of grain boundary, it can improve or decrease solar cell performance, respectively.

Finally, we uncovered the presence of aligned point defects and stacking faults in FAPbI₃. The nature and distribution of crystal defects in photoactive perovskites is a topic that has received much attention, but the exact nature of these defects is still unclear (41). Most studies and simulations have focused on point defects and their influence on charge-carrier properties (51). As revealed in Fig. 4D, aligned point defects in the form of vacancies on the Pb-I sublattice are indeed present in the FAPbI₃ crystal. We find here that stacking faults (Fig. 4E) and edge dislocations (Fig. 4F) are also very common in FAPbI₃.

Interestingly, the stacking faults shown in Fig. 4E correspond to a shift of half a unit cell, connecting Pb-I columns with I[−] columns rather than with Pb-I columns. Edge dislocations are also present, as highlighted in the green rectangles in Fig. 4F. The dislocations are dissociated in a direction perpendicular to their glide plane, and are thus referred to as climb-dissociated, with an apparent region of PbI₂ forming the planar fault that joins the dislocations.

Discussion

Resolving the atomic structure of FAPbI₃ thin films allows us to explain many of the intriguing properties of MHPs that had already been well established in the literature but had remained mysterious. The observed loss of FA⁺ through an intermediate perovskite structure displaying a checkered FA⁺ vacancy pattern provides direct clues to the ability of MHPs to sustain deviations in stoichiometry, and to the regenerative capacity of hybrid organic-inorganic MHPs. This mechanism is likely to be general, given that gradual changes have also been observed in photoluminescence and cathodoluminescence spectra when MHPs are exposed to photon (52) and electron (53) irradiation, respectively; such changes do not occur in all-inorganic photoactive perovskites (53). It can also be presumed to underpin the readily regenerative properties of photobleached MAPbI₃ solar cells when they are exposed to gaseous MAI (13); such a mechanism would be consistent with the lead iodide perovskite framework remaining even after substantial loss of the organic A-cation. This highly adaptable and resilient structure can thus accommodate large stoichiometry changes without collapsing, which, along with its potential for regeneration, is a clear advantage for applications. Such resilience of structure, induced by ion migration, may well be unique to MHPs, as compared to other more covalent crystalline solids, and is likely to be the origin of the “radiation hardness” of MHPs, which

makes them useful in space applications and as gamma and x-ray detectors.

In addition, the observed atomically coherent nature of perovskite/PbI₂ interfaces explains how the presence of PbI₂ can enhance the crystal quality of MHP films (16) by acting as seeds for crystal growth while not adversely affecting the local optoelectronic properties of the material. Higher-energy charge carriers generated in PbI₂ can easily transfer to the lower-energy perovskite, thereby contributing to the overall photogenerated current (54), which is relatively unaffected by the perovskite/PbI₂ boundaries. The highly ordered nature of the perovskite/PbI₂ boundary is therefore a core mechanism underpinning the facile fabrication of high-performance solar cells. In essence, this feature forms the MHP analog of the native SiO₂-Si interface, which is benign and is probably the main reason why our electronics industry is primarily based on silicon.

The presence of predominantly sharp, strain-free boundaries between individual FAPbI₃ grains is a vital clue to the reasons for the high performance of perovskite solar cells based on polycrystalline films. Grain boundaries in other polycrystalline semiconductor solar cells are generally detrimental to performance, decreasing the short-circuit current density and introducing trap states (55, 56). The revelation of such sharp boundaries in evaporated FAPbI₃ films, with some amorphous exceptions, therefore presents a first step toward understanding the exact nature of grain boundaries in perovskite solar cells, through which improved control can ensure the highest possible solar cell performance and stability. In addition, our acquisition of truly atomic-resolution images, enabled through careful choices of image acquisition and filtering approaches, will form the basis of investigations into the microstructure of MHPs across a wide range of different stoichiometries. For example, the addition of small fractions of Cs to replace some of the FA cations in FAPbI₃ has been shown to enhance material stability and crystallinity (57), which may link with the nature of grain boundaries.

Finally, stacking faults have been shown to act as recombination sites in dislocation-free silicon (58), and edge dislocations decrease the photoluminescence intensity in multicrystalline silicon (59). Our identification of aligned point defects in the form of vacancies on the FA⁺ sublattice and climb-dissociated dislocations in FAPbI₃ provides first insights into the specific forms taken by such defects in MHPs. Such information will pave the way for future modeling of defects based on first principles, which can reveal the impact such defects have on the optoelectronic properties of MHPs, and the subsequent performance of perovskite solar cells.

Materials and methods

Sample preparation

FAPbI₃ and MAPbI₃ thin films were deposited on the carbon-coated side of 300-mesh copper TEM grids. Immediately before deposition, the grids were cleaned with O₂ plasma for 0.3 min. CH(NH₂)₂I (FAI) and PbI₂ were co-evaporated in a customized thermal evaporation chamber described previously (60). Heating of the sources started when the chamber pressure fell below 5×10^{-6} mbar. The PbI₂ rate was kept constant at 0.23 \AA s^{-1} , measured using a gold-plated quartz microbalance (QMB). For the “PbI₂-rich” FAPbI₃ films with substantial excess of PbI₂, the FAI rate was set to 40% of the rate used for the “near-stoichiometric” FAPbI₃ films. This resulted in FAI crucible temperatures of $\sim 150^\circ\text{C}$ during the deposition of the PbI₂-rich FAPbI₃ films and 180°C during the deposition of the near-stoichiometric FAPbI₃ films. During deposition, the pressure increased to 1×10^{-5} to 2×10^{-5} mbar for the PbI₂-rich FAPbI₃ films, and to 5×10^{-5} to 9×10^{-5} mbar for the near-stoichiometric FAPbI₃ films. The duration of the overall perovskite deposition was controlled such that the film thickness was 90 nm for the PbI₂-rich FAPbI₃ films and 30 nm for the near-stoichiometric FAPbI₃ films. After deposition, all films were annealed at 170°C for 1 min. The films were stored in a nitrogen-filled glove box after preparation and transported to the microscope in a triple-sealed nitrogen atmosphere to prevent exposure to moisture.

CH₃NH₃PbI₃ (MAPbI₃) thin films were prepared using a recently published method (32). CH₃NH₃I (MAI, Greatcell) and PbI₂ (99.999% metal base, Sigma-Aldrich) were co-evaporated in the same custom Lesker thermal evaporator chamber. Heating of the sources started when the chamber pressure fell below 5×10^{-6} mbar. During the subsequent evaporation, the pressure increased to 1×10^{-5} to 2×10^{-5} mbar. The PbI₂ rate was kept constant at 0.30 \AA s^{-1} using a gold-plated QMB; the MAI temperature was controlled to aim for an overall MAPbI₃ deposition rate of 0.45 \AA s^{-1} , as measured by a QMB located next to the substrate. For the depositions used in this work, this involved ramping the MAI crucible from 175.1°C to 175.6°C over 1070 s, for 1.1 g of MAI initially in the crucible. The substrate was cooled to 0°C during the deposition; the sensors were cooled to 17°C .

All samples were loaded rapidly into the microscope, being in contact with air for less than 1 min. Perovskite film thicknesses were determined from calibration of material deposition rates of the two sources according to QMB readings inside the vacuum chamber, supported by film depositions on hard transparent quartz substrates for which film thickness readings were determined from a combination of Dektak profilometer readings and optical transmission measurements.

Image acquisition

Microscopy for imaging and condition development was performed on two microscopes. LAADF and HAADF imaging was performed on a JEOL ARM-200F cold FEG, Cs probe-corrected STEM, using 200 kV acceleration voltage and 20 to 21 mrad convergence angle. An annular dark field detector at 8 cm camera length was used for signal acquisition, with inner and outer collection angles of 33 and 120.77 mrad for LAADF imaging, and 72.80 and 235.75 mrad for HAADF imaging, respectively. A JEOL ARM-300F C_s probe-corrected STEM at 300 kV was used for imaging condition development. Both room-temperature and cryogenic conditions were examined, and the cryogenic conditions did not noticeably reduce the beam sensitivity of the FAPbI₃ thin films, so all imaging was done at room temperature. Similarly, no noticeable difference in beam sensitivity was found between 200 and 300 kV acceleration voltage, but the material was found to damage faster at 80 kV. All imaging presented in the manuscript was at 200 kV. All alignments and focusing were done away from the areas imaged to reduce electron beam-induced damage to the material. All micrographs were obtained without tilting the sample to reduce the beam damage. The specific image acquisition conditions for all images are listed in table S1.

REFERENCES AND NOTES

- H.-S. Kim *et al.*, Lead iodide perovskite sensitized all-solid-state submicron thin film mesoscopic solar cell with efficiency exceeding 9%. *Sci. Rep.* **2**, 591 (2012). doi: [10.1038/srep00591](https://doi.org/10.1038/srep00591); pmid: [22912919](https://pubmed.ncbi.nlm.nih.gov/22912919/)
- M. Liu, M. B. Johnston, H. J. Snaith, Efficient planar heterojunction perovskite solar cells by vapour deposition. *Nature* **501**, 395–398 (2013). doi: [10.1038/nature12509](https://doi.org/10.1038/nature12509); pmid: [24025775](https://pubmed.ncbi.nlm.nih.gov/24025775/)
- H. Zhou *et al.*, Interface engineering of highly efficient perovskite solar cells. *Science* **345**, 542–546 (2014). doi: [10.1126/science.1254050](https://doi.org/10.1126/science.1254050); pmid: [25082698](https://pubmed.ncbi.nlm.nih.gov/25082698/)
- N. J. Jeon *et al.*, Solvent engineering for high-performance inorganic-organic hybrid perovskite solar cells. *Nat. Mater.* **13**, 897–903 (2014). doi: [10.1038/nmat4014](https://doi.org/10.1038/nmat4014); pmid: [24997740](https://pubmed.ncbi.nlm.nih.gov/24997740/)
- J.-P. Correa-Baena *et al.*, The rapid evolution of highly efficient perovskite solar cells. *Energy Environ. Sci.* **10**, 710–727 (2017). doi: [10.1039/C6EE03397K](https://doi.org/10.1039/C6EE03397K)
- M. Saliba *et al.*, Incorporation of rubidium cations into perovskite solar cells improves photovoltaic performance. *Science* **354**, 206–209 (2016). doi: [10.1126/science.aah5557](https://doi.org/10.1126/science.aah5557); pmid: [27708053](https://pubmed.ncbi.nlm.nih.gov/27708053/)
- D. P. McMeekin *et al.*, A mixed-cation lead mixed-halide perovskite absorber for tandem solar cells. *Science* **351**, 151–155 (2016). doi: [10.1126/science.aad5845](https://doi.org/10.1126/science.aad5845); pmid: [26744401](https://pubmed.ncbi.nlm.nih.gov/26744401/)
- D. W. deQuilletes *et al.*, Photo-induced halide redistribution in organic-inorganic perovskite films. *Nat. Commun.* **7**, 11683 (2016). doi: [10.1038/ncomm11683](https://doi.org/10.1038/ncomm11683); pmid: [27216703](https://pubmed.ncbi.nlm.nih.gov/27216703/)
- C. C. Stoumpos, C. D. Malliakas, M. G. Kanatzidis, Semiconducting tin and lead iodide perovskites with organic cations: Phase transitions, high mobilities, and near-infrared photoluminescent properties. *Inorg. Chem.* **52**, 9019–9038 (2013). doi: [10.1021/ic401215x](https://doi.org/10.1021/ic401215x); pmid: [23834108](https://pubmed.ncbi.nlm.nih.gov/23834108/)
- T. Baikie *et al.*, A combined single crystal neutron/X-ray diffraction and solid-state nuclear magnetic resonance study of the hybrid perovskites CH₃NH₃PbX₃ (X = I, Br and Cl). *J. Mater. Chem. A* **3**, 9298–9307 (2015). doi: [10.1039/C5TA01125F](https://doi.org/10.1039/C5TA01125F)
- M. U. Rothmann, W. Li, J. Etheridge, Y. B. Cheng, Microstructural Characterisations of Perovskite Solar Cells—From Grains to Interfaces: Techniques, Features, and Challenges. *Adv. Energy Mater.* **7**, 1700912 (2017). doi: [10.1002/aenm.201700912](https://doi.org/10.1002/aenm.201700912)

12. K. Domanski, E. A. Alharbi, A. Hagfeldt, M. Grätzel, W. Tress, Systematic investigation of the impact of operation conditions on the degradation behaviour of perovskite solar cells. *Nat. Energy* **3**, 61–67 (2018). doi: [10.1038/s41560-017-0060-5](https://doi.org/10.1038/s41560-017-0060-5)
13. L. Hong *et al.*, Improvement and Regeneration of Perovskite Solar Cells via Methylamine Gas Post-Treatment. *Adv. Funct. Mater.* **27**, 1703060 (2017). doi: [10.1002/adfm.201703060](https://doi.org/10.1002/adfm.201703060)
14. D. H. Cao *et al.*, Remnant PbI₂, an unforeseen necessity in high-efficiency hybrid perovskite-based solar cells? *APL Mater.* **2**, 091101 (2014). doi: [10.1063/1.4895038](https://doi.org/10.1063/1.4895038)
15. H.-Y. Hsu *et al.*, Optimization of PbI₂/MAPbI₃ perovskite composites by scanning electrochemical microscopy. *J. Phys. Chem. C* **120**, 19890–19895 (2016). doi: [10.1021/acs.jpcc.6b07850](https://doi.org/10.1021/acs.jpcc.6b07850)
16. Y. C. Kim *et al.*, Beneficial effects of PbI₂ incorporated in organo-lead halide perovskite solar cells. *Adv. Energy Mater.* **6**, 1502104 (2016). doi: [10.1002/aenm.201502104](https://doi.org/10.1002/aenm.201502104)
17. J.-W. Lee *et al.*, The role of grain boundaries in perovskite solar cells. *Mater. Today Energy* **7**, 149–160 (2018). doi: [10.1016/j.mtener.2017.07.014](https://doi.org/10.1016/j.mtener.2017.07.014)
18. R. Long, J. Liu, O. V. Prezhdo, Unravelling the effects of grain boundary and chemical doping on electron-hole recombination in CH₃NH₃PbI₃ perovskite by time-domain atomistic simulation. *J. Am. Chem. Soc.* **138**, 3884–3890 (2016). doi: [10.1021/jacs.6b00645](https://doi.org/10.1021/jacs.6b00645); pmid: 26930494
19. T. Niu *et al.*, Stable High-Performance Perovskite Solar Cells via Grain Boundary Passivation. *Adv. Mater.* **30**, e1706576 (2018). doi: [10.1002/adma.201706576](https://doi.org/10.1002/adma.201706576); pmid: 29527750
20. S. Jariwala *et al.*, Local crystal misorientation influences non-radiative recombination in halide perovskites. *Joule* **3**, 3048–3060 (2019). doi: [10.1016/j.joule.2019.09.001](https://doi.org/10.1016/j.joule.2019.09.001)
21. C. L. Davies *et al.*, Impact of the organic cation on the optoelectronic properties of formamidinium lead triiodide. *J. Phys. Chem. Lett.* **9**, 4502–4511 (2018). doi: [10.1021/acs.jpclett.8b01628](https://doi.org/10.1021/acs.jpclett.8b01628); pmid: 30036475
22. S. K. Yadavalli *et al.*, Mechanisms of Exceptional Grain Growth and Stability in Formamidinium Lead Triiodide Thin Films for Perovskite Solar Cells. *Acta Mater.* **193**, 10–18 (2020). doi: [10.1016/j.actamat.2020.03.036](https://doi.org/10.1016/j.actamat.2020.03.036)
23. A. Knight, L. M. Herz, Preventing phase segregation in mixed-halide perovskites: A perspective. *Energy Environ. Sci.* **13**, 2024–2046 (2020). doi: [10.1039/D0EE00788A](https://doi.org/10.1039/D0EE00788A)
24. O. Breitenstein, J. Rakotonianina, M. H. Al Rifai, M. Werner, Shunt types in crystalline silicon solar cells. *Prog. Photovolt. Res. Appl.* **12**, 529–538 (2004). doi: [10.1002/ppp.544](https://doi.org/10.1002/ppp.544)
25. A. Stoffers *et al.*, Complex Nanotwin Substructure of an Asymmetric Σ 9 Tilt Grain Boundary in a Silicon Polycrystal. *Phys. Rev. Lett.* **115**, 235502 (2015). doi: [10.1103/PhysRevLett.115.235502](https://doi.org/10.1103/PhysRevLett.115.235502); pmid: 26684123
26. Y. Yu *et al.*, Atomic resolution imaging of halide perovskites. *Nano Lett.* **16**, 7530–7535 (2016). doi: [10.1021/acs.nanolett.6b03331](https://doi.org/10.1021/acs.nanolett.6b03331); pmid: 27960472
27. D. Zhang *et al.*, Atomic-resolution transmission electron microscopy of electron beam-sensitive crystalline materials. *Science* **359**, 675–679 (2018). doi: [10.1126/science.aao0865](https://doi.org/10.1126/science.aao0865); pmid: 29348363
28. Y. Li *et al.*, Unravelling degradation mechanisms and atomic structure of organic-inorganic halide perovskites by cryo-EM. *Joule* **3**, 2854–2866 (2019). doi: [10.1016/j.joule.2019.08.016](https://doi.org/10.1016/j.joule.2019.08.016)
29. M. U. Rothmann *et al.*, Structural and chemical changes to CH₃NH₃PbI₃ induced by electron and gallium ion beams. *Adv. Mater.* **30**, e1800629 (2018). doi: [10.1002/adma.201800629](https://doi.org/10.1002/adma.201800629); pmid: 29700861
30. S. Chen *et al.*, Atomic scale insights into structure instability and decomposition pathway of methylammonium lead iodide perovskite. *Nat. Commun.* **9**, 4807 (2018). doi: [10.1038/s41467-018-07177-y](https://doi.org/10.1038/s41467-018-07177-y); pmid: 30442950
31. R. Hooley, A. Brown, A. Kulak, F. Meldrum, R. Brydson, *J. Phys. Conf. Ser.* **902**, 012005 (2017). doi: [10.1021/acsenergylett.0c00183](https://doi.org/10.1021/acsenergylett.0c00183); pmid: 32296733
32. K. B. Lohmann *et al.*, Control over crystal size in vapor deposited metal-halide perovskite films. *ACS Energy Lett.* **5**, 710–717 (2020). doi: [10.1021/acscenergylett.0c00183](https://doi.org/10.1021/acscenergylett.0c00183); pmid: 32296733
33. J. Borchert *et al.*, Large-area, highly uniform evaporated formamidinium lead triiodide thin films for solar cells. *ACS Energy Lett.* **2**, 2799–2804 (2017). doi: [10.1021/acscenergylett.7b00967](https://doi.org/10.1021/acscenergylett.7b00967)
34. J. Gonnissen *et al.*, Optimal experimental design for the detection of light atoms from high-resolution scanning transmission electron microscopy images. *Appl. Phys. Lett.* **105**, 063116 (2014). doi: [10.1063/1.4892884](https://doi.org/10.1063/1.4892884)
35. L. Oesinghaus *et al.*, Toward tailored film morphologies: The origin of crystal orientation in hybrid perovskite thin films. *Adv. Mater. Interfaces* **3**, 1600403 (2016). doi: [10.1002/admi.201600403](https://doi.org/10.1002/admi.201600403)
36. H. J. Jung *et al.*, Stability of halide perovskite solar cell devices: In situ observation of oxygen diffusion under biasing. *Adv. Mater.* **30**, e1802769 (2018). doi: [10.1002/adma.201802769](https://doi.org/10.1002/adma.201802769); pmid: 30133013
37. T. W. Kim *et al.*, Self-Organized Superlattice and Phase Coexistence inside Thin Film Organometal Halide Perovskite. *Adv. Mater.* **30**, 1705230 (2018). doi: [10.1002/adma.201705230](https://doi.org/10.1002/adma.201705230); pmid: 29318666
38. Z. Chen, Z.-G. Gu, W.-Q. Fu, F. Wang, J. Zhang, A confined fabrication of perovskite quantum dots in oriented MOF thin film. *ACS Appl. Mater. Interfaces* **8**, 28737–28742 (2016). doi: [10.1021/acsami.6b11712](https://doi.org/10.1021/acsami.6b11712); pmid: 27723302
39. M. Xiao *et al.*, A fast deposition-crystallization procedure for highly efficient lead iodide perovskite thin-film solar cells. *Angew. Chem.* **126**, 10056–10061 (2014). doi: [10.1002/ange.201405334](https://doi.org/10.1002/ange.201405334); pmid: 25047967
40. Y. Zhou, H. Sternlicht, N. P. Padture, Transmission electron microscopy of halide perovskite materials and devices. *Joule* **3**, 641–661 (2019). doi: [10.1016/j.joule.2018.12.011](https://doi.org/10.1016/j.joule.2018.12.011)
41. J. Ran *et al.*, Electron-Beam-Related Studies of Halide Perovskites: Challenges and Opportunities. *Adv. Energy Mater.* **10**, 1903191 (2020). doi: [10.1002/aenm.201903191](https://doi.org/10.1002/aenm.201903191)
42. R. F. Egerton, Radiation damage to organic and inorganic specimens in the TEM. *Micron* **119**, 72–87 (2019). doi: [10.1016/j.micron.2019.01.005](https://doi.org/10.1016/j.micron.2019.01.005); pmid: 30684768
43. R. F. Egerton, P. Li, M. Malac, Radiation damage in the TEM and SEM. *Micron* **35**, 399–409 (2004). doi: [10.1016/j.micron.2004.02.003](https://doi.org/10.1016/j.micron.2004.02.003); pmid: 15120123
44. P. A. Beckmann, A review of polytypism in lead iodide. *Cryst. Res. Technol.* **45**, 455–460 (2010). doi: [10.1002/crat.201000066](https://doi.org/10.1002/crat.201000066)
45. H.-S. Duan *et al.*, The identification and characterization of defect states in hybrid organic-inorganic perovskite photovoltaics. *Phys. Chem. Chem. Phys.* **17**, 112–116 (2015). doi: [10.1039/C4CP04479G](https://doi.org/10.1039/C4CP04479G); pmid: 25354141
46. X. Wen *et al.*, Defect trapping states and charge carrier recombination in organic-inorganic halide perovskites. *J. Mater. Chem. C* **4**, 793–800 (2016). doi: [10.1039/C5TC03109E](https://doi.org/10.1039/C5TC03109E)
47. D. W. deQuilettes *et al.*, Impact of microstructure on local carrier lifetime in perovskite solar cells. *Science* **348**, 683–686 (2015). doi: [10.1126/science.aaa5333](https://doi.org/10.1126/science.aaa5333); pmid: 25931446
48. Q. Chen *et al.*, Controllable self-induced passivation of hybrid lead iodide perovskites toward high performance solar cells. *Nano Lett.* **14**, 4158–4163 (2014). doi: [10.1021/nl501838y](https://doi.org/10.1021/nl501838y); pmid: 24960309
49. Y. Shao, Z. Xiao, C. Bi, Y. Yuan, J. Huang, Origin and elimination of photocurrent hysteresis by fullerene passivation in CH₃NH₃PbI₃ planar heterojunction solar cells. *Nat. Commun.* **5**, 5784 (2014). doi: [10.1038/ncomms5784](https://doi.org/10.1038/ncomms5784); pmid: 25503258
50. A. F. Castro-Méndez, J. Hidalgo, J. P. Correa-Baena, The role of grain boundaries in perovskite solar cells. *Adv. Energy Mater.* **9**, 1901489 (2019). doi: [10.1002/aenm.201901489](https://doi.org/10.1002/aenm.201901489)
51. J. M. Ball, A. Petrozza, Defects in perovskite-halides and their effects in solar cells. *Nat. Energy* **1**, 16149 (2016). doi: [10.1038/nenergy.2016.149](https://doi.org/10.1038/nenergy.2016.149)
52. A. Merdasa *et al.*, Super-resolution luminescence microspectroscopy reveals the mechanism of photoinduced degradation in CH₃NH₃PbI₃ perovskite nanocrystals. *J. Phys. Chem. C* **120**, 10711–10719 (2016). doi: [10.1021/acs.jpcc.6b03512](https://doi.org/10.1021/acs.jpcc.6b03512)
53. C. Xiao *et al.*, Mechanisms of Electron-Beam-Induced Damage in Perovskite Thin Films Revealed by Cathodoluminescence Spectroscopy. *J. Phys. Chem. C* **119**, 26904–26911 (2015). doi: [10.1021/acs.jpcc.5b09698](https://doi.org/10.1021/acs.jpcc.5b09698)
54. J. B. Patel *et al.*, Photocurrent spectroscopy of perovskite solar cells over a wide temperature range from 15 to 350 K. *J. Phys. Chem. Lett.* **9**, 263–268 (2018). doi: [10.1021/acs.jpcclett.7b02935](https://doi.org/10.1021/acs.jpcclett.7b02935); pmid: 29260569
55. J. D. Zook, Effects of grain boundaries in polycrystalline solar cells. *Appl. Phys. Lett.* **37**, 223–226 (1980). doi: [10.1063/1.91832](https://doi.org/10.1063/1.91832)
56. F. Greuter, G. Blatter, Electrical properties of grain boundaries in polycrystalline compound semiconductors. *Semicond. Sci. Technol.* **5**, 111–137 (1990). doi: [10.1088/0268-1242/5/2/001](https://doi.org/10.1088/0268-1242/5/2/001)
57. J.-W. Lee *et al.*, Formamidinium and Cesium Hybridization for Photo- and Moisture-Stable Perovskite Solar Cell. *Adv. Energy Mater.* **5**, 1501310 (2015). doi: [10.1002/aenm.201501310](https://doi.org/10.1002/aenm.201501310)
58. L. Kimerling, H. Leamy, J. Patel, The electrical properties of stacking faults and precipitates in heat-treated dislocation-free Czochralski silicon. *Appl. Phys. Lett.* **30**, 217–219 (1977). doi: [10.1063/1.89355](https://doi.org/10.1063/1.89355)
59. H. Sugimoto, M. Inoue, M. Tajima, A. Ogura, Y. Ohshita, Analysis of intra-grain defects in multicrystalline silicon wafers by photoluminescence mapping and spectroscopy. *Jpn. J. Appl. Phys.* **45**, L641–L643 (2006). doi: [10.1143/JJAP.45.L641](https://doi.org/10.1143/JJAP.45.L641)
60. J. Borchert *et al.*, Impurity Tracking Enables Enhanced Control and Reproducibility of Hybrid Perovskite Vapor Deposition. *ACS Appl. Mater. Interfaces* **11**, 28851–28857 (2019). doi: [10.1021/acsami.9b07619](https://doi.org/10.1021/acsami.9b07619); pmid: 31314481

ACKNOWLEDGMENTS

We thank the David Cockayne Centre for Electron Microscopy, University of Oxford, for access and support in the use of the JEOL ARM200F instrument (proposal number EP/K040375/1) and additional instrument provision from the Henry Royce Institute (grant reference EP/R010145/1). We also thank Diamond Light Source for access and support in use of the electron Physical Science Imaging Centre (E02, MG21734) that contributed to the results presented here. L.M.H. and M.B.J. thank the Humboldt Foundation for research awards. **Funding:** Supported by the UK Engineering and Physical Sciences Research Council (EPSRC) through grant EP/P033229/1 and through the EPSRC CDT for New and Sustainable Photovoltaics. **Author contributions:** M.U.R. designed and carried out the electron microscope studies and analyzed and interpreted data. J.S.K. participated in the EM studies and provided insights in data interpretation and during editing. J.B. and K.B.L. prepared the samples. C.M.O. and A.A.S. assisted with the EM studies. L.C. performed the multislice simulations. H.J.S. contributed to data analysis and interpretation. M.B.J., P.D.N., and L.M.H. designed the study, interpreted data, and supervised M.U.R., who wrote the manuscript with input from all co-authors. **Competing interests:** H.J.S. is the chief scientific officer and a director of Oxford PV Ltd., a company that is commercializing perovskite photovoltaic technologies. **Data and materials availability:** The datasets and numerical analysis sets described in this manuscript are available for download from the Oxford University Research Archive, under the following DOI: 10.5287/bodleian:pr10G00gZ.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/eabb5940/suppl/DC1
Figs. S1 to S16
Table S1
References (61–68)

5 March 2020; accepted 3 September 2020
10.1126/science.abb5940

RESEARCH ARTICLE SUMMARY

MICROBIOTA

Multi-omics analyses of radiation survivors identify radioprotective microbes and metabolites

Hao Guo, Wei-Chun Chou, Yunjia Lai, Kaixin Liang, Jason W. Tam, W. June Brickey, Liang Chen, Nathan D. Montgomery, Xin Li, Lauren M. Bohannon, Anthony D. Sung, Nelson J. Chao, Jonathan U. Peled, Antonio L. C. Gomes, Marcel R. M. van den Brink, Matthew J. French, Andrew N. Macintyre, Gregory D. Sempowski, Xianming Tan, R. Balfour Sartor, Kun Lu, Jenny P. Y. Ting*

INTRODUCTION: The toxicity of high-dose ionizing radiation is associated with the induction of both chronic and acute radiation syndromes that occur after partial or total body radiation and can be further characterized into hematopoietic, gastrointestinal, and cerebrovascular syndromes. The intestine is the major target of radiation and the biggest niche for gut microbiota. Although there are sporadic descriptive studies showing a potential correlation between the gut microbiota and radiation-induced damage, the detailed underpinnings of this relationship remain obscure. In addition, medical intervention to counteract radiation injury is still a global challenge despite decades of rigorous research.

RATIONALE: Over the last decade, numerous investigations have demonstrated highly di-

verse gut microbiota between individuals and significant correlations of gut microbiota with multiple diseases. Gut microbes, as well as microbe-derived metabolites represented by short-chain fatty acids (SCFAs) and tryptophan metabolites, have essential roles in regulating host metabolism and immunity. The imbalance or dysbiosis of a microbial community is associated with potential diseases, risks, or even to the clear onset of clinical symptoms. We have previously corroborated the biological importance of gut microbiota and certain bacteria (e.g., *Lachnospiraceae*) together with SCFAs in attenuating colitis and obesity. It has also been reported that SCFAs and tryptophan metabolites can reduce proinflammatory cytokines such as tumor necrosis factor- α , interleukin-6, and interferon- γ and promote the anti-inflammatory cytokines, all

of which are vital mediators of radiation-induced damage. These findings raise the possibility that the gut microbiota and metabolites play a key role in the regulation of disease susceptibility after radiation challenge.

RESULTS: We found that a small percentage of mice could survive a high dose of radiation and live a normal life span. These “elite-survivors” harbored a distinct gut microbiome that developed after radiation. Taking advantage of this finding, we used a combination of fecal engraftment and dirty cage sharing to demonstrate that the microbiota from elite-survivors provided substantial radioprotection in both germ-free and conventionally housed recipients, characterized by enhanced survival and ameliorated clinical scores. An unbiased microbiome analysis identified *Lachnospiraceae* and *Enterococcaceae* as the most enriched bacteria in elite-survivors. Monoassociation analysis provided direct evidence for the protective role of *Lachnospiraceae* and *Enterococcaceae* in promoting hematopoiesis and attenuating gastrointestinal damage. Clinical relevance in humans was supported by an analysis of leukemia patients who were exposed to whole-body radiation. The elevated abundance of *Lachnospiraceae* and *Enterococcaceae* was associated with fewer adverse effects in a highly statistically significant fashion. Treatment with SCFAs, especially propionate, rendered mice resistant to radiation, mediated by attenuation of DNA damage and reactive oxygen species release both in hematopoietic and gastrointestinal tissues. Further, an untargeted metabolomics study revealed a realm of metabolites that were affected by radiation and selectively increased in elite-survivors. Among these, two tryptophan pathway metabolites, ¹H-indole-3-carboxaldehyde (I3A) and kynurenic acid (KYNA), provided long-term radioprotection in vivo.

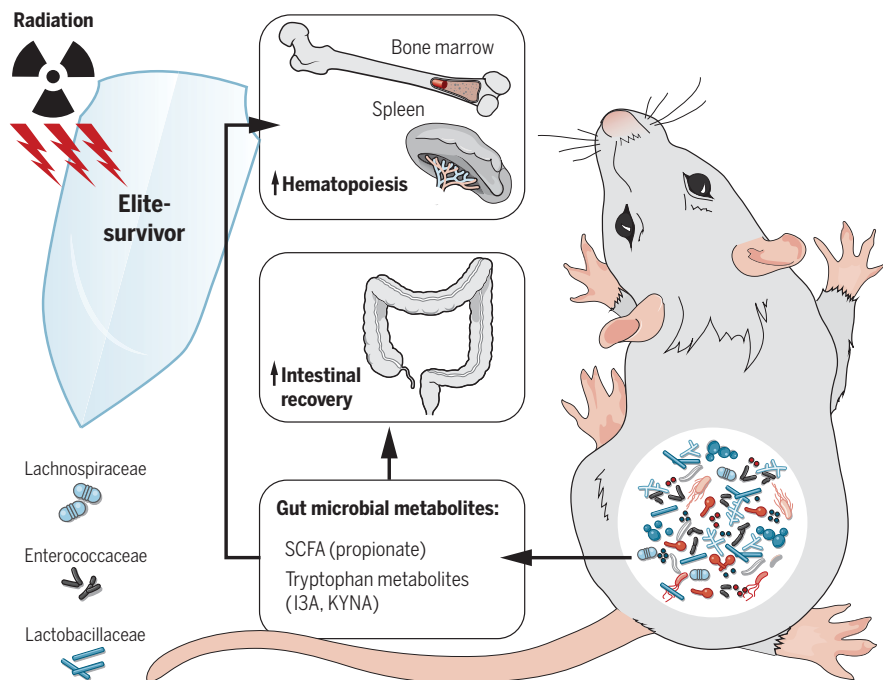
CONCLUSION: Our findings emphasize a crucial role for the gut microbiota as a master regulator of host defense against radiation, capable of protecting both the hematopoietic and gastrointestinal systems. *Lachnospiraceae* and *Enterococcaceae*, together with downstream metabolites represented by propionate and tryptophan pathway members, contribute substantially to radioprotection. This study sheds light on the pivotal role that the microbiota-metabolite axis plays in generating broad protection against radiation and provides promising therapeutic targets to treat the adverse side effects of radiation exposure. ■

The list of author affiliations is available in the full article online.

*Corresponding author. Email: jenny_ting@med.unc.edu

Cite this article as: H. Guo et al., *Science* 370, eaay9097 (2020). DOI: 10.1126/science.aay9097

S READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.aay9097>



Gut microbiota and metabolites mediate radioprotection. Gut microbes, especially *Lachnospiraceae* and *Enterococcaceae* along with bacteria-derived metabolites represented by SCFA (propionate) and tryptophan pathway members (I3A and KYNA), tune host resistance against high doses of radiation by facilitating hematopoiesis and gastrointestinal recovery.

RESEARCH ARTICLE

MICROBIOTA

Multi-omics analyses of radiation survivors identify radioprotective microbes and metabolites

Hao Guo¹, Wei-Chun Chou^{1,2}, Yunjia Lai³, Kaixin Liang^{1,4}, Jason W. Tam^{1*}, W. June Brickey^{1,5}, Liang Chen^{1,2,5*}, Nathan D. Montgomery^{1,6}, Xin Li¹, Lauren M. Bohannon⁷, Anthony D. Sung⁷, Nelson J. Chao⁷, Jonathan U. Peled^{8,9}, Antonio L. C. Gomes^{8,9}, Marcel R. M. van den Brink^{8,9}, Matthew J. French¹⁰, Andrew N. Macintyre¹⁰, Gregory D. Sempowski¹⁰, Xianming Tan¹, R. Balfour Sartor¹¹, Kun Lu³, Jenny P. Y. Ting^{1,2,5†}

Ionizing radiation causes acute radiation syndrome, which leads to hematopoietic, gastrointestinal, and cerebrovascular injuries. We investigated a population of mice that recovered from high-dose radiation to live normal life spans. These “elite-survivors” harbored distinct gut microbiota that developed after radiation and protected against radiation-induced damage and death in both germ-free and conventionally housed recipients. Elevated abundances of members of the bacterial taxa *Lachnospiraceae* and *Enterococcaceae* were associated with postradiation restoration of hematopoiesis and gastrointestinal repair. These bacteria were also found to be more abundant in leukemia patients undergoing radiotherapy, who also displayed milder gastrointestinal dysfunction. In our study in mice, metabolomics revealed increased fecal concentrations of microbially derived propionate and tryptophan metabolites in elite-survivors. The administration of these metabolites caused long-term radioprotection, mitigation of hematopoietic and gastrointestinal syndromes, and a reduction in proinflammatory responses.

Radiation-induced multiorgan injury represents a sizeable unmet medical need that occurs during cancer radiotherapy and accidental exposure. Substantial federal efforts have been made to develop therapeutic radioprotectants for routine clinical use, but radiation injury remains a long-standing and unresolved problem (1). Acute radiation syndrome or sickness (ARS) occurs after partial or total body exposure to ionizing radiation, and death is frequently attributed to poor hematopoietic recovery and/or death of epithelial cells lining the gastrointestinal (GI) tract. The GI tract is the fastest-renewing adult tissue, making it highly sensitive to radiation. It also harbors a diverse microbial

community comprising 10 to 100 trillion microorganisms, raising the possibility that the gut microbiota is involved in responses to radiation-induced damage (2–7). Although there have been some studies showing associations between gut microbiota and radiation-induced damage (6, 8, 9), possible mechanisms are poorly understood.

Our work stemmed from an observation that a small percentage of specific pathogen-free (SPF) C57BL/6 mice survived a high dose of total body radiation and lived a normal life span up to 600 days. These mice have the same genetic background as those that did not survive radiation, thus prompting us to look for differences in the microbiota, which have been implicated in radioprotection (4–9). To better understand the interactions between the gut microbiome and radiation-induced damage, we performed microbiota adoptive transfer experiments as well as monoassociation studies to investigate the function of the gut bacterial community and specific bacterial strains in a radiation model. Both biased and unbiased metabolomics profiling were applied to uncover potential metabolites that play a vital role in radioprotection.

Distinct microbiome in elite-survivors

On administration of a high dose of total body radiation (9.2 Gy), we unexpectedly observed that 5 to 15% of exposed SPF C57BL/6 mice recovered and survived long term (up to 600 days) (Fig. 1A). These mice are referred to as “elite-survivors.” To determine whether the gut mi-

crobiota was specific to the elite-survivors, we performed high-throughput gene-sequencing analysis of 16S ribosomal RNA (rRNA) in fecal bacterial DNA isolated from age-matched control mice and elite-survivors at day 290 after radiation. Rarefaction analysis comparing bacterial diversity within individual subjects revealed that elite-survivors harbored a distinct bacterial community relative to that of controls (Fig. 1B). These data were further quantified by UniFrac dissimilarity distance analysis (Fig. 1C) and supported by heatmap analysis of bacterial operational taxonomic units (Fig. 1D; details of the bacterial taxa are listed in table S1). These results led to the speculation that specific gut microbiota may protect against radiation.

Radioprotection mediated by microbiota

To investigate more stringently the contribution of microbiota to radioprotection, we performed a “cohousing” assay. Fecal microbiota samples were passively transferred from elite-survivors through a dirty cage-sharing experiment (Fig. 2A). This experimental design was necessary because elite-survivors were old male mice, so a traditional cohousing approach with young male recipients might lead to fighting and injury to the older mice. Instead, dirty cages, which contained fecal material and used bedding from elite-survivors or age-matched control mice, were used to house SPF C57BL/6 mice. Weekly for 8 weeks, recipients were transferred into fresh dirty cages and then exposed to a high dose of total body radiation (8.2 to 9.2 Gy). Nearly 75% of mice housed in dirty cages previously occupied by elite-survivors (ES-Recip) survived high radiation, whereas only 20% of those housed in dirty cages from controls (Ctrl-Recip) survived (Fig. 2B). Clinical score (encompassing the seven body parameters listed in table S2, which were proportional to disease severity) was significantly lower in ES-Recip (Fig. 2C) mice. Histological studies were performed in a double-blinded fashion by a licensed pathologist. Reduced stromal injury, decreased hematopoietic cell death marked by caspase 3, and enhanced proliferation indicated by Ki67 were observed in bone marrow specimens from ES-Recip compared with Ctrl-Recip mice (fig. S1, A to C, upper panels). Splenic architecture was near normal in ES-Recip mice, with the white pulp containing well-developed, lymphocyte-rich follicles and the red pulp exhibiting extramedullary hematopoiesis. Conversely, appreciable atrophy and lymphocyte depletion were observed in Ctrl-Recip spleen samples (fig. S1A, lower panels). A reduction in cleaved caspase 3 staining and immunoblotting and increased Ki67 staining were also observed in spleens from ES-Recip mice (fig. S1, B and C, lower panels, and fig. S1D). These results indicate that the microbiota

¹Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ²Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ³Department of Environmental Sciences and Engineering, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ⁴School of Dentistry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ⁵Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ⁶Department of Pathology and Laboratory Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ⁷Division of Hematologic Malignancies and Cellular Therapy/BMT, Department of Medicine, Duke University, Durham, NC, USA. ⁸Adult Bone Marrow Transplantation Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. ⁹Weill Cornell Medical College, New York, NY, USA. ¹⁰Duke Human Vaccine Institute, Duke University, Durham, NC, USA. ¹¹Center for Gastrointestinal Biology and Disease, Department of Medicine, Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

*Present address: AbbVie Inc., Boston, MA, USA.

†Corresponding author. Email: jenny.ting@med.unc.edu

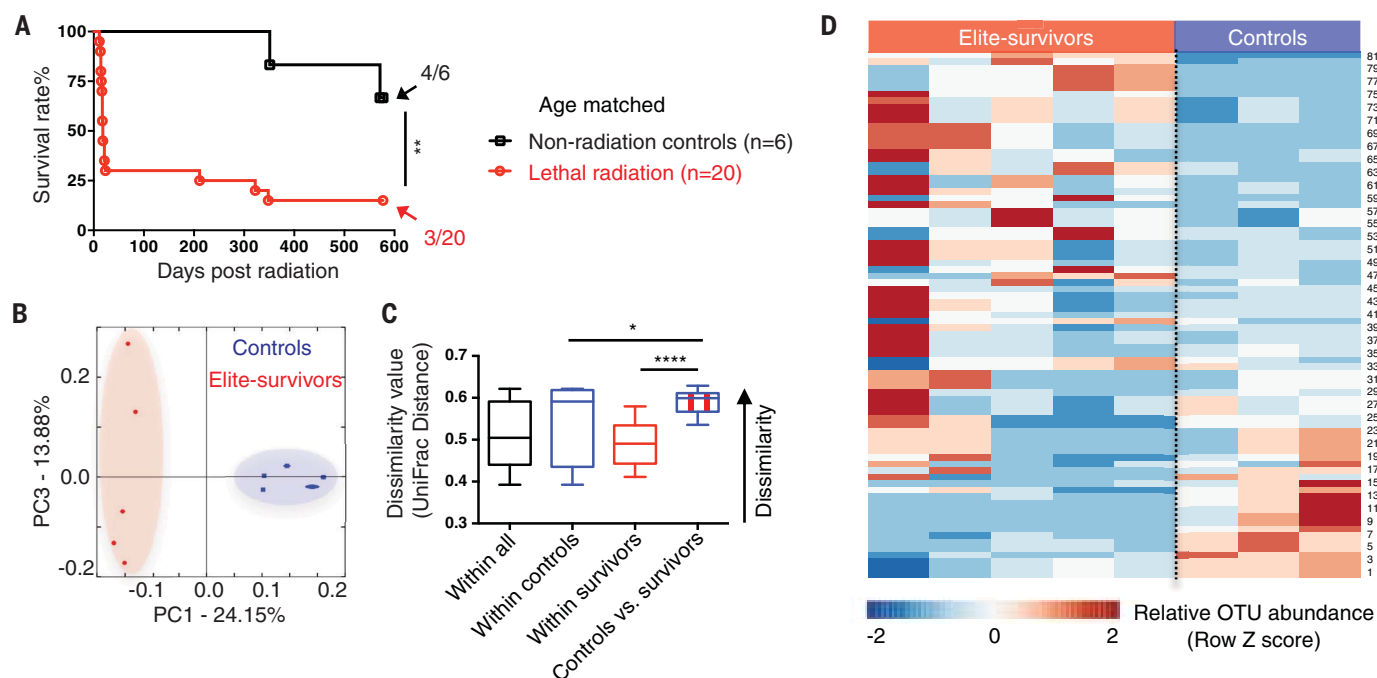


Fig. 1. Radiation elite-survivors harbor a distinct gut microbiota compared with age-matched controls. (A) SPF C57BL/6 mice received a high dose of total body radiation and were monitored for >600 days. Fecal samples were collected from age-matched controls and elite-survivors at day 290 after radiation. Fractions indicate the number of mice that survived to the end of the experiment. Data are pooled from two independent experiments. (B and C) PCA plot

showing microbial compositional differences (B) as quantified by UniFrac distance (C). (D) Heatmap of sequenced bacterial operational taxonomic unit (OTU) abundances. A detailed bacterial taxa list is provided in table S1. Fecal samples were collected from two independent experiments. Each symbol represents one mouse. Error bars indicate SEM. *P < 0.05, **P < 0.01, ****P < 0.0001 determined by log-rank (Mantel Cox) test (A) and Student's t test (C).

from elite-survivors mediates the protection of hosts against a high dose of radiation by promoting hematopoietic proliferation and reducing cell death.

We profiled bacterial 16S rRNA genes in the feces from both donors and recipients after dirty cage sharing. Principal component analysis (PCA) and UniFrac dissimilarity distance analysis showed differences in microbiome composition in ES-Recip mice (Fig. 2, D and E) compared with Ctrl-Recip mice. Furthermore, the microbiome compositions of the recipients were similar to those of their respective donors, indicating that the dirty cage-sharing strategy was efficient in exchanging gut microbiota from donors to recipients (Fig. 2D). Specifically, a significant decrease in the abundance of members of the order *Erysipelotrichales* was accompanied by an increase in the abundances of *Bacteroidales* and *Clostridiales* in elite-survivors compared with controls. This pattern was recapitulated in ES-Recip mice compared with Ctrl-Recip mice (Fig. 2, F and G).

To confirm the relationship between microbiota and radioprotection, we performed a fecal microbiota transplantation (FMT) experiment in which germ-free (GF) or SPF C57BL/6 mice were reconstituted with fecal materials from elite-survivors or age-matched controls (Fig. 2H and fig. S2A) (10). FMT from elite-

survivors resulted in elevated survival rates among GF and SPF mice, which also showed reduced clinical scores compared with that from control mice (Fig. 2, I and J, and fig. S2, B to D), and in an altered microbiome composition as shown by PCA plot and UniFrac dissimilarity distance (Fig. 2, K and L). Bacterial abundances were directly compared between GF ES-Recip and Ctrl-Recip mice. Linear discriminant analysis (LDA) effect size analysis showed that eight taxa were enriched and five were reduced in ES-Recip mice, with an LDA score >0.2 (Fig. 2M). A volcano plot flagged *Lachnospiraceae* as the most enriched and represented bacteria in ES-Recip mice (Fig. 2N).

We undertook a clinical study to discover whether the bacterial taxa that we identified as radioprotective in our mouse studies might be of relevance to humans. We collected fecal samples from 21 leukemia patients at the start of total body radiation given as a prehematopoietic stem cell transplantation conditioning (detailed clinical patient information is listed in table S3). The incidence and severity of GI side effects were correlated with the extent of diarrhea observed. 16S rRNA sequencing was used to identify the gut microbiome composition. We analyzed the abundance of *Lachnospiraceae* and *Enterococcaceae* in all patients (Fig. 3, A and B) during a 30-day ob-

servation period. Patients with shorter diarrheal duration (<10 days) had significantly higher abundances of *Lachnospiraceae* and *Enterococcaceae* than patients with longer diarrheal duration (>10 days). *Lactobacillaceae* also showed an increased trend in patients with shorter diarrheal duration, but the value did not reach statistical significance. There was also a statistically significant negative correlation between the abundance of *Lachnospiraceae* and the occurrence of the GI adverse effects in patients (Fig. 3B). The patient data should be interpreted with caution because these are limited by the small sample size and only show an association of specific bacteria with reduced adverse symptoms; moreover, the study was not designed to show causation.

Postradiation selection of microbiota in mice

Unlike humans, who have diverse health indicators, the GF and SPF mice used in our study should have similar basal microbiome compositions. To address the question of how elite-survivors harbor a radioprotective microbiota, we investigated whether a distinct microbiota predominated in elite-survivors before radiation or if it was selected after radiation (Fig. 3C). A total of 57 naïve SPF C57BL/6 mice were given a high dose of total body radiation (8.2 Gy), and fecal samples were collected from each mouse before (day -1) and at days 2, 7, 21,

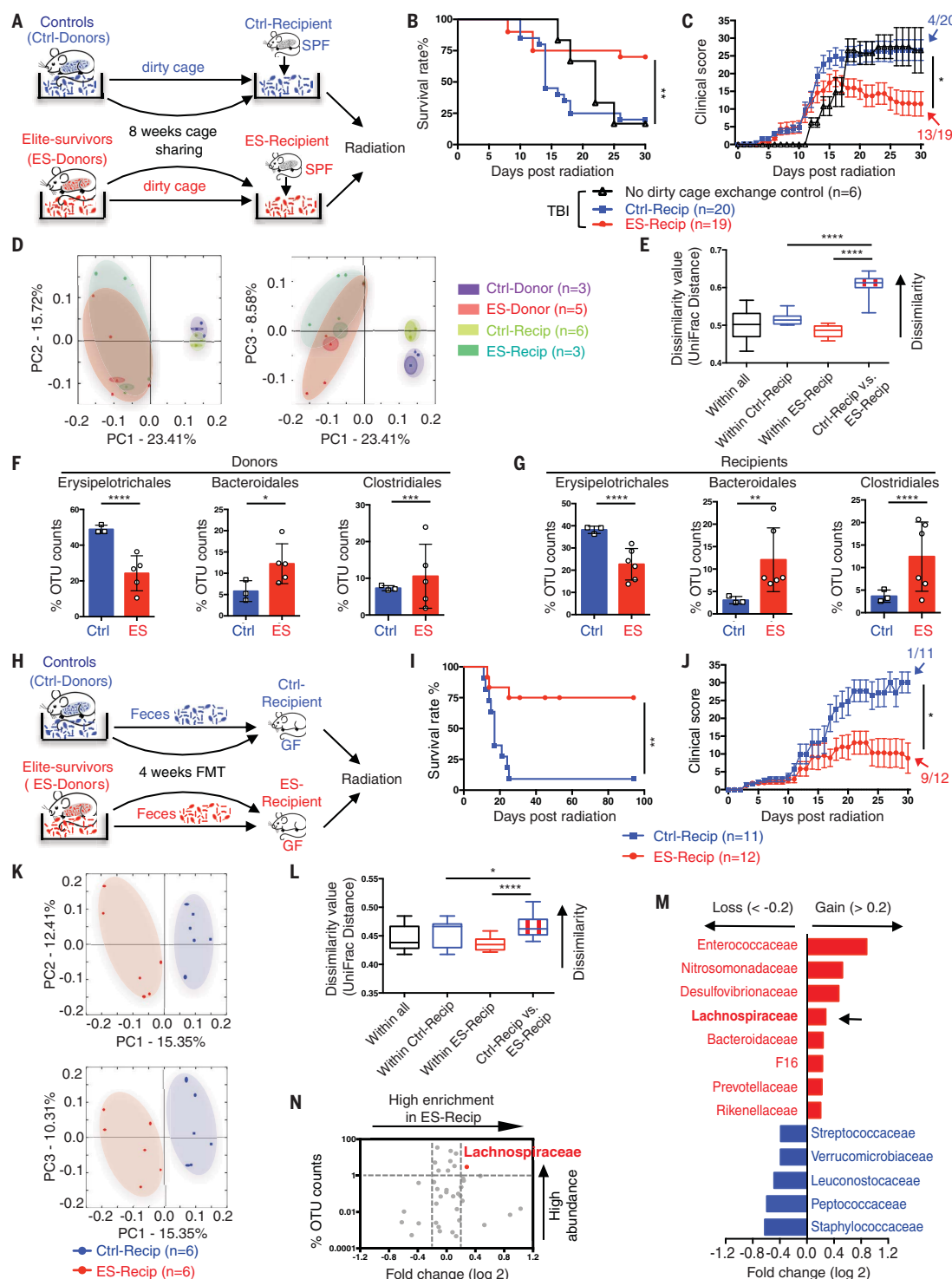


Fig. 2. Gut microbiota from elite-survivors protect GF and SPF recipients from radiation-induced death. (A) Illustration of dirty cage-sharing experiment. (B and C) Survival rates (B) and clinical scores (C). Significance was found between the Ctrl-Recip and ES-Recip groups. Fecal samples were collected from recipients after 8 weeks of dirty cage sharing. (D and E) PCA plot (D) and UniFrac distance (E) of microbial composition. (F and G) Composite results of substantially changed bacterial groups were identified from donors (F) and recipients (G). (H) Illustration of FMT experiment. (I and J) Survival rates (I) and clinical scores (J). Fecal samples were collected after 4 weeks of FMT from recipients. (K and L) PCA plot (K) and UniFrac distance (L). (M) LDA effect size analysis [LDA significant

threshold (log 2) $\geq \pm 0.2$] identified taxonomic biomarkers within recipients. Red bars, enriched within ES-Recip; blue bars, enriched within Ctrl-Recip. (N) Volcano plot displaying relative abundance distribution of microbial OTUs. x axis, log2 relative abundance; y axis, microbial OTU%. Each symbol represents one mouse or bacterial taxa (N). Fractions indicate the number of mice that survived to the end of the experiment [(C) and (J)]. Data are pooled from three [(B) and (C)] or four [(I) and (J)] independent experiments. Error bars indicate SEM, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$ determined by log-rank (Mantel Cox) test [(B) and (I)], Mann-Whitney test for area under the curve (AUC) [(C) and (J)], Student's t test [(E) and (L)], and two-way ANOVA [(F) and (G)].

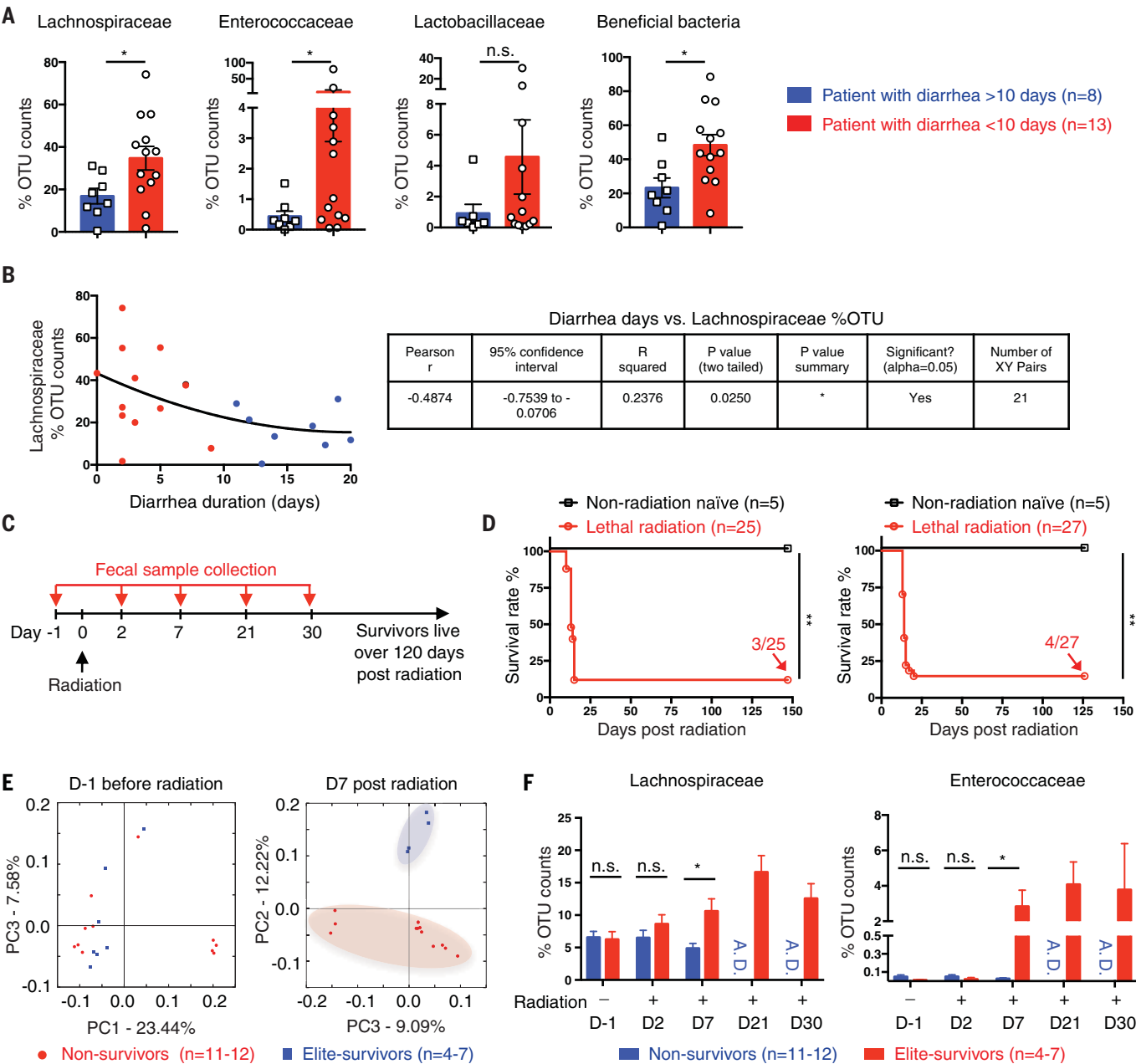


Fig. 3. Gut microbiota is significantly associated with radioprotection both in patients and mice. (A and B) Fecal samples were collected from 21 leukemia patients at the start of total body radiation given as a prehematopoietic stem cell transplantation conditioning. The gut microbiome was detected by 16S rRNA sequencing and compared in patients stratified by duration of diarrhea experienced. (A) Relative abundance of *Lachnospiraceae*, *Enterococcaceae*, *Lactobacillaceae*, or a sum of these three beneficial bacteria. (B) Correlation between diarrheal duration and *Lachnospiraceae* relative abundance. Blue dots, patients with diarrhea > 10 days; red dots, patients with diarrhea <10 days. (C) SPF C57BL/6 mice received a high dose of total body radiation and were monitored

for >120 days. Fecal samples were collected from each single mouse at the indicated time points. Mice that survived over day 30 after radiation are referred to as elite-survivors and mice that died are referred to as non-survivors. (D) Survival rates ($n = 25$ in experiment 1 and $n = 27$ in experiment 2). Fractions indicate the number of mice that survived to the end of the experiment. Fecal 16S rRNA was sequenced. (E) PCA plot. (F) *Lachnospiraceae* and *Enterococcaceae* relative abundances. A.D., all mice died. Each symbol represents one patient [(A) and (B)] or one mouse [(C) to (E)]. Error bars indicate SEM, * $P < 0.05$, ** $P < 0.01$ determined by Mann-Whitney test [(A) and (F)] and log-rank (Mantel Cox) test (D).

and 30 after radiation treatment. Survival was monitored for >120 days (Fig. 3D). Mice that survived >30 days were considered elite-survivors. 16S rRNA sequencing was performed on fecal samples from elite-survivors

and those that died after radiation treatment (non-survivors). Before radiation (day -1), PCA showed that there was no difference in the bacterial compositions between elite-survivors and non-survivors. However, by day 7 after ra-

diation, elite-survivors harbored a bacterial community that was distinct from those that subsequently died (Fig. 3E). Representatives of *Lachnospiraceae* and *Enterococcaceae* showed greater abundance in elite-survivors by day 7

after radiation compared with mice that died (Fig. 3F). After 7 days, elite-survivors continued to show a greater abundance of these bacterial taxa, whereas nonsurvivors progressively died before day 21. These data indicate that radiation treatment had a selective effect on *Lachnospiraceae* and *Enterococcaceae* taxa in a host “survivor phenotype.”

For monoassociation studies, we selected bacteria that were (i) identifiable to the genus or family level and showed increased abundance in the ES-Recip group (Fig. 2M), (ii) cultivatable for functional studies in vitro and in vivo (11), and (iii) typed strains to ensure reproducibility (11). We tested radioprotection mediated by a mixture of 23 individual strains in the family *Lachnospiraceae*. We also tested *Enterococcus faecalis* and *Bacteroides fragilis*. These were compared with strains that showed reduced abundance in ES-Recip mice from the FMT experiment (Fig. 2M), including *Staphylococcus aureus*, *Staphylococcus saprophyticus*, and *Streptococcus agalactiae*. We also tested *Lactobacillus rhamnosus*, which has been previously reported to be radioprotective (5, 7). We inoculated SPF C57BL/6 mice with the selected bacteria by oral gavage twice a week for 8 weeks (Fig. 4A). The bacterial growth medium brain heart infusion (BHI) was used as a control. After bacterial reconstitution, recipients received a high dose of total body radiation (8.2 to 8.5 Gy). Mice inoculated with *Lachnospiraceae* elicited the greatest improvement in survival rate and clinical score (Fig. 4, B to E). Mice given *E. faecalis* and *L. rhamnosus* showed a 40 to 60% increase in survival rate (Fig. 4B). Mice gavaged with *B. fragilis*, *S. aureus*, *S. saprophyticus*, or *S. agalactiae* showed no improvement of outcome compared with the BHI controls (Fig. 4, B and C, and fig. S3).

Recipients of *Lachnospiraceae* (Lachno-Recip) also showed significantly increased bone marrow cellularity and splenic recovery of the white and red pulp regions (Fig. 4F). Control mice showed more colonic crypt distortion and atrophy, with gaps between the crypt bases and the muscularis mucosa, a common epithelial response to injury (Fig. 4G, indicated by arrows). By contrast, the crypts of mice given *Lachnospiraceae* remained closely attached to the muscularis mucosa. In control mice given BHI, the intestinal villi were flattened, but this effect was rescued by administration of *Lachnospiraceae* (Fig. 4G). Fluorescein isothiocyanate (FITC)-dextran assay indicated that after radiation, mice receiving *Lachnospiraceae* showed reduced gut permeability compared with BHI controls (Fig. 4H).

The radioprotection apparently afforded by *Lachnospiraceae* raises the possibility that these bacteria may not only be of therapeutic value to patients receiving radiotherapy but may also induce an unintended radioprotective

effect to tumor cells. To test whether inoculation of *Lachnospiraceae* attenuated the antitumor efficacy of radiotherapy, we introduced *Lachnospiraceae* or BHI into SPF mice for 8 weeks as described above, and then subcutaneously injected B16 (melanoma) or EL4 (lymphoma) cells into the abdomen (fig. S4A). *Lachnospiraceae* treatment did not affect tumor growth in either tumor model (fig. S4, B and C). To test the effect of *Lachnospiraceae* on radiotherapy, we transplanted *Lachnospiraceae* into mice, followed by inoculation of B16 tumor cells. Ten days later, we treated tumor-bearing mice with 10 Gy of localized radiation at the tumor site and monitored survival and tumor growth (fig. S4D). *Lachnospiraceae* treatment did not negatively affect the efficacy of radiation in reducing tumor growth (fig. S4, E to G).

Metabolite mediators of radioprotection

Members of the *Lachnospiraceae* are known for their ability to synthesize short-chain fatty acids (SCFAs) by fermentation of dietary polysaccharides (12). SCFAs are important substrates for maintaining intestinal epithelium and for regulating the immune system and inflammatory response (13, 14). We tested whether enhanced SCFA production plays a role in the ameliorative effect of *Lachnospiraceae* on ARS. Concentrations of three SCFAs (acetate, butyrate, and propionate) were greater in fecal samples from elite-survivors or ES-Recip mice from the dirty cage-sharing experiment compared with controls (fig. S5, A and B). The effects of butyrate approached significance ($P = 0.054$), showing a clear upward trend, and the levels of propionate and total SCFAs were significantly higher in ES-Recip mice.

SPF C57BL/6 mice were treated with acetate-, butyrate-, or propionate-supplemented water and then given 8.0 to 8.2 Gy of total body radiation (Fig. 5A). The 30-day survival rate was 79% for the propionate-treated group compared with 28% for the control group (Fig. 5, B and C). Acetate and butyrate provided more modest protection (Fig. 5, B and C). To confirm the role of SCFAs in microbiota-mediated radioprotection, we measured the amount of SCFAs produced by several strains of *Lachnospiraceae* that were either high- or low-SCFA producers (fig. S6, A and B). Subsequent treatment with three high-producer strains protected 100% of mice against radiation, whereas the three low-producer strains only provided 50% protection (fig. S6, C to E).

Histologic analyses revealed that propionate treatment resulted in elevated bone marrow cellularity and splenic white and red pulp recovery (Fig. 5D). Additionally, propionate attenuated radiation-induced loss of granulocyte-macrophage progenitors (GMPs), common myeloid progenitors (CMPs), and megakaryocyte-erythroid progenitors (MEPs)

(Fig. 5E) (15). Furthermore, alcian blue and periodic acid-Schiff (AB/PAS) staining of intracellular mucin glycoproteins within goblet cells (16) showed increased mucus thickness and crypt length in propionate recipients compared with controls (Fig. 5F).

We investigated whether SCFAs mediate radioprotection in mice by alleviating the effects of DNA damage and the release of reactive oxygen species (ROS) generated by ionizing radiation. After SCFA treatment, a high dose of total body radiation (8.4 Gy) was administered and 24 hours later, primary intestinal epithelial cells, intestinal intraepithelial lymphocytes, and bone marrow stem cells were isolated. Levels of DNA damage-related proteins, including γ H2AX (Fig. 5G), p53, and 53BP1 (fig. S7), were monitored by Western blot. Phosphorylation of these markers was reduced in butyrate- and propionate-treated groups compared with controls in both GI and hematopoietic tissues. Intracellular ROS levels in bone marrow stem cells (Fig. 5H) were significantly enhanced after radiation, whereas butyrate or propionate reduced the ROS levels by 50 to 60% of that found in non-radiation-naïve mice. Finally, because the presence of *Lachnospiraceae* provided long-term protection, we also found a protective effect of propionate >400 days (Fig. 5I).

To identify whether metabolites other than SCFAs have radioprotective benefits, an unbiased metabolomics approach was performed on fecal samples from elite-survivors and age-matched control mice (17, 18). Total ion chromatogram metabolomic cloudplot and PCA plot showed very distinct metabolite profiles between elite-survivors and controls (Fig. 6, A and B). From 3787 ion features that were altered between these two groups of mice, we identified 141 individual structures that changed substantially (tables S4 and S5 and fig. S8A). The most highly enriched metabolites from the feces of elite-survivors clustered in the tryptophan metabolic pathway, with a fivefold change seen in level of indole-3-carboxaldehyde (I3A) and an eightfold change in the level of kynurenic acid (KYNA) (Fig. 6, C and D, and fig. S8B). These tryptophan metabolites were tested for radiomitigation in vivo, and both metabolites were found to significantly improve survival and reduce the clinical scores of radiation-treated SPF mice (Fig. 6, E to G). Specifically, both I3A- and KYNA-treated groups showed survival rates of ~75%, with the protection persisting for >200 days after radiation (Fig. 6H).

Discussion

We have discovered that a gut microbiome-metabolome network can offer substantial protection against radiation-induced damage in mice. Our studies showed that the presence of *Lachnospiraceae* and *Enterococcaceae* was

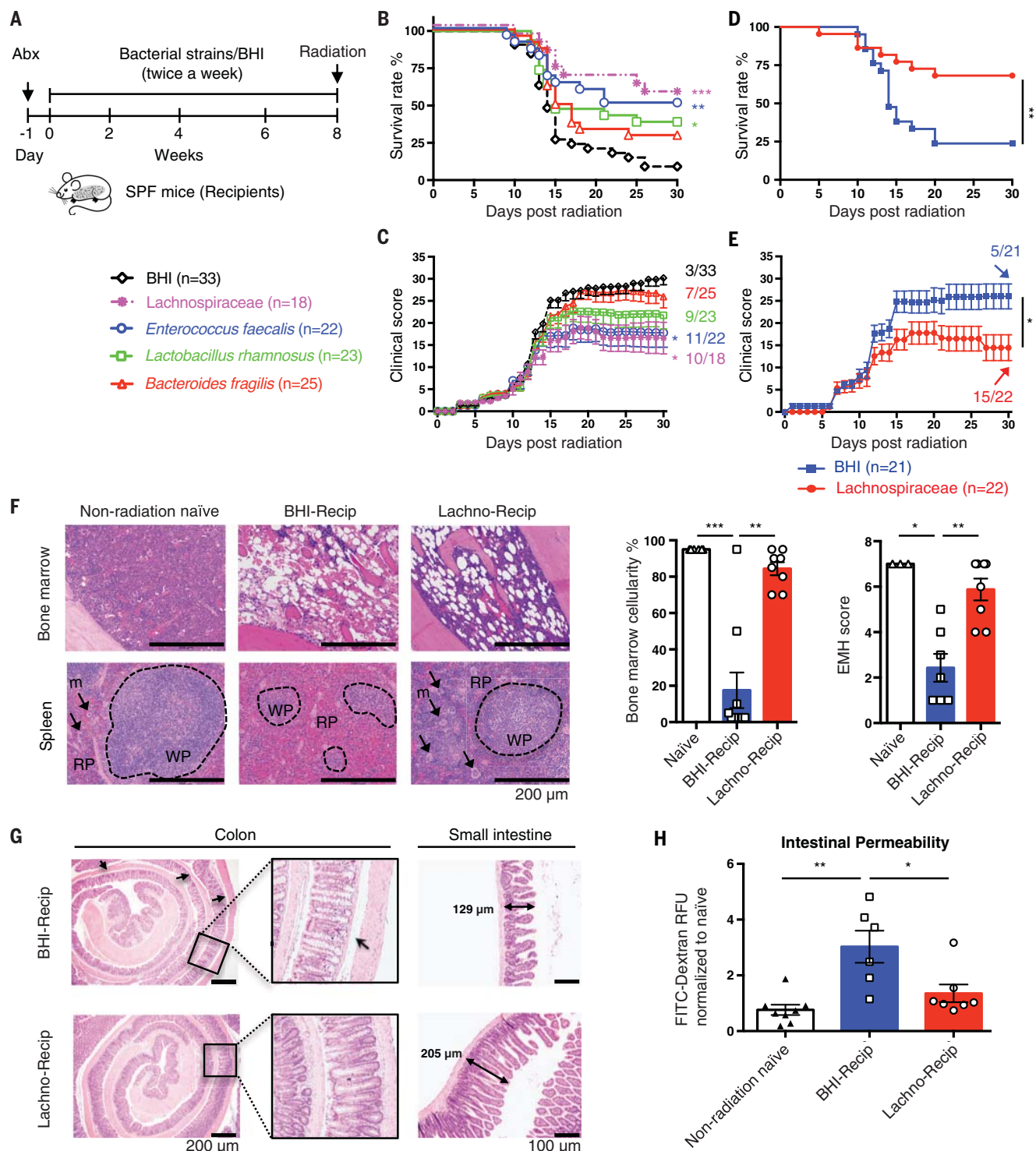


Fig. 4. Administration of *Lachnospiraceae* ameliorates radiation-induced syndromes. (A) Illustration of transfer experiments with bacteria [*Lachnospiraceae*] (a mixture of 23 individual strains in the family *Lachnospiraceae*), *E. faecalis*, *L. rhamnosus*, and *B. fragilis*] compared with media control (BHI). (B and C) Survival rates (B) and clinical scores (C). Significance is shown between the BHI group and each bacteria group. (D and E) Survival rates (D) and clinical scores (E) from *Lachnospiraceae* transfer experiments. Fractions indicate the number of mice that survived to the end of the experiment. (F) Femurs and spleens were stained with H&E and quantified for bone marrow cellularity and spleen extramedullary hematopoiesis scores. White pulp (WP), black dashed circles; red pulp (RP), area outside of WP;

megakaryocytes (m), black arrows. (G) Colons and small intestines were collected at day 1 after radiation. Arrows indicate gaps between crypt bases and muscularis mucosa. Double-headed arrows indicate the dimensions measured for villi length in small intestines. Representative histology from two independent experiments with seven to eight samples per group is shown. (H) Gut permeability was tested at day 1 after radiation. Fluorescence intensity in sera of each sample was normalized to non-radiation-naïve group average. Data are pooled from three [(B) to (E)] or two [(F) and (H)] independent experiments. Error bars indicate SEM, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ determined by log-rank (Mantel Cox) test [(B) and (D)], Mann-Whitney test for AUC [(C) and (E)], and Mann-Whitney test [(F) and (H)].

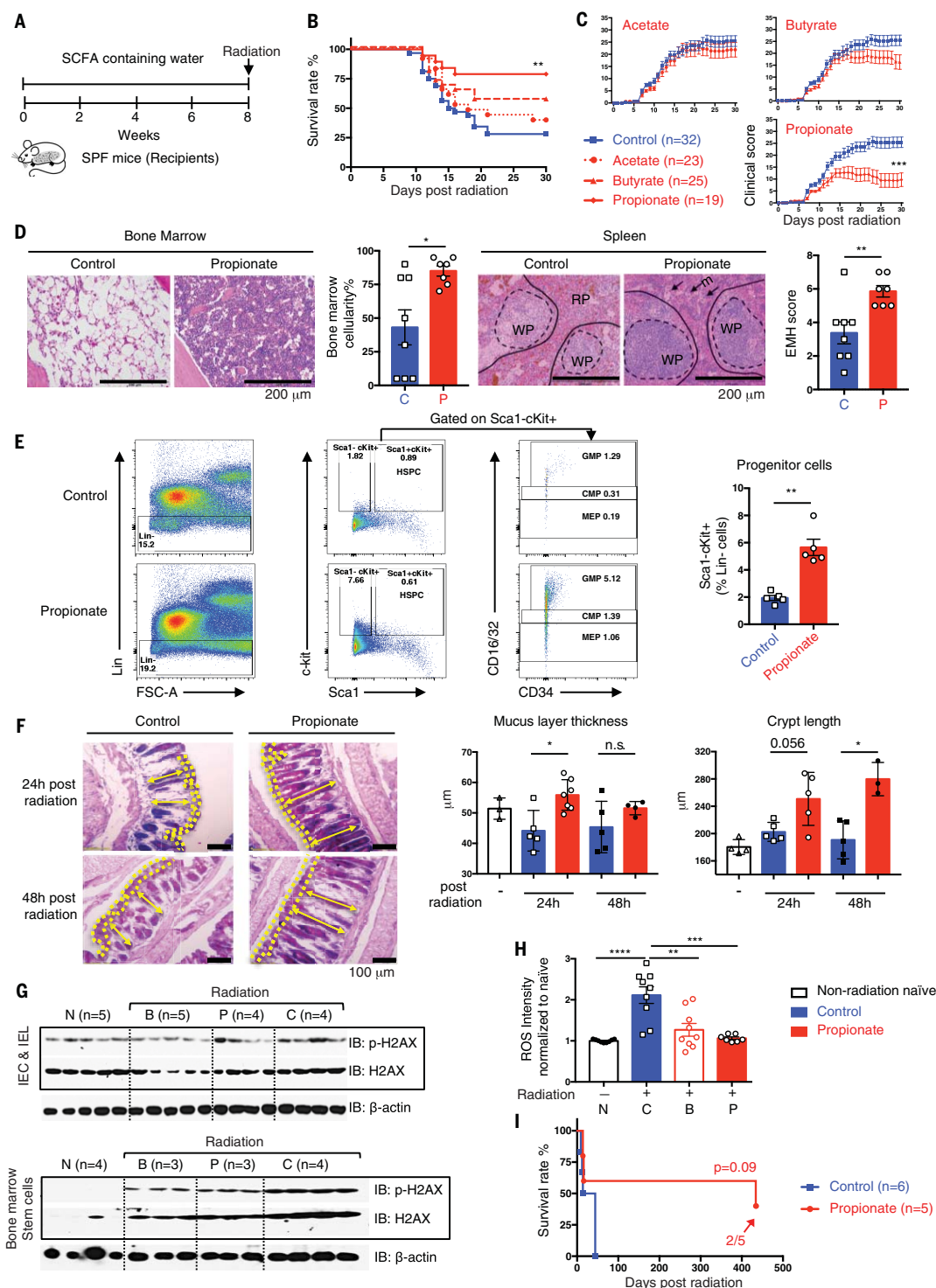


Fig. 5. Commensal-associated SCFAs suppress radiation-induced death and damage. (A) Schematic of SCFA treatment. (B and C) Survival rates (B) and clinical scores (C). (D) Femurs and spleens were stained with H&E and quantified as described in Fig. 4F. (E) Flow cytometric analysis of hematopoietic progenitor cells from bone marrow (CMPs, Lin⁻Sca1⁺cKit⁺CD16/32^{int}; GMPs, Lin⁻Sca1⁺cKit⁺CD16/32^{hi}; MEPs, Lin⁻Sca1⁺cKit⁺CD16/32^{lo}). (F) Representative images of AB/PAS staining in colon. Mucus layer, area between yellow dash lines; crypt length, yellow double-headed arrow. Mucus layer thickness and crypt length were quantified. (G) Intestinal epithelial cells, intestinal intraepithelial lymphocytes, and bone marrow stem cells were isolated (N, non-radiation naive; B, butyrate; P, propionate; C, control) at

24 hours after radiation. Phosphorylated and total H2AX were detected by Western blot. (H) ROS levels in bone marrow stem cells were detected. Fluorescence intensity of each sample was normalized to non-radiation-naïve group average. (I) Long-term survival in propionate compared with control treated groups were monitored for >400 days after radiation. Fractions indicate the number of mice that survived to the end of the experiment. Data are pooled from three [(B) and (C)] or two [(D) and (H)] independent experiments or represent two independent experiments [(E) to (G)]. Error bars indicate SEM, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$ determined by log-rank (Mantel Cox) test [(B) and (I)], Mann-Whitney test for AUC (C), and Mann-Whitney test [(D), (E), (F), and (H)].

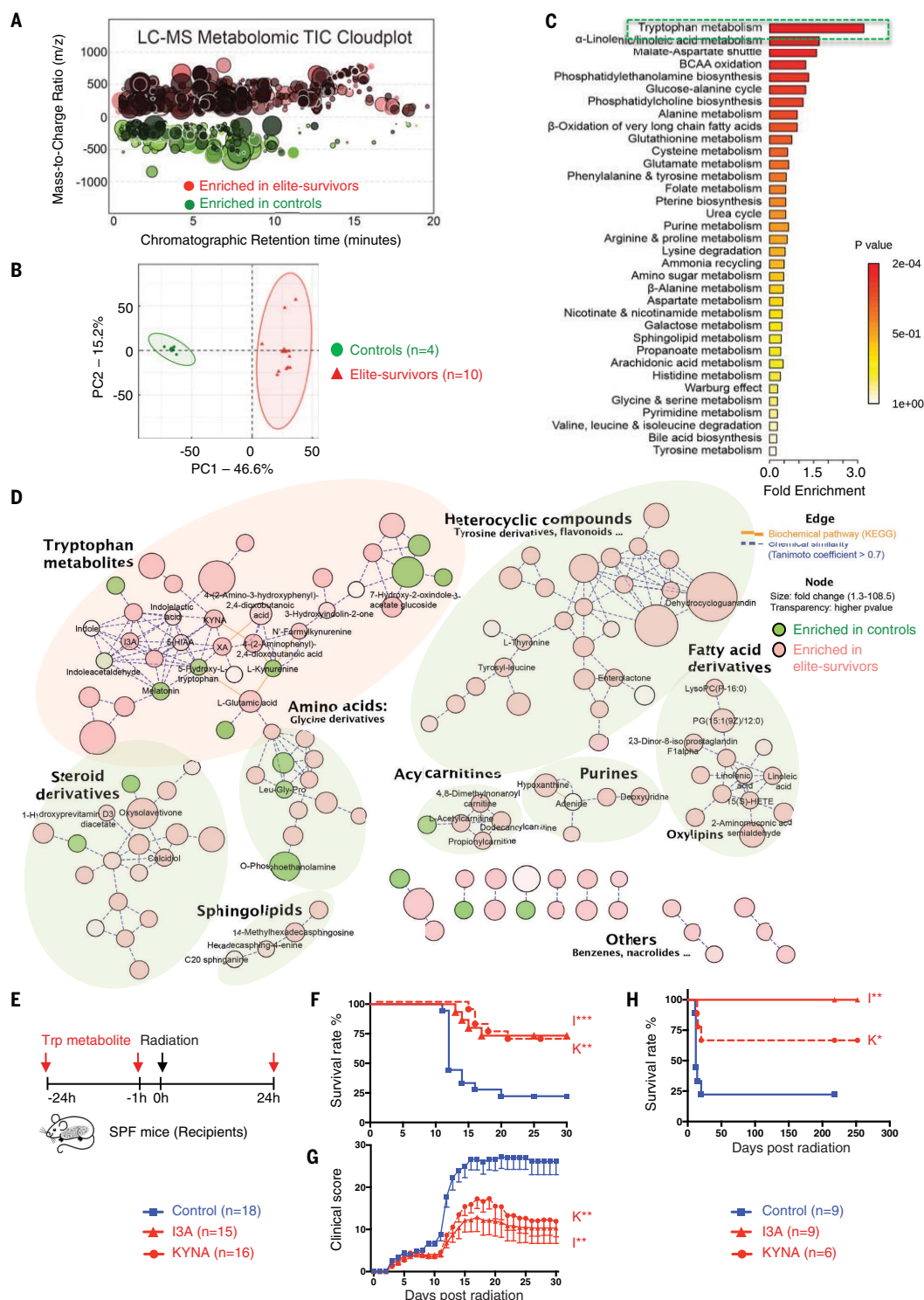


Fig. 6. Untargeted metabolomics reveals tryptophan metabolites as potent radioprotectants. Metabolite profiles were measured in fecal samples of elite-survivors and control mice at day 290 after total body radiation. (**A** and **B**) Total ion chromatogram metabolomic cloudplot ($P < 0.01$) (**A**) and PCA plot (**B**) of fecal metabolites. (**C**) Metabolite set enrichment analysis identified and interpreted metabolites in biochemical contexts. (**D**) Metabolic network integrated biochemical pathways and chemical relationships of all detected metabolites. Identified metabolites are represented by circular nodes, with lower transparency reflecting lower *P* values from Welch's *t* test. Red nodes, metabolites with higher abundance in elite-survivors; green nodes, those higher

in controls. Orange lines connecting metabolites symbolize Kyoto Encyclopedia of Genes and Genomes reactant pair links; green lines symbolize chemical similarity with a Tanimoto coefficient score > 0.7 . Tryptophan metabolites are highlighted by a pale pink shadow; others are distinguished by pale green shadows. (**E**) Schematic of tryptophan metabolite treatment. (**F** and **G**) Survival rates (**F**) and clinical scores (**G**). (**H**) Long-term survival in tryptophan metabolites compared with control treated groups were monitored for >200 days after radiation. Data are pooled from three [(**F**) and (**G**)] or two (**H**) independent experiments. Error bars indicate SEM, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ determined by log-rank (Mantel Cox) test [(**F**) and (**H**)] and Mann-Whitney test for AUC (**G**).

associated with restoration of hematopoiesis and gastrointestinal repair. These two taxa of bacteria were also found to be more abundant in leukemia patients undergoing radiotherapy who showed mild gastrointestinal dysfunction. Propionate, as well as two tryptophan pathway metabolites, I3A and KYNA, were elevated in elite-survivors and provided long-term protection against radiation. This investigation shows that the life-promoting radioprotection of both acute and delayed radiation damage can be provided by intestinal microbiota and their metabolites.

Lachnospiraceae and the microbial metabolite butyrate play an ameliorating role in colitis, food allergy, and graft-versus-host disease (10, 19, 20). SCFAs are metabolized from gut microbial fermentation of indigestible foods in the distal intestine (21), and can directly enter the systemic circulation and act as signaling molecules on peripheral tissues (22). Our work shows that SCFAs can directly promote hematopoiesis and intestinal repair after radiation. Broad inhibition of systemic inflammation, either by reducing the levels of proinflammatory cytokines or by inducing the anti-inflammatory cytokine interleukin-10, mediated by SCFAs may also provide radioprotection (23–26). In addition, this study shows that the tryptophan metabolites I3A and KYNA protect against ARS. Tryptophan metabolites are also known to attenuate proinflammatory cytokine responses (27, 28), and were shown by our untargeted metabolomics approach to be elevated in radiation survivors. Tryptophan metabolites can directly reduce radiation-induced damage and are candidate radioprotective agents. In addition to tryptophan metabolites, we also saw elevations of other metabolite families in the elite-survivor mice. Future studies will be needed to investigate whether these compounds also have radioprotective effects.

Treatment options for the adverse side effects of radiation exposure are limited and expensive, and our study indicates a promising route for further translational investigations.

Materials and Methods

Mice

All experiments used littermate controls or their immediate descendants. Male and female wild-type C57BL/6J mice were obtained from The Jackson Laboratory and house-raised at the University of North Carolina-Chapel Hill (UNC) for at least nine generations. Conventionally raised mice were bred and housed in SPF conditions, and GF animals were generated and housed in the National Gnotobiotic Rodent Resource Center of UNC in AAALAC-accredited facilities. Six- to eight-week-old mice were used in experiments unless otherwise noted. Both male and female mice were used, and both genders responded the same

to our bacteria or metabolites. More details are provided in the supplementary materials.

Total body radiation

Nonanesthetized mice were placed in ventilated plastic pie cages and exposed to 8.0 to 9.2 Gy of total body radiation (dose was dependent on sex, age, and SPF/GF). After radiation, mice were housed in sterile autoclaved cages and provided standard chow and water ad libitum unless otherwise noted. Mice were monitored for changes in body weight and other body parameters through 30 days after radiation unless otherwise noted. A clinical score was determined using a cumulative scoring system (table S2) based on weight loss, temperature change, physical appearance, posture, mobility, food consumption, and hydration (29). The last observation carried forward method was used when interpreting the clinical score plots. More details are provided in the supplementary materials.

Elite-survivors

Wild-type C57BL/6J mice received a high dose of total body radiation as described above. Age-matched littermates without radiation treatment were used as controls. After 30 days of monitoring, mice that showed no radiation-induced damage or injury were considered as elite-survivors. At day 290 after radiation, fecal samples were collected for microbiome analyses. Dirty cages from elite-survivors and controls were reserved for the dirty cage-sharing experiment and feces were collected for the FMT experiment.

Dirty cage-sharing experiment

Dirty cages containing feces and used bedding from age-matched controls or elite-survivors were collected every week. Six- to 8-week-old SPF C57BL/6J mice (recipients) were kept in the dirty cages, with mice being transferred to new dirty cages weekly, over a span of 8 weeks. One day after the final dirty cage-housing period, recipients were treated with a high dose of total body radiation. Fecal DNA was collected before radiation for microbiome analysis.

FMT experiment

Feces were collected from controls or elite-survivors. Freshly collected feces were diluted in saline at a ratio of 40 mg feces/ml saline, homogenized, and filtered with a stainless steel sieve (pore size: 0.25 mm). Total bacterial protein concentration was determined by BCA assay to normalize the number of bacterial cells. Fecal mixtures were mixed with 10% autoclaved glycerol, aliquoted, and frozen at -80°C until use. GF or SPF C57BL/6J mice were administered with 200 μl of the fecal mixture by oral gavage twice every week for 4 weeks (GF recipients) or 8 weeks (SPF re-

cipients). Twenty-four hours after the last FMT, recipients were treated with a high dose of total body radiation. Fecal DNA was collected before radiation for microbiome analysis.

Bacteria strain administration

The bacteria taxa termed “*Lachnospiraceae*” in this study were a mixture of 23 strains that belong to the family *Lachnospiraceae* as previously reported (11) unless otherwise noted. These bacteria were cultured in an anaerobic chamber in BHI broth supplemented with 5% fetal bovine serum, 0.01% L-cysteine, and 1% corn starch. *E. faecalis* (ATCC #19433), *L. rhamnosus* (ATCC #53103), *B. fragilis* (ATCC #23745), *S. aureus* (ATCC #25904), *S. saprophyticus* (ATCC #49907), and *S. agalactiae* (ATCC #13813) were purchased from ATCC and cultured according to the manufacturer's instructions. SPF C57BL/6J mice first received antibiotic treatment (20 mg of streptomycin per mouse) by oral gavage. One day later, mice were gavaged with the indicated strains twice a week for 8 weeks. BHI medium was used as a vehicle control. After the last administration, recipients were treated with a high dose of total body radiation.

SCFA treatment

Sodium acetate (Sigma, #S5636), sodium butyrate (Sigma, #303410), or sodium propionate (Sigma, #P1880) (200 mM each) was provided to SPF C57BL/6J mice in drinking water for 8 weeks, and then mice were treated with a high dose of total body radiation. SCFA treatment continued during the monitoring process.

Tryptophan metabolites treatment

I3A (Sigma, #129445) and KYNA (Sigma, #K3375) were dissolved in dimethyl sulfoxide (DMSO). I3A and KYNA (200 mg/kg) were administered by oral gavage into SPF C57BL/6J mice in three doses 1 day before radiation, 1 hour before radiation, and 1 day after radiation. For each gavage, the injection volume was 75 to 100 μl , not to exceed 5 ml/kg (DMSO/body weight).

Histopathology

To assess tissue pathology, spleens, femurs, small intestines, and colons were collected and fixed in 10% neutral-buffered formalin, paraffin embedded, and sectioned (4- μm thickness) at the UNC Lineberger Animal Histopathology Core Facility. Before paraffin embedding, femurs underwent an additional decalcification step in Immunocal (StatLab, McKinney, TX). Slides were stained with hematoxylin and eosin (H&E) and semiquantitatively scored by a board-certified pathologist who was blinded to experimental conditions. Slides of spleen and femur sections were also stained with Ki67 antibody (Cell Signaling Technology, #12202) or cleaved caspase 3 antibody (Cell Signaling

Technology, #9661). More details are provided in the supplementary materials.

Clinical patient data

Human stool samples were obtained from 21 leukemia patients at the start of total body radiation given as pretransplantation conditioning before allogeneic hematopoietic stem cell transplantation (see table S3 for detailed information). DNA extraction, polymerase chain reaction amplification of genomic 16S ribosomal RNA V4 and V5 regions, and sequencing were performed as previously described (30).

Tumor inoculation and radiotherapy

Mice were inoculated subcutaneously with 100,000 B16F10 cells or 200,000 EL4 cells in 50% matrigel (Corning, #354234) on the abdomen flank. Tumor volume was measured every 3 days and was calculated using the formula $(\text{length} \times \text{width}^2) \times 0.5$. Mouse survival rate was recorded. In accordance with the approved IACUC protocol, mice showing any of the following symptoms were humanely euthanized: (i) weight loss >20%, (ii) ulcerated tumors, or (iii) size reaching 2 cm in any dimension in a single tumor. For radiotherapy, after 10 days of tumor inoculation, mice were anesthetized by isoflurane and immobilized on the operating table by medical tape. Tumor site-localized radiation was performed using an x-ray irradiator with a dose of 10 Gy.

Immunostaining of mucins and measurement of mucus layer and crypt length

Briefly, ~2-cm sections of jejunum and colons were harvested and mounted in CryoMount embedding medium, snap-frozen in liquid nitrogen, and stored at -80°C until use. Four-micrometer-thick sections were cut and mounted onto poly-L-lysine-coated microscope slides, allowed to air dry, and stored at -80°C . Mucus and goblet cell immunostaining was performed using the AB/PAS kit according to the manufacturer's protocol (Polysciences, #25086). Inner mucus thickness and crypt length were measured in at least 10 different areas of one slide and in more than five mice per group.

Intestine permeability detection

At day 1 after radiation, mice were fasted but allowed water for 4 hours. FITC-dextran (4 kD, Sigma, #46944) was given orally (50 mg/100 g body weight). At 2 hours after gavage, blood was collected and the fluorescence of serum samples was measured using a microplate reader (excitation, 490 nm; emission, 520 nm). Fluorescence intensity of each sample was normalized to the non-radiation-naïve group average.

Bone marrow stem cell, intestinal epithelial cell, and intestinal intraepithelial lymphocyte isolation

Bone marrow stem cells were obtained using a magnetic cell sorting system (MACS, Miltenyi

Biotech, #130-091-224) according to the manufacturer's protocols. Intestinal epithelial cells and intestinal intraepithelial lymphocytes were isolated from fresh colons using a digestion buffer containing 100 μl of DNase (1/50, Sigma, #DN25-1G), 20 μl Collagenase D (1/250, Roche, #11088874103) and 100 μl Dispase (1/50, Stem-cell, #07913). More details are provided in the supplementary materials.

Immunoblot analysis

Bone marrow stem cells, intestinal epithelial cells, intestinal intraepithelial lymphocytes or whole spleen tissues were homogenized in radio-immunoprecipitation buffer (Boston BioProducts, #BP-115) containing complete protease inhibitor (Roche, #11697498001) and PhosphoStop (Roche, #4906845001) per the manufacturers' protocols. Whole-cell lysates were subjected to SDS-polyacrylamide gel electrophoresis and immunoblot analysis. The following primary antibodies were used at a dilution of 1:1000: anti-histone H2AX (Cell Signaling Technology, #7631), anti-phospho-histone H2AX (Cell Signaling Technology, #9718), anti-p53 (Cell Signaling Technology, #2524), anti-phospho-p53 (Cell Signaling Technology, #9284), anti-53BP1 (Invitrogen, #PA1-16566), anti-phospho-53BP1 (Abcam, #ab70323), anti-cleaved caspase 3 (Cell Signaling Technology, #9664), and anti-actin-HRP (C-11) (Santa Cruz Biotechnology, #sc-1615). Goat antirabbit horseradish peroxidase (HRP) (Jackson Immune Research Laboratories, #111-035-144) and goat anti-mouse-IgG HRP (Jackson Immune Research Laboratories, #115-035-146) were used as secondary antibodies. Protein densitometry was quantified with ImageJ software.

ROS detection

Intracellular ROS levels in bone marrow stem cells were detected using the DCFDA Cellular ROS Detection Assay Kit (Abcam, #ab113851) according to the manufacturer's protocol. The average level of ROS in the non-radiation-naïve group was calculated, and all sample results were normalized to this average. More details are provided in the supplementary materials.

Flow cytometry analysis

Bone marrow was flushed from femurs and then red blood cells were lysed with ammonium chloride-potassium lysing buffer (Lonza, #10-548E). Cells were labeled with phycoerythrin (PE)/Cy7-conjugated anti-CD16/32 clone 93 (BioLegend, #101317), V450 lineage cocktail (anti-CD3e clone 500A2, anti-CD11b clone MI/70, anti-B220 clone RA3-6B2, anti-Ly-76 clone TER-II9, anti-Ly6G/C clone RB6-8C5, BD Biosciences, #561301), PE-conjugated anti-Sca1 clone E13-161.7 (BD Biosciences, #553336), and FITC-conjugated anti-c-kit clone 2B8 (BD Biosciences, #553354). Labeled cells were fixed in 4% paraformaldehyde (ThermoFisher, #AAJ19943K2)

and then analyzed using an LSRII flow cytometer (BD Biosciences). The resultant data were gated using FlowJo version 10.4 software (BD Biosciences). Hematopoietic stem and progenitor cells were defined as Lin[−]Sca1⁺c-kit⁺, CMPs as Lin[−]Sca1[−]c-kit⁺CD16/32^{int}, GMPs as Lin[−]Sca1[−]c-kit⁺CD16/32^{hi}, and MEPs as Lin[−]Sca1[−]c-kit⁺CD16/32^{low}.

Fecal DNA extraction, 16S rRNA gene sequencing, and data analysis

Fecal sample preparation and 16S rRNA sequencing were conducted as previously described (10, 31). Fecal DNA samples were amplified by polymerase chain reaction (PCR) using bar-coded primer pairs targeting the V1–V2 or V3–V4 region of the 16S rRNA gene. PCR amplicons were sequenced using the Mi-Seq Illumina sequencer. The resulting bacterial sequence fragments were analyzed in QIIME version 1.8.0. More details are provided in the supplementary materials.

SCFAs detection by gas chromatography–mass spectrometry quantitation

Fecal SCFA contents were quantified as described in Zheng *et al.* (32) with modifications. The instrumental analysis was conducted on a 7820A GC-5977B MSD system (Agilent Technologies, Santa Clara, CA, USA). The relative standard deviations of both intraday and interday precision of analysis were <10%. The limit of quantitation was determined to be 1 ng on the column for acetic acid, 1 ng for propionic acid, and 0.1 ng for butyric acid. More details are provided in the supplementary materials.

Untargeted metabolomics

Untargeted metabolomics of mouse feces was performed on a high-resolution accurate mass spectrometry-based platform coupling to a cutting-edge integrated chemoinformatics pipeline as modified from a previous publication (33). The system consisted of a Vanquish UHPLC and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfacing with a heated electrospray ionization source and a hybrid quadrupole-orbitrap mass analyzer operated at 70,000 mass resolution (Waltham, MA, USA). Stringent quality assurance and quality control procedures were applied, including timely mass calibration, sample randomization, and intermittent quality control injection.

All MS1 raw data were processed in XCMS (Scripps, La Jolla, CA) for peak picking, peak alignment, gap filling, and sample normalization. Compound identification was implemented in MS-DIAL 2.94 and MS-FINDER 3.12. All identifications were manually curated and validated as available using an in-house metabolomic database established based on authentic chemical standards. Confidence level of annotation was assigned to each structure as required by the

Metabolomics Standards Initiative. For pathway analysis and visualization, MetaboAnalyst version 4.0 was used to perform metabolite enrichment analysis and to create a heatmap showing a selection of 50 annotated structures with largest group difference; a MetaMapp approach was used to compute a metabolomic network clustering by biochemical and chemical relationship that was further visualized in CytoScape version 3.7.1. More details are provided in the supplementary materials.

Statistics

All values are presented in the figures as mean \pm SEM, with $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$; n.s., no significance. The number of animals (n) used in the experiments is indicated in the figures. For all data, one dot or lane represents one mouse or one patient. Linear fitting and normalization were performed in GraphPad Prism software. If not specified in the figure legends, statistical significance between two groups was determined by unpaired, two-tailed Student's t test, log-rank (Mantel Cox) test, or Mann-Whitney test, and significance between more than two groups was determined using two-way ANOVA in GraphPad Prism with the default parameter based on experience.

REFERENCES AND NOTES

- M. Hauer-Jensen, J. W. Denham, H. J. Andreyev, Radiation enteropathy: Pathogenesis, treatment and prevention. *Nat. Rev. Gastroenterol. Hepatol.* **11**, 470–479 (2014). doi: [10.1038/nrgastro.2014.46](#); pmid: [24686268](#)
- A. K. Shadad, F. J. Sullivan, J. D. Martin, L. J. Egan, Gastrointestinal radiation injury: Prevention and treatment. *World J. Gastroenterol.* **19**, 199–208 (2013). doi: [10.3748/wjg.v19.i2.199](#); pmid: [23345942](#)
- A. L. Kau, P. P. Ahern, N. W. Griffin, A. L. Goodman, J. I. Gordon, Human nutrition, the gut microbiome and the immune system. *Nature* **474**, 327–336 (2011). doi: [10.1038/nature10213](#); pmid: [21677749](#)
- P. A. Crawford, J. I. Gordon, Microbial regulation of intestinal radiosensitivity. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13254–13259 (2005). doi: [10.1073/pnas.0504830102](#); pmid: [16129828](#)
- M. A. Ciorba et al., Lactobacillus probiotic protects intestinal epithelium from radiation injury in a TLR-2/cyclo-oxygenase-2-dependent manner. *Gut* **61**, 829–838 (2012). doi: [10.1136/gutjnl-2011-300367](#); pmid: [22027478](#)
- A. Wang et al., Gut microbial dysbiosis may predict diarrhea and fatigue in patients undergoing pelvic cancer radiotherapy: A pilot study. *PLOS ONE* **10**, e0126312 (2015). doi: [10.1371/journal.pone.0126312](#); pmid: [25955845](#)
- T. E. Riehl et al., Lactobacillus rhamnosus GG protects the intestinal epithelium from radiation injury through release of lipoteichoic acid, macrophage activation and the migration of mesenchymal stem cells. *Gut* **68**, 1003–1013 (2019). doi: [10.1136/gutjnl-2018-316226](#); pmid: [29934438](#)
- C. Manichanh et al., The gut microbiota predispose to the pathophysiology of acute proctodistal proctitis. *Am. J. Gastroenterol.* **103**, 1754–1761 (2008). doi: [10.1111/j.1572-0241.2008.01868.x](#); pmid: [18564125](#)
- Y. S. Kim, J. Kim, S. J. Park, High-throughput 16S rRNA gene sequencing reveals alterations of mouse intestinal microbiota after radiotherapy. *Anaerobe* **33**, 1–7 (2015). doi: [10.1016/j.anaerobe.2015.01.004](#); pmid: [25600706](#)
- L. Chen et al., NLRP12 attenuates colon inflammation by maintaining colonic microbial diversity and promoting protective commensal bacterial growth. *Nat. Immunol.* **18**, 541–551 (2017). doi: [10.1038/ni.3690](#); pmid: [28288099](#)
- A. E. Reeves, M. J. Koenigsnecht, I. L. Bergin, V. B. Young, Suppression of Clostridium difficile in the gastrointestinal tracts of germfree mice inoculated with a murine isolate from the family Lachnospiraceae. *Infect. Immun.* **80**, 3786–3794 (2012). doi: [10.1128/IAI.00647-12](#); pmid: [22890996](#)
- N. Reichardt et al., Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *ISME J.* **8**, 1323–1335 (2014). doi: [10.1038/ismej.2014.14](#); pmid: [24553467](#)
- P. M. Smith et al., The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* **341**, 569–573 (2013). doi: [10.1126/science.1241165](#); pmid: [23828891](#)
- N. Arpaia et al., Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature* **504**, 451–455 (2013). doi: [10.1038/nature12726](#); pmid: [24226773](#)
- J. H. Kim et al., NRF2-mediated Notch pathway activation enhances hematopoietic reconstitution following myelosuppressive radiation. *J. Clin. Invest.* **124**, 730–741 (2014). doi: [10.1172/JCI70812](#); pmid: [24463449](#)
- X. Gao et al., Chronic stress promotes colitis by disturbing the gut microbiota and triggering immune system response. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E2960–E2969 (2018). doi: [10.1073/pnas.1720696115](#); pmid: [29531080](#)
- M. G. Rooks, W. S. Garrett, Gut microbiota, metabolites and host immunity. *Nat. Rev. Immunol.* **16**, 341–352 (2016). doi: [10.1038/nri.2016.42](#); pmid: [27231050](#)
- M. Levy, E. Blacher, E. Elinav, Microbiome, metabolites and host immunity. *Curr. Opin. Microbiol.* **35**, 8–15 (2017). doi: [10.1016/j.mib.2016.10.003](#); pmid: [27883933](#)
- N. D. Mathewson et al., Gut microbiome-derived metabolites modulate intestinal epithelial cell damage and mitigate graft-versus-host disease. *Nat. Immunol.* **17**, 505–513 (2016). doi: [10.1038/ni.3400](#); pmid: [26998764](#)
- T. Feehley et al., Healthy infants harbor intestinal bacteria that protect against food allergy. *Nat. Med.* **25**, 448–453 (2019). doi: [10.1038/s41591-018-0324-z](#); pmid: [30643289](#)
- G. T. Macfarlane, S. Macfarlane, Bacteria, colonic fermentation, and gastrointestinal health. *J. AOAC Int.* **95**, 50–60 (2012). doi: [10.5740/jaoacint.SGE.Macfarlane](#); pmid: [22468341](#)
- E. E. Canfora, J. W. Jocken, E. E. Blaak, Short-chain fatty acids in control of body weight and insulin sensitivity. *Nat. Rev. Endocrinol.* **11**, 577–591 (2015). doi: [10.1038/nrendo.2015.128](#); pmid: [26260141](#)
- M. C. Maa et al., Butyrate reduced lipopolysaccharide-mediated macrophage migration by suppression of Src enhancement and focal adhesion kinase activity. *J. Nutr. Biochem.* **21**, 1186–1192 (2010). doi: [10.1016/j.jnutbio.2009.10.004](#); pmid: [20149623](#)
- L. Liu et al., Butyrate interferes with the differentiation and function of human monocyte-derived dendritic cells. *Cell. Immunol.* **277**, 66–73 (2012). doi: [10.1016/j.cellimm.2012.05.011](#); pmid: [22698927](#)
- P. V. Chang, L. Hao, S. Offermanns, R. Medzhitov, The microbial metabolite butyrate regulates intestinal macrophage function via histone deacetylase inhibition. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2247–2252 (2014). doi: [10.1073/pnas.1322269111](#); pmid: [24390544](#)
- A. Trompette et al., Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nat. Med.* **20**, 159–166 (2014). doi: [10.1038/nm.3444](#); pmid: [24390308](#)
- M. Włodarska et al., Indoleacrylic acid produced by commensal Peptostreptococcus species suppresses inflammation. *Cell Host Microbe* **22**, 25–37.e6 (2017). doi: [10.1016/j.chom.2017.06.007](#); pmid: [28704649](#)
- S. Krishnan et al., Gut microbiota-derived tryptophan metabolites modulate inflammatory response in hepatocytes and macrophages. *Cell Rep.* **23**, 1099–1111 (2018). doi: [10.1016/j.celrep.2018.03.109](#); pmid: [29694888](#)
- C. J. Kurkjian et al., The toll-like receptor 2/6 agonist, FSL-1 lipopeptide, therapeutically mitigates acute radiation syndrome. *Sci. Rep.* **7**, 17355 (2017). doi: [10.1038/s41598-017-17729-9](#); pmid: [29230065](#)
- J. U. Peled et al., Microbiota as predictor of mortality in allogeneic hematopoietic cell transplantation. *N. Engl. J. Med.* **382**, 822–834 (2020). doi: [10.1056/NEJMoA1900623](#); pmid: [32101664](#)
- A. D. Truax et al., The inhibitory innate immune sensor NLRP12 maintains a threshold against obesity by regulating gut microbiota homeostasis. *Cell Host Microbe* **24**, 364–378.e6 (2018). doi: [10.1016/j.chom.2018.08.009](#); pmid: [30212649](#)
- X. Zheng et al., A targeted metabolomic protocol for short-chain fatty acids and branched-chain amino acids. *Metabolomics* **9**, 818–827 (2013). doi: [10.1007/s11306-013-0500-6](#); pmid: [23997757](#)
- Z. Lai et al., Identifying metabolites by integrating metabolome databases with mass spectrometry cheminformatics. *Nat. Methods* **15**, 53–56 (2018). doi: [10.1038/nmeth.4512](#); pmid: [29176591](#)

ACKNOWLEDGMENTS

Funding: This project was supported by NIAID U19-A1067798, R35-CA232109, T32-CA009156, NMSS CA10068, and P01-DK094779 (J.P.Y.T. and R.B.S.). We thank the following for core and technical support: UNC Flow Cytometry Core Facility, Microbiome Core Facility and National Gnotobiotic Rodent Resource Center [supported by NIH P40-OD010995, P30-DK34987, P30-DK056350, the Crohn's and Colitis Foundation and the Center for Gastrointestinal Biology and Disease (CGIBD)]. Animal histopathology was performed within the LCCC Animal Histopathology Core Facility at UNC [supported in part by an NCI Center Core Support Grant (CA16086)]. The metabolomics study was performed within the UNC Biomarker Mass Spectrometry Facility and the UNC Center for Environmental Health and Susceptibility (supported by NIEHS ES024950 and P30-ES010126). Clinical data were provided by the Duke Adult Blood and Marrow Transplant Program (supported by R21AG066388 (A.D.S., L.M.B., and N.J.C.), an ASH Scholar Award (A.D.S.), and Duke Pepper Center P30-AG028716-13 Mini#6 (A.D.S.) and Memorial Sloan Kettering Cancer Center [supported by K08HL43189 (J.U.P.), NCI P30 CA008748 (J.U.P.), R01-CA228358 (M.R.M.v.d.B.), R01-CA228308 (M.R.M.v.d.B.), R01-HL147584 (M.R.M.v.d.B.), P30 CA008748 (M.R.M.v.d.B.), Project 2 of P01-CA023766 (M.R.M.v.d.B.), Project 2 of P01-AG052359 (M.R.M.v.d.B.), and U01A124275 (M.R.M.v.d.B.)]. Flow cytometry was performed at the Duke Human Vaccine Institute Flow Cytometry Facility in the Immunology Unit of Duke Regional Biocointainment Laboratory (supported by UC6-A1058607). **Author contributions:** H.G. and J.P.Y.T. designed the experiments and wrote the manuscript. H.G. performed the animal experiments. W.-C.C. contributed to mouse harvests and flow cytometry. J.W.T. and W.J.B. contributed to mouse harvests and manuscript revisions. N.D.M. and H.G. performed the histopathological analyses. M.J.F., A.N.M., and G.D.S. performed flow cytometry. H.G., L.C., and X.T. contributed to microbiome analysis. Y.L. and K.Lu performed targeted and untargeted metabolomics. H.G., K.X.L., and X.L. performed Western blot. L.M.B., A.D.S., N.J.C., J.U.P., A.L.C.G., and M.R.M.v.d.B. provided clinical patients data. R.B.S. provided GF mice. All contributing authors have agreed to the submission of this manuscript for publication. **Competing interests:** J.P.Y.T. is a cofounder of IMMvention Therapeutics and GoldCrest Bio. H.G. and W.J.B. hold equity shares in GoldCrest Bio. J.P.Y.T., H.G., and W.J.B. are named as inventors on U.S. Provisional Patent Application Serial No. 62/779,776: "Lachnospiraceae Mitigates Against Radiation-Induced Hematopoietic/Gastrointestinal Injury And Death, And Promotes Cancer Control By Radiation." M.R.M.v.d.B. has received research support from Seres Therapeutics; has consulted, received honorarium from, or participated in advisory boards for Seres Therapeutics, Novartis, Evelo, Jazz Pharmaceuticals, Therakos, Amgen, Magenta Therapeutics, Merck & Co, Inc., Acute Leukemia Forum (ALF), and DKMS Medical Council (Board); has IP Licensing with Seres Therapeutics, Juno Therapeutics; and has stock options from Smart Immune. J.U.P. reports research funding, intellectual property fees, and travel reimbursement from Seres Therapeutics and consulting fees from DaVolterra. **Data and materials availability:** All sequencing data have been deposited to the European Nucleotide Archive (<https://www.ebi.ac.uk/ena/browser/home>) under the accession numbers PRJEB37999, PRJEB38109, PRJEB38110, PRJEB38111, PRJEB38112, PRJEB38115, PRJEB38116, PRJEB38117, PRJEB38118, and PRJEB38119. All other data generated or analyzed during this study are included in this article and its supplementary materials. All radiotherapy participants provided informed consent before their participation in the studies, which were approved by the institutional review board at Duke University (Pro00050975, Pro0006268, and Pro00078566). The Lachnospiraceae strains were kindly provided by V. B. Young (University of Michigan–Ann Arbor), and the authors do not have the right to disseminate them.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/eaay9097/suppl/DC1
Materials and Methods
Figs. S1 to S8
References
Tables S1 to S5
MDAR Reproducibility Checklist

[View/request a protocol for this paper from Bio-protocol.](#)

30 July 2019; resubmitted 13 April 2020
Accepted 3 September 2020
[10.1126/science.aay9097](#)

RESEARCH ARTICLE

PLANT SCIENCE

Receptor kinase module targets PIN-dependent auxin transport during canalization

Jakub Hajný^{1,2}, Tomáš Prát¹, Nikola Rydza³, Lesia Rodriguez¹, Shutang Tan¹, Inge Verstraeten¹, David Domjan¹, Ewa Mazur^{3,4}, Elwira Smakowska-Luzan^{5,6}, Wouter Smet^{7,8}, Eliana Mor^{7,8}, Jonah Nolf^{7,8}, BaoJun Yang^{7,8}, Wim Grunewald^{7,8}, Gergely Molnár^{1,9}, Youssef Belkadir⁶, Bert De Rybel^{7,8}, Jiří Friml^{1*}

Spontaneously arising channels that transport the phytohormone auxin provide positional cues for self-organizing aspects of plant development such as flexible vasculature regeneration or its patterning during leaf venation. The auxin canalization hypothesis proposes a feedback between auxin signaling and transport as the underlying mechanism, but molecular players await discovery. We identified part of the machinery that routes auxin transport. The auxin-regulated receptor CAMEL (Canalization-related Auxin-regulated Malectin-type RLK) together with CANAR (Canalization-related Receptor-like kinase) interact with and phosphorylate PIN auxin transporters. *camel* and *canar* mutants are impaired in PIN1 subcellular trafficking and auxin-mediated PIN polarization, which macroscopically manifests as defects in leaf venation and vasculature regeneration after wounding. The CAMEL-CANAR receptor complex is part of the auxin feedback that coordinates polarization of individual cells during auxin canalization.

Plant development flexibly adapts the plant's architecture and physiology to an ever-changing environment. Much of this adaptive development is characterized by self-organization of patterning processes, such as integration of new organs with the preexisting vascular network, rise of complex leaf venation patterns, and flexible vasculature regeneration around a wound.

Formation of organized vasculature from originally uniform tissues involves coordinated polarization of individual cells. The canalization hypothesis proposes that the plant hormone auxin acts as a polarizing cue by means of its directional intercellular flow and feedback between auxin signaling and transport (1). Auxin transport is mediated by polarly localized PIN auxin transport proteins (2), and thus, auxin signaling coordinating the repolarization of PINs in individual cells can generate auxin transport channels, demarcating the future position of forming vasculature.

The emergence of PIN-expressing auxin channels preceding vasculature formation has been observed in different plant species to connect newly formed organs (3) or lateral shoot branches (4) with a preexisting vasculature network, also during leaf venation (5), in embryogenesis (6), and during regeneration after wounding (7, 8). Similar PIN-expressing auxin channels arise from an artificial local auxin source, revealing that auxin is the necessary and sufficient signal for channel formation (4, 8).

It remains enigmatic how such auxin feedback on subcellular PIN localization that leads to coordinated tissue polarization can integrate directional and positional cues. Auxin transcriptionally regulates PIN expression (9) and inhibits PIN endocytic recycling (10), which may explain auxin-mediated PIN repolarization through de novo secretion and at a differential endocytosis rate of PIN proteins from the plasma membrane, leading to establishment of polarity (11). A mechanistic model of auxin canalization (12) predicts that PIN polarization away from the auxin source can arise from a combination of intracellular, transcriptional auxin signaling that regulates PIN abundance, and cell-surface auxin signaling that regulates PIN internalization rate (10), stabilizing PINs at the given cell side. This mechanism would sense an auxin gradient throughout the tissue and translate it into tissue polarization. Additionally, a (so far, elusive) short-range signaling mechanism would mediate coordination between individual cells during this process.

We identified the CAMEL (Canalization-related Auxin-regulated Malectin-type RLK)–

CANAR (Canalization-related Receptor-like kinase) cell-surface receptor complex, acting downstream of the canonical TIR1/AFB-WRKY23 auxin signaling, which is required for the auxin effect on PIN trafficking and polarity in individual cells as well as for coordinated tissue polarization during vascular formation and regeneration.

Results

Identification of potential auxin canalization regulators downstream of WRKY23

To gain more insight into the auxin regulation of PIN polarity, we designed a microarray experiment to find genes downstream of the *WRKY DNA-BINDING Protein 23* (*WRKY23*) transcription factor (TF) that regulates auxin-mediated PIN repolarization (fig. S1, A and B) (13). We obtained 14 genes as auxin-inducible, potential targets of WRKY23 (Fig. 1A and table S1). To identify among them regulators of auxin canalization, we isolated or used previously described corresponding loss-of-function mutants. We analyzed the vasculature formation during leaf development as a classical process that requires auxin feedback on PIN-dependent auxin transport (5). Typically, wild-type (WT) cotyledons form a conserved pattern of four loops as in most mutants, with exception of mutants in *AT5G40780* and *AT1G05700*, where we observed strong venation defects: (*AT5G40780/lht-1*: 41% and *AT1G05700/camel-1*: 45%; where % stands for any type of abnormality deviating from typical WT pattern) (fig. S1C).

Thus, we identified *AT5G40780* and *AT1G05700* acting downstream of TIR1/AFB-WRKY23 and being required for canalization-based processes such as leaf venation.

Malectin-type LRR receptor-like kinase CAMEL downstream of auxin signaling

AT1G05700 encodes a previously uncharacterized member of the leucine-rich repeat (LRR) receptor-like kinase (RLK) family from the subfamily I, and its extracellular domain consists of a large Malectin-like domain and three LRR repeats (Fig. 1C and fig. S1D). We named *AT1G05700* “CAMEL”.

Quantitative reverse transcription polymerase chain reaction (RT-PCR) on lines where *WRKY23* is either targeted inducibly to the nucleus (*35S::WRKY23-GR*) or engineered into a transcriptional repressor (*35S::WRKY23-SRDX*) (14) confirmed that *CAMEL* mRNA amounts increased after activation of WRKY23 and decreased after its repression (Fig. 1D). The JASPAR database of TFs (15) also predicted WRKY23 among the top candidates binding to a 2000-base pair (bp) *CAMEL* promoter (table S2). We confirmed an activation of *CAMELpro* by WRKY23 using the luciferase-based reporter system in *Nicotiana benthamiana*. A coexpression of *35S::WRKY23* and *CAMELpro::LUC* led to activation of luciferase activity (Fig. 1F). This

¹Institute of Science and Technology (IST), 3400 Klosterneuburg, Austria. ²Laboratory of Growth Regulators, The Czech Academy of Sciences, Institute of Experimental Botany and Palacký University, Šlechtitelů 27, CZ-78371 Olomouc, Czech Republic.

³Mendel Centre for Plant Genomics and Proteomics, Central European Institute of Technology (CEITEC), Masaryk University, Brno, Czech Republic. ⁴Faculty of Natural Sciences, Institute of Biology, Biotechnology and Environmental Protection, University of Silesia in Katowice, Jagiellońska 28, 40-032 Katowice, Poland.

⁵Laboratory of Biochemistry, Wageningen University, Stippeneng 4, 6708 Wageningen, the Netherlands. ⁶Gregor Mendel Institute (GMI), Austrian Academy of Sciences, Vienna BioCenter (VBC), Vienna, Austria. ⁷Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium. ⁸VIB Center for Plant Systems Biology, 9052 Ghent, Belgium.

⁹Department of Applied Genetics and Cell Biology, University of Natural Resources and Life Sciences, (BOKU), Vienna, Austria.

*Corresponding author. Email: jiri.friml@ist.ac.at

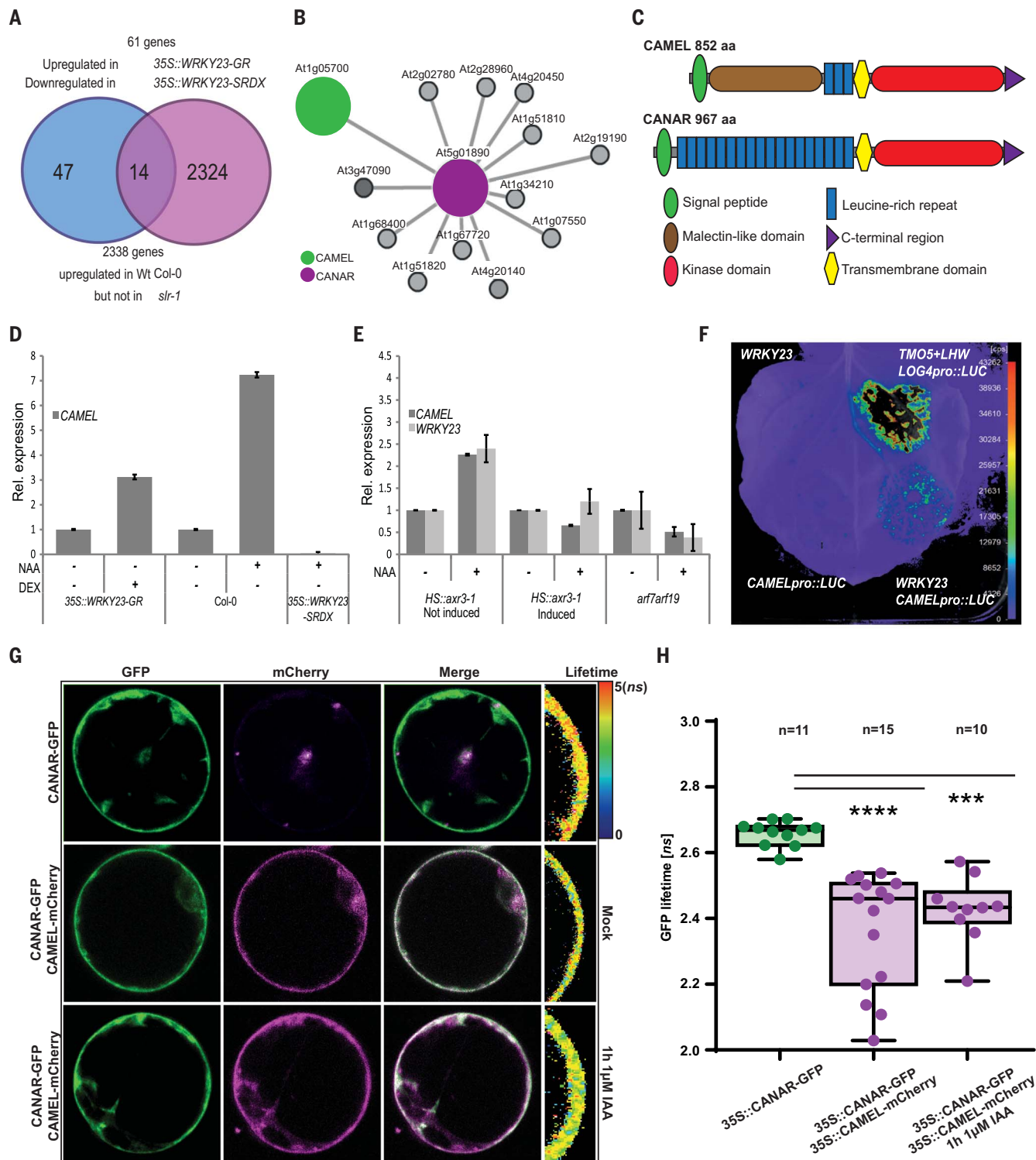


Fig. 1. CAMEL expression is regulated by WRKY23 and depends on the TIR1/AFB pathway. (A) Scheme of the microarray experimental setup to identify auxin-regulated genes downstream of the TIR1/AFB-WRKY23 signaling module. (B) Map of physical interactions between extracellular domains of putative interactors of CANAR from a previous study (16), with CAMEL being one of high confidence, illustrated with the BAR interaction viewer. (C) Schematic representation of the domain organization of CAMEL and CANAR. (D and E) Quantitative RT-PCR experiments showing that (D) *CAMEL* expression depends on WRKY23 and (E) auxin-mediated up-regulation of *CAMEL* requires the

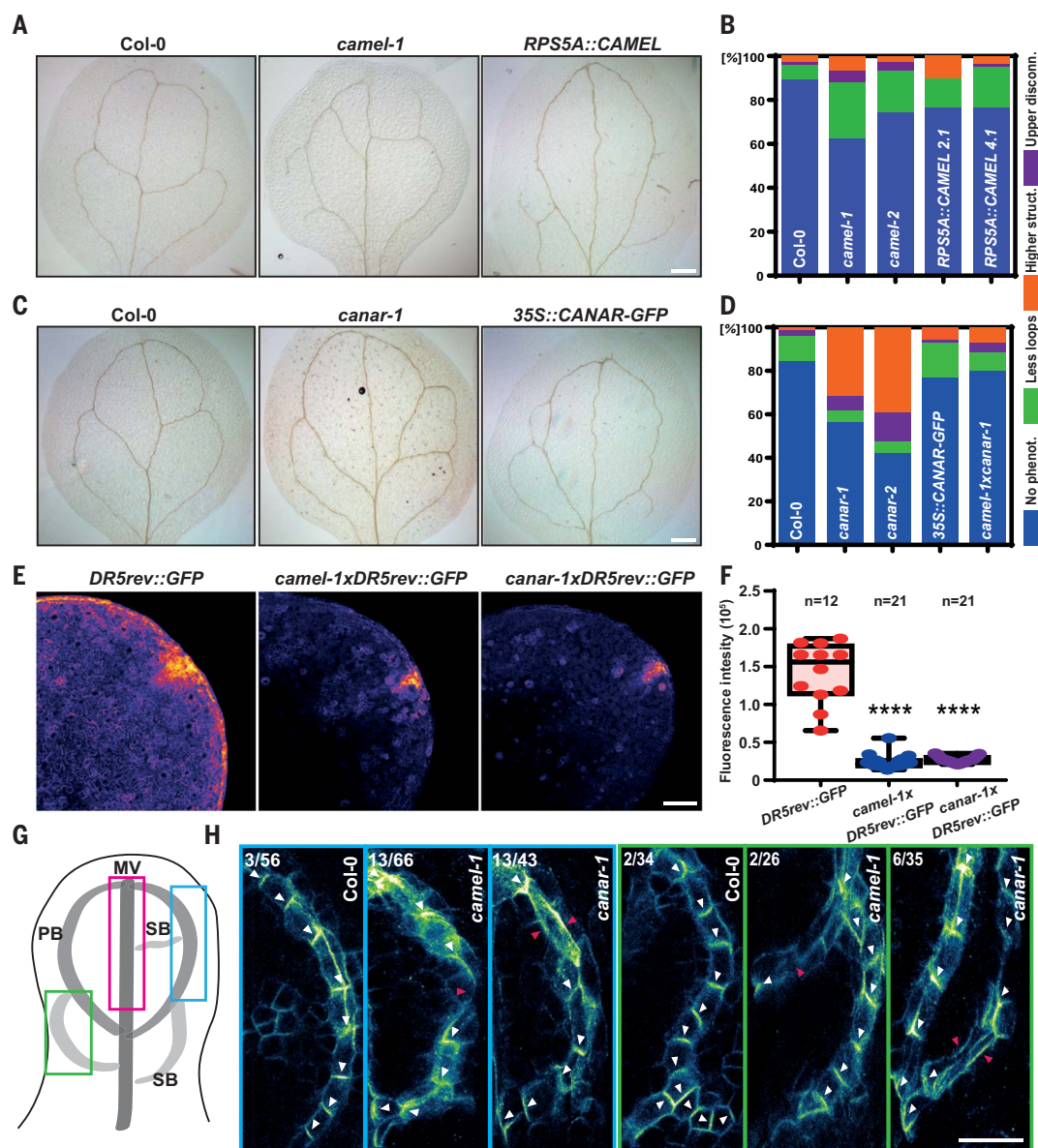
TIR1/AFB activity. (F) Luciferase assay in *Nicotiana benthamiana*. 35S::WRKY23 coexpressed with CAMELpro::LUC, and negative control of 35S::WRKY23 or CAMEL::LUC was indicated, respectively. (G and H) FRET-FLIM analysis of transiently expressed 35S::CANAR-GFP and 35S::CAMEL-mCherry in protoplasts. The GFP fluorescence lifetime was calculated as described in the supplementary materials, materials and methods, and the heatmap indicates the fluorescent lifetime values. A one-way analysis of variance (ANOVA) test compared marked sets of data (***P < 0.001; ****P < 0.0001). *n* denotes the number of scored protoplasts.

Fig. 2. Abnormal patterning and PIN1 polarization during leaf venation in *camel* and *canar* mutants. (A and C)

Representative images of venation patterning defects in cotyledons associated with (A) *camel-1* or (C) *canar-1*, respectively. Scale bar, 100 μ m.

(B and D) Quantification of venation defects in *camel* and *canar* mutants ($n > 75$ of each genotype). Scored categories: No phenotype, less loops, higher structure (including extra loops or branches), and upper disconnections. (E) DR5rev::GFP signal distribution and (F) intensity in cotyledons of *camel-1/canar-1* mutants ($n > 12$ of each genotype). Scale bar, 100 μ m.

(G and H) Coordinated PIN1 polarity in Col-0 and defective PIN1 polarity in *camel-1/canar-1* mutants. Colored boxes in (G) indicate positions of close-ups in (H). (H) Representative images of PIN1 immunolocalization in first leaves. The numbers in the left top corners indicate an incidence of observed PIN1 defective polarity to a total amount of analyzed leaves. White arrowheads indicate a typical PIN1 polar localization. Red arrowheads indicate a defective PIN1 polarity. Evaluation criteria are described in the supplementary materials, materials and methods. Scale bar, 10 μ m



supports that *CAMEL* is a downstream gene of WRKY23.

As shown previously, the TIR1/AFB auxin pathway is required for auxin-mediated PIN polarity rearrangements and canalization-based development (4, 8), and this goes in part through *WRKY23* (13). Consistently with this, *CAMEL* transcription, similar to *WRKY23*, is induced by auxin in a time- and dose-dependent manner (fig. S1, E and F), and this auxin effect is not observed in the *35S::WRKY23-SRDX* line or in mutants defective in transcriptional auxin signaling (*HS::axr3-1* and *arf7arf19*) (Fig. 1, D and E). Furthermore, *CAMELpro* contains six auxin-responsive elements (fig. S1G), suggesting additional auxin regulation, possibly directly by auxin response factors (ARFs), which are also supported by fast up-regulation of *CAMEL* by auxin (fig. S1F).

Thus, *CAMEL* is transcriptionally regulated by WRKY23 downstream of the TIR1/AFB-ARF signaling module.

CAMEL-CANAR signaling complex acts at the cell surface

To identify other receptor-like kinases that might interact with *CAMEL*, we used a map of physical interactions between extracellular domains of *Arabidopsis* LRR-RLKs (16). A large class VII LRR-RLK was retrieved as a single, high-confidence interactor (Fig. 1, B and C, fig. S1H). Given its association with *CAMEL* in canalization processes, we named it “CANAR”.

To determine whether *CAMEL* and *CANAR* interact in vivo as full-length proteins, we used three approaches. First, we immunoprecipitated *CANAR*-green fluorescent protein (GFP) using antibodies to GFP and performed tandem mass

spectrometry to identify proteins coimmunoprecipitated with our bait. We obtained a list with possible interactors, including *CAMEL* among the top 10 interactors (table S3). Next, we tested the *CAMEL-CANAR* interaction using bimolecular fluorescence complementation (BiFC) in *Nicotiana benthamiana* leaves. Coinfiltration of *CAMEL* and *CANAR* resulted in a positive signal in both combinations [*35S::CAMEL*-(C)CFP + *35S::CANAR*-(N)CFP and vice versa]. A negative control *35S::TMK2*-(C)CFP + *35S::CANAR*-(N)CFP did not show a signal (fig. S1I).

To test the *CAMEL-CANAR* interaction more quantitatively, we used Förster resonance energy transfer combined with fluorescence lifetime imaging microscopy (FRET-FLIM). We coexpressed *35S::CANAR-GFP* and *35S::CAMEL-mCherry* in *Arabidopsis* root protoplasts and detected a reduction in the GFP

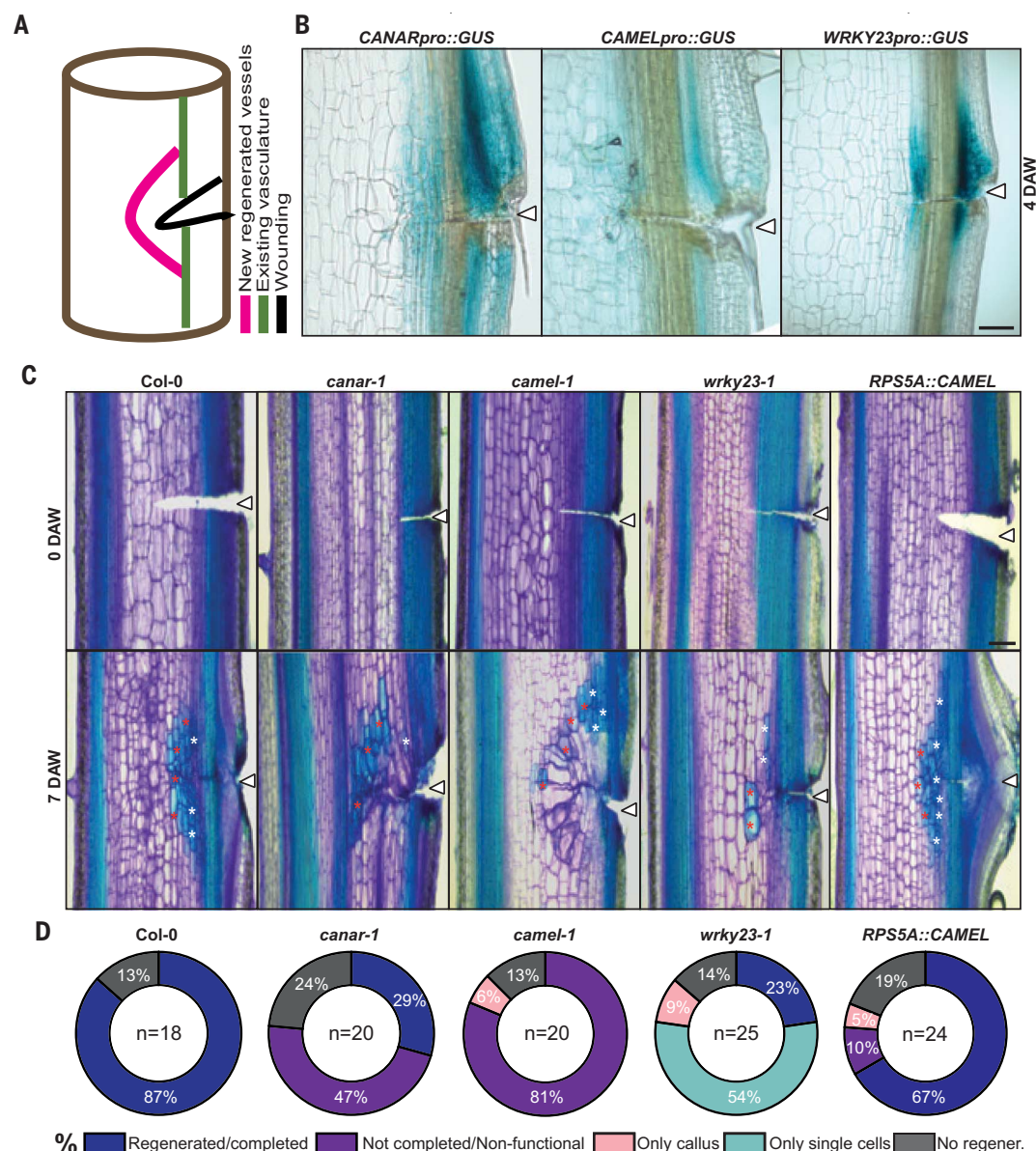


Fig. 3. Defective vasculature regeneration after wounding in *camel*, *canar*, and *wrky23-1* mutants. (A) Scheme of the experiment for analyzing vasculature regeneration after wounding. (B) Expression of *CANARpro::GUS*, *CAMELpro::GUS*, and *WRKY23pro::GUS* 4 days after wounding. The wound site is indicated with a white arrowhead. Scale bar, 100 μ m. (C) Wounded stems of *canar-1*, *camel-1*, and *wrky23-1* mutants and *RPS5A::CAMEL* overexpressing line at 0 and 7 days after wounding (DAW). Stems are stained with toluidine blue to visualize newly regenerated blue-vessel cells (white asterisks) and lignified parenchyma cells (red asterisks). The wound site is indicated with a white arrowhead. Scale bar, 100 μ m. (D) Quantification of regeneration defects in (C). *n* denotes the number of evaluated plants.

fluorescence lifetime as compared with that of *35S::CANAR-GFP* expressed alone (Fig. 1, G and H). Auxin treatment had no effect on CAMEL-CANAR interaction. Spatial resolution of FRET-FLIM experiments suggested that both proteins are localized and form complexes at the cell surface (Fig. 1G).

These observations showed constitutive, auxin-insensitive interaction between CAMEL and CANAR.

CAMEL and CANAR in leaf venation

Next, we tested genetically an involvement of CAMEL-CANAR complex in auxin-mediated canalization. We isolated publicly available transferred DNA (T-DNA) insertional loss-of-function mutants in both genes, *camel-1/-2* and *canar-1/-2* (fig. S2, A to D), and generated gain-of-function transgenic lines overexpressing *CAMEL* and *CANAR* under the constitu-

tive *RPS5A* and *35S* promoters, respectively (fig. S2, E and F).

The *camel-1* and *-2* mutants showed abnormal vascular patterning, with disconnected upper loops, extra branches, and extra or missing loops. Two independent overexpression lines also exhibited vasculature defects, albeit with lesser frequency (Fig. 2, A and B). Also, both *canar-1* and *-2* mutant alleles showed similar, even more pronounced vasculature abnormalities, whereas *35S::CANAR-GFP* exhibited overall less frequent defects (Fig. 2, C and D). The double mutant *camel-1xcanar-1* exhibited largely rescued venation, implying a possible antagonistic action of CAMEL and CANAR (Fig. 2, A and C, and fig. S2G). CAMEL/CANAR function appears to be rather specific to vasculature formation because no other obvious growth defects were observed (fig. S2, H to P).

Next, we tested whether these venation defects are linked to altered auxin distribution and auxin transport. We analyzed the expression of the auxin-responsive reporter *DR5rev::GFP*, which can be correlated with auxin distribution (17). In both *camel-1* and *canar-1* mutants, DR5 activity was decreased compared with that in the control (Fig. 2, E and F). When analyzing the basipetal (rootward) auxin transport in hypocotyls, we observed that both *camel-1* and *canar-1* mutants have reduced auxin flow (fig. S2Q). Given that formation of PIN1-expressing polarized channels has been linked to vein formation (5), we examined PIN1 polarity in young first leaves by means of staining with antibody to PIN1. In WT leaves, coordinated PIN1 polarity defining an auxin-transporting channel was observed with rare PIN1 polarity abnormalities in primary and secondary branches. By contrast,

both *camel-1* and *canar-1* mutants showed higher incidence of PIN1 polarity defects (Fig. 2H, red arrowheads) in primary and *canar-1* also in secondary branches (Fig. 2, G and H), whereas no defects were observed in midvein for any of the tested genotypes (fig. S2R).

These observations show that the CAMEL and CANAR mediate vasculature development during leaf venation and coordinate PIN1 polarities during this process.

CAMEL and CANAR in vasculature regeneration after wounding

Another classical example of a canalization-mediated process is vasculature regeneration after wounding when new vasculature is generated that circumvents the wound (7, 8).

We interrupted the vasculature in the *Arabidopsis thaliana* inflorescence stem with a horizontal cut (Fig. 3A) and analyzed *GUS* expression at 1 to 7 days after wounding. Both CAMEL and CANAR as well as their upstream regulator WRKY23 showed promoter activities during the regeneration (Fig. 3B and fig. S3A). Next, we analyzed the extent of vasculature regeneration in loss-of-function mutants and overexpressing lines visualized with toluidine blue (TBO) staining of regenerated vasculature. In wild type, vasculature was fully developed, and both newly regenerated vessel cells (Fig. 3C, white asterisk) and lignified parenchyma cells (Fig. 3C, red asterisk) stained in blue were visible. All tested mutants showed defective regeneration caused by the inability to form a continuous strand of regenerated cells. *RPS5A::CAMEL* showed less frequent defects, and *35S::CANAR-GFP* exhibited even improved regeneration over that of the wild type (Fig. 3, C and D, and fig. S3, B and C). Similar defects were observed also for flexible formation of auxin transport channels. In *PIN1-GFP*, but not in *canar-1xPIN1-GFP*, PIN1-GFP-expressing channel (fig. S3D, yellow arrows), circumventing the wound was formed. In contrast to leaf venation, *camel-1xcanar-1* double mutant showed regeneration defects comparable with those of the individual mutants (fig. S3, B and C).

To analyze more directly auxin-mediated formation of auxin transport channels, we used external, local auxin application (18). Application of an indole-3-acetic acid (IAA) droplet on the stem side below the wound (marked in magenta) led to the formation of a PIN1-expressing channel that connected with the preexisting vasculature already 2 days after wounding in wild type, an effect not observed in *canar-1* background even after 4 days after wounding (fig. S3E). Accordingly, the similar, newly formed vascular strands (as revealed with TBO staining) could be observed only in wild type (fig. S3F).

These results revealed a role for CAMEL and CANAR as well as their upstream regulator

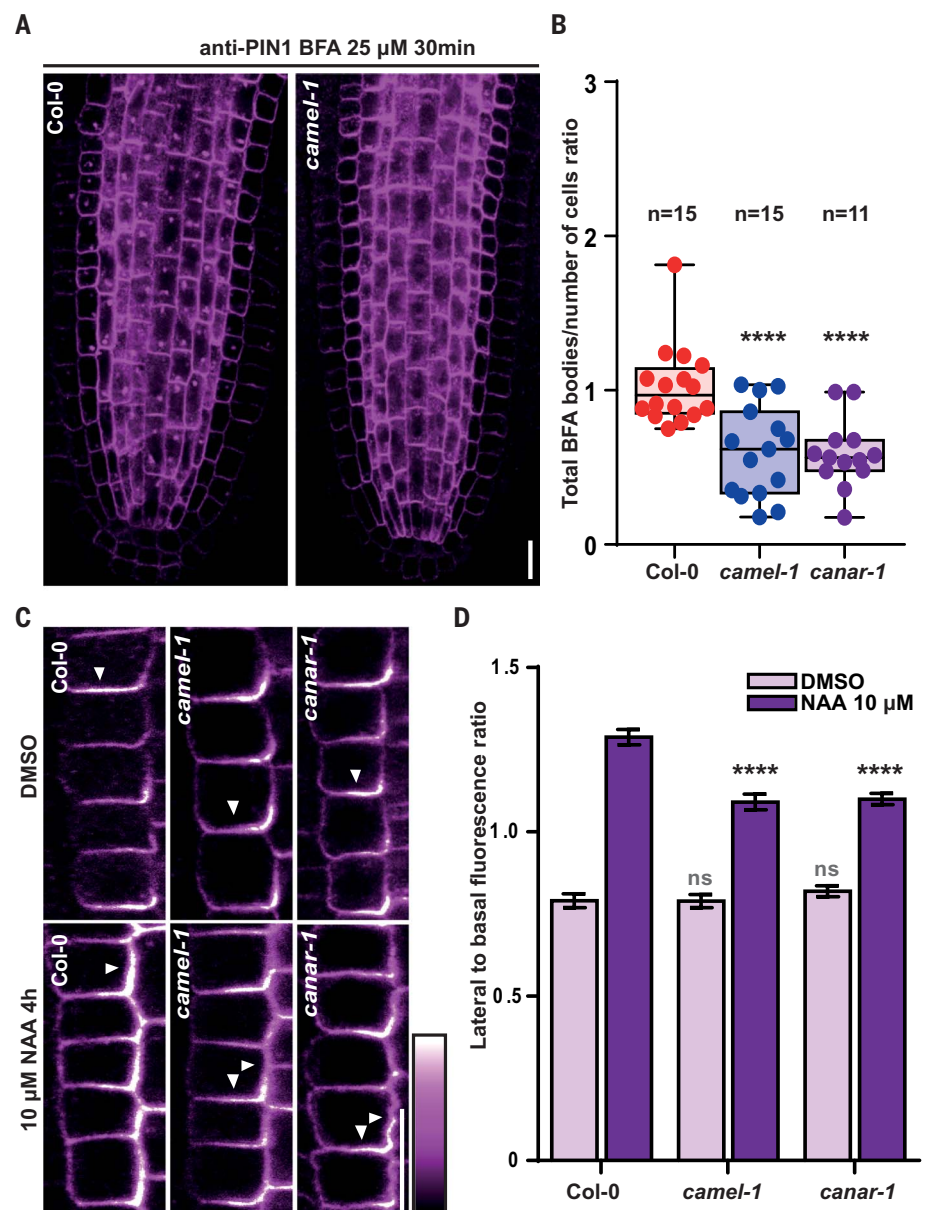


Fig. 4. Subcellular trafficking and auxin feedback on PIN polarity is compromised in *camel* and *canar* mutants. (A) Representative confocal images of primary root stele cells after immunostaining PIN1 in Wt, *camel-1*, and *canar-1*. Seedlings were BFA-treated (25 μ M) for 30 min. Scale bar, 10 μ m. (B) Quantitative evaluation of (A) shows ratio of total number of BFA bodies per total number of cells per root. *n* denotes the number of evaluated seedlings (*****P* < 0.0001). (C) Immunolocalization of PIN1 in endodermis of root meristem after 4 hours of 10 μ M naphthaleneacetic acid (NAA) treatment. Scale bar, 10 μ m. (D) Quantitative evaluation of (C) shows mean ratio of PIN1 lateral-to-basal signal intensity ratio in endodermal cells. Error bars indicate standard error. The experiment was carried out three times; one representative experiment is presented. A one-way ANOVA test compared marked datasets (*****P* < 0.0001; *n* > 80 cells corresponding to a minimum of 10 roots per treatment, and experiments were imaged using comparable settings).

WRKY23 in the flexible regeneration of vasculature after wounding.

CAMEL and CANAR in trafficking and polarity of PIN auxin transporters

Because auxin feedback on PIN polarity is one of the main features of canalization, and PIN polar localization is linked to its constitutive

endocytic recycling (11, 12), we tested whether CAMEL and CANAR are involved in these processes. PIN endocytic recycling can be indirectly visualized by means of PIN intracellular aggregation in response to treatment with the trafficking inhibitor Brefeldin A (BFA) (19). Anti-PIN1 immunostaining in roots showed that after BFA treatment, PIN1

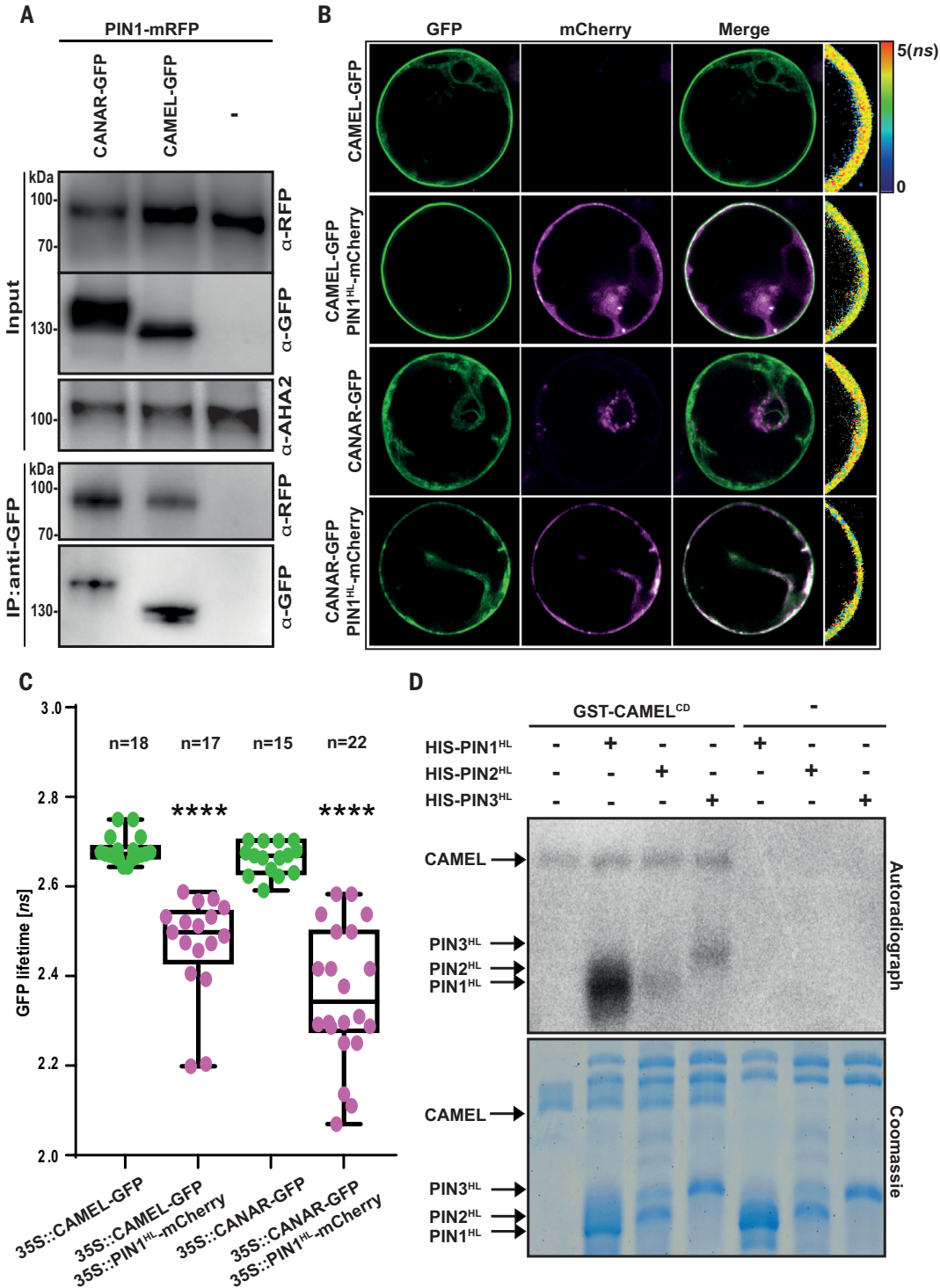


Fig. 5. The CAMEL-CANAR signaling module directly targets PIN1. (A) Coimmunoprecipitation assay from *Arabidopsis* root protoplasts. The protein complex of CAMEL-PIN1 and CANAR-PIN1 was coimmunoprecipitated by antibody-to-GFP beads. Antibody to AHA2 was used as the loading control. The experiment was carried out three times. (B and C) FRET-FLIM analysis of transiently expressed 35S::CAMEL-GFP and 35S::CANAR-GFP with 35S::PIN1^{HL}-mCherry (HL, hydrophilic loop) in protoplasts. The GFP fluorescence lifetime was calculated as described in the supplementary materials, materials and methods, and the heatmap indicates fluorescent lifetime values. A one-way ANOVA test compared the marked data sets (*****P* < 0.0001). *n* denotes the number of scored protoplasts. (D) Autoradiograph of an in vitro kinase phosphorylation assay of PIN1/2/3^{HL} by CAMEL^{CD} (CD, cytoplasmic domain). Aliquots of the samples were separated by SDS-polyacrylamide gel electrophoresis and exposed to autoradiography. The Coomassie Brilliant Blue was used as the loading control and to indicate the presence of the respective recombinant proteins. The blots shown are representative for three biological replicates.

intracellular aggregation was reduced in *camel-1* and *canar-1* mutants (Fig. 4, A and B). The same phenomenon was observed for *camel-1xPIN2-GFP* and *canar-1xPIN2-GFP* crosses (fig. S4, A and B), indicating a defect in the PIN endocytic recycling.

The auxin effect on PIN polarity can be approximated by the repolarization of PIN1 from the basal to the inner lateral side in the

root endodermis cells (8). Anti-PIN1 immunolocalization revealed that after auxin treatment, PIN1 repolarization was reduced in *camel* and *canar* mutants (Fig. 4, C and D, and fig. S4, C to E).

These results imply that the CAMEL-CANAR complex not only plays a role in the canalization-related development at the level of organs and tissues but also targets PIN1 in individual cells,

regulating its subcellular trafficking and auxin feedback on PIN polarity.

CAMEL-CANAR receptor complex targets and phosphorylates PIN auxin transporters

To get insight into the mechanism of CAMEL-CANAR action and downstream processes, we immunoprecipitated CAMEL-GFP from seedlings to identify the interactome of CAMEL.

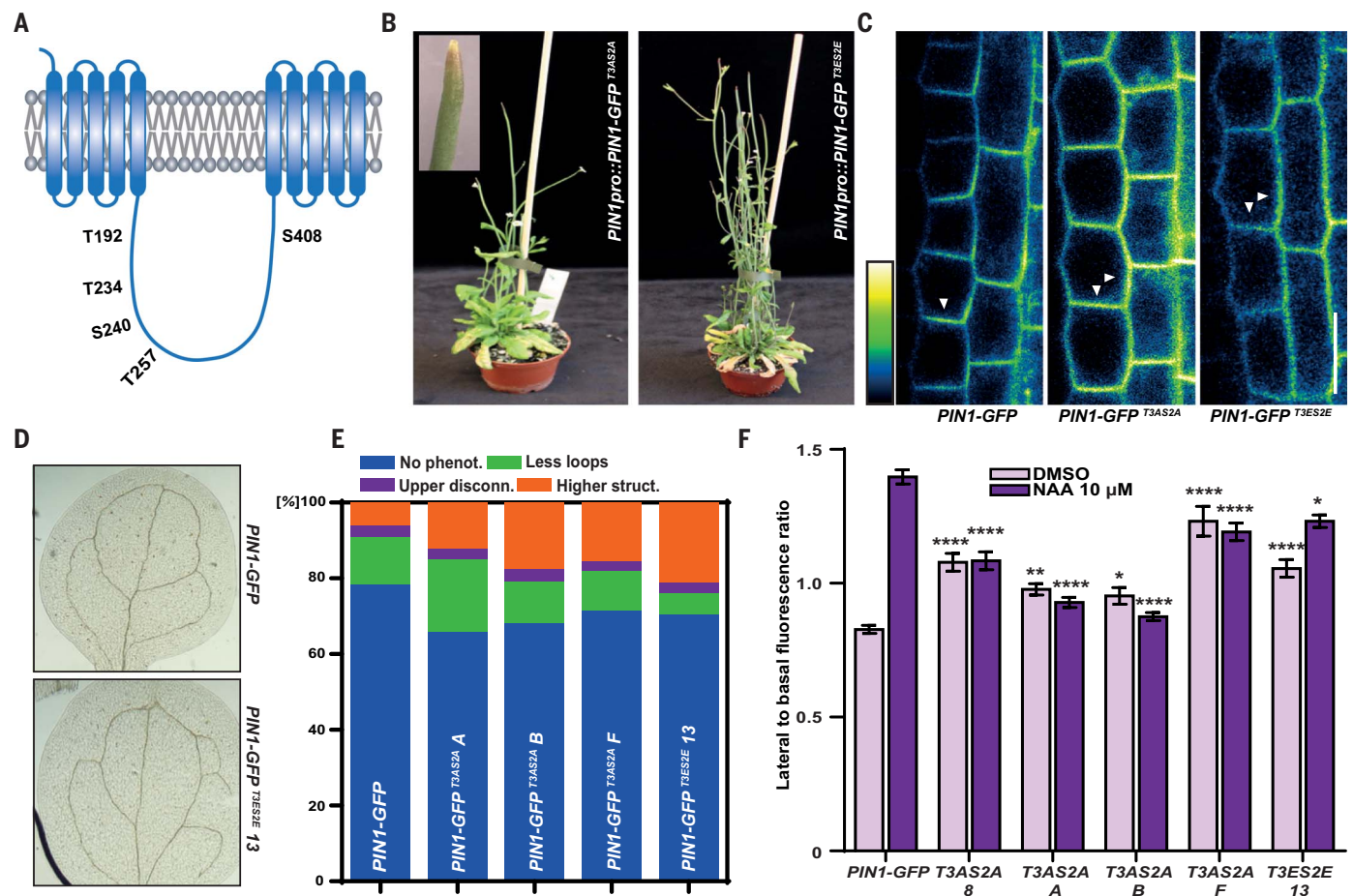


Fig. 6. CAMEL-targeted phosphosites in the PIN1 cytoplasmic loop are important for PIN polarity and venation. (A) Schematic representation of PIN1 in the plasma membrane, with marked positions of phosphosites in the cytoplasmic loop that are targeted by CAMEL. (B) Phenotypes of *PIN1pro::PIN1-GFP^{T3AS2A}* and *PIN1pro::PIN1-GFP^{T3ES2E}*, 35 days old. (C) Subcellular localization of PIN1-GFP, PIN1-GFP^{T3AS2A}, and PIN1-GFP^{T3ES2E} in root meristem endodermal cells. White arrows indicate the predominant subcellular localization. Scale bar, 100 μ m. (D) Representative images of vasculature phenotype in cotyledons of *PIN1pro::PIN1-GFP*, *PIN1pro::PIN1-GFP^{T3ES2E}* (line 13). (E) Quantification of vasculature defects in *PIN1pro::PIN1-GFP*,

PIN1pro::PIN1-GFP^{T3AS2A} (lines A, B, and F), and *PIN1pro::PIN1-GFP^{T3ES2E}* (line 13) ($n > 68$ of each genotype). Scored categories are normal vasculature, less loops, higher structure (including extra loops or branches), and upper disconnections. (F) Quantitative evaluation of (fig. S6D) shows the mean lateral-to-basal ratio of PIN1-GFP signal in endodermal cells. Error bars indicate standard errors. The experiment was carried out three times; one representative experiment is presented. A one-way ANOVA test was performed to compare marked datasets ($*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$; $n > 40$ cells corresponding to a minimum of eight roots per treatment, and experiments were imaged using comparable settings).

Coimmunoprecipitated proteins were analyzed by use of mass spectrometry (table S4). Among the list of putative interactors, multiple PIN proteins were found.

To confirm the interaction with PIN1, we transiently coexpressed *35S::CAMEL-GFP* + *35S::PIN1-mRFP* and *35S::CANAR-GFP* + *35S::PIN1-mRFP* in *Arabidopsis* root protoplasts. PIN1-mRFP coimmunoprecipitated with both CANAR-GFP and CAMEL-GFP (Fig. 5A). Furthermore, we performed FRET-FLIM in root protoplasts expressing *35S::CAMEL-GFP* or *35S::CANAR-GFP*. The lifetime of CAMEL-GFP and CANAR-GFP was reduced after co-expression with *35S::PIN1^{HL}-mCherry* (HL, hydrophilic loop), further confirming an interaction between CAMEL/CANAR and PIN1 (Fig. 5, B and C).

Because CAMEL and CANAR are expected to act as kinases and PIN phosphorylation is a well-established mode of regulation of PIN activity and polar localization (2), we tested the ability of CAMEL and CANAR to phosphorylate PINs. We therefore performed an *in vitro* kinase assay by incubating purified PIN1^{HL}, PIN2^{HL}, or PIN3^{HL} with purified cytoplasmic kinase domains of CAMEL and CANAR with radiolabeled adenosine triphosphate (ATP). We detected phosphorylation of PIN loops by CAMEL, with PIN1^{HL} being the best substrate (Fig. 5D). However, CANAR did not show kinase activity (fig. S5A). Lack of kinase activity can be explained by losing an aspartic acid from the conserved HRD motif in the catalytic core similarly to other known pseudokinases: BIR2, GHR1, PRK5 (fig. S5B).

Considering the lack of CANAR kinase activity, constitutive CAMEL-CANAR interaction, and complementation of leaf vasculature defects of *camel-lxcanar-1* double mutant, we hypothesize that CANAR might be a negative regulator of CAMEL. This is further supported by the ability of CANAR kinase domains to reduce CAMEL autophosphorylation and kinase activity toward PIN1 (fig. S5C).

To test relevance of CAMEL-mediated PIN1 phosphorylation, we analyzed *in vitro* kinase reaction using mass spectrometry and identified five mostly conserved putative phosphosites in PIN1^{HL} (Fig. 6, A to C, and fig. S5D). These sites seem distinct because they are not shared by any previously reported kinase phosphorylating PIN loop such as PID/WAGs, D6PK, or MPKs (Fig. S5D), and when mutated, they decreased

the ability of CAMEL kinase domain to phosphorylate PIN1^{HL} (fig. S6, A and B).

We generated phosphodead *PIN1-GFP^{T3AS2A}* and phosphomimic *PIN1-GFP^{T3ES2E}* constructs through the substitution of three threonine and two serine to alanine or glutamic acid, respectively, placed them under control of the native *PIN1* promoter, and introduced them into WT plants. Positive, GFP-expressing transformants for both constructs showed already in the first-generation naked inflorescence stems (7 of 20 for *PIN1-GFP^{T3AS2A}* and 3 of 18 for *PIN1-GFP^{T3ES2E}*), which is strongly reminiscent of *pin1* loss of function (Fig. 6B). Other positive plants did not show strong phenotypes and produced seeds, which allowed analysis in the next generation. Venation in cotyledons of both *PIN1-GFP^{T3AS2A}* and *PIN1-GFP^{T3ES2E}* lines exhibited increased incidence of vascular abnormalities (Fig. 6, D and E, and fig. S6C). All positive transformants in the first generation for *PIN1^{T3AS2A}* (4 of 4) and *PIN1^{T3ES2E}* (2 of 2) that showed naked inflorescence stems exhibited no vasculature regeneration after wounding, which was characterized by fragmented vessel cells, nonfunctional parenchyma cell connections, or extensive callus formation in the wound (fig. S6E). To test the role of the identified phosphosites in canalization, we tested auxin's effect on mutated PIN1 variants. Whereas PIN1-GFP is localized in roots predominantly basally, both *PIN1-GFP^{T3AS2A}* and *PIN1-GFP^{T3ES2E}* showed more apolar localization already, without any treatments (Fig. 6C and fig. S6D). When immunolocalized with antibody to GFP, both *PIN1-GFP^{T3AS2A}* and *PIN1-GFP^{T3ES2E}*, already partially polarized to the inner-lateral side in the mock situation, did not show any further polarity changes after auxin application (Fig. 6F and fig. S6D).

The CAMEL-CANAR complex interacts with PINs, and CAMEL is capable of phosphorylat-

ing their cytosolic loops. Effects of phosphomimic and phosphodead mutations in the PIN1 loop support the relevance of these phosphorylations for auxin transport and auxin canalization. The stronger defects in lines that carry PIN1 with mutated CAMEL-targeted phosphorylation sites as compared with the *camel/canar* mutants suggest that these phosphosites are shared by other kinases that control auxin transport in different developmental contexts.

Discussion

In this study, we provided mechanistic insights into how auxin controls its own directional cell-to-cell transport and identified molecular components of a so-called auxin canalization mechanism underlying flexible and self-organizing formation of auxin channels, guiding vasculature formation. Identification of the CAMEL-CANAR complex downstream of transcriptional auxin signaling and its direct regulation of PIN-dependent auxin transport provide a potential means for how to integrate global auxin signals with a so far hypothetical short-range signaling for the coordination of cell polarities during plant adaptive development.

REFERENCES AND NOTES

1. T. Berleth, T. Sachs, *Curr. Opin. Plant Biol.* **4**, 57–62 (2001).
2. M. Adamowski, J. Friml, *Plant Cell* **27**, 20–32 (2015).
3. E. Benková et al., *Cell* **115**, 591–602 (2003).
4. J. Balla, P. Kalousek, V. Reinöhl, J. Friml, S. Procházka, *Plant J.* **65**, 571–577 (2011).
5. E. Scarpella, D. Marcos, J. Friml, T. Berleth, *Genes Dev.* **20**, 1015–1027 (2006).
6. H. S. Robert et al., *Curr. Biol.* **23**, 2506–2512 (2013).
7. E. Mazur, E. Benková, J. Friml, *Sci. Rep.* **6**, 33754 (2016).
8. M. Sauer et al., *Genes Dev.* **20**, 2902–2911 (2006).
9. A. Vieten et al., *Development* **132**, 4521–4531 (2005).
10. T. Paciorek et al., *Nature* **435**, 1251–1256 (2005).
11. M. Glanc, M. Fendrych, J. Friml, *Nat. Plants* **4**, 1082–1088 (2018).
12. K. Wabnick et al., *Mol. Syst. Biol.* **6**, 447 (2010).
13. T. Prát et al., *PLOS Genet.* **14**, e1007177 (2018).
14. W. Grunewald et al., *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1554–1559 (2012).

15. A. Khan et al., *Nucleic Acids Res.* **46** (D1), D260–D266 (2018).
16. E. Smakowska-Luzan et al., *Nature* **553**, 342–346 (2018).
17. J. Friml et al., *Nature* **426**, 147–153 (2003).
18. E. Mazur, I. Kulik, J. Hajný, J. Friml, *New Phytol.* **226**, 1375–1383 (2020).
19. N. Geldner, J. Friml, Y.-D. Stierhof, G. Jürgens, K. Palme, *Nature* **413**, 425–428 (2001).

ACKNOWLEDGMENTS

We acknowledge M. Glanc and Y. Zhang for providing entry clones; Vienna Biocenter Core Facilities (VBCF) for recombinant protein production and purification; Vienna Biocenter Mass spectrometry Facility, Bioimaging, and Life Science Facilities at IST Austria and Proteomics Core Facility CEITEC for a great assistance. **Funding:** This project received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement 742985) and Austrian Science Fund (FWF): I 3630-B25 to J.F. and by grants from the Austrian Academy of Science through the Gregor Mendel Institute (Y.B.) and the Austrian Agency for International Cooperation in Education and Research (D.D.); the Netherlands Organization for Scientific Research (NWO: VIDI-864.13.001) (W.S.); the Research Foundation–Flanders (FWO: Odysseus II G0D0515N) and a European Research Council grant (ERC; StG TORPEDO: 714055) to B.D.R., B.Y., and E.M.; and the Hertha Firnberg Programme postdoctoral fellowship (T-947) from the FWF Austrian Science Fund to E.S.-L.; J.H. is the recipient of a DOC Fellowship of the Austrian Academy of Sciences at IST Austria. **Author contributions:** J.F. and W.G. conceived and designed the experiments. J.H. and J.F. wrote the paper, with help from B.D.R. and Y.B.; J.H., T.P., L.R., S.T., I.V., D.D., N.R., E.Ma., E.S.-L., W.S., E.Mo., B.Y., B.D.R., Y.B., and J.N. conducted experiments and contributed to the study design. G.M., J.H., and I.V. analyzed the data. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. Raw microarray data from this article can be found in the EMBL ArrayExpress repository under accession no. E-MTAB-9563.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/550/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S6
Tables S1 to S6
References (20–45)
MDAR Reproducibility Checklist

[View/request a protocol for this paper from Bio-protocol.](#)

22 November 2019; accepted 8 September 2020
10.1126/science.aba3178

ANCIENT DOG GENOMICS

Origins and genetic legacy of prehistoric dogs

Anders Bergström^{1*}, Laurent Frantz^{2,3*}, Ryan Schmidt^{4,5}, Erik Ersmark^{6,7}, Ophélie Lebrasseur^{8,9}, Linus Girdland-Flink^{10,11}, Audrey T. Lin^{8,12,13}, Jan Stora¹⁴, Karl-Göran Sjögren¹⁵, David Anthony^{16,17}, Ekaterina Antipina¹⁸, Sariah Amir¹⁹, Guy Bar-Oz²⁰, Vladimir I. Bazaliiskii²¹, Jelena Bulatović²², Dorcas Brown¹⁶, Alberto Carmagnini², Tom Davy¹, Sergey Fedorov²³, Ivana Fiore^{24,25}, Deirdre Fulton²⁶, Mietje Germonpré²⁷, James Haile²⁸, Evan K. Irving-Pease^{8,29}, Alexandra Jamieson⁸, Luc Janssens³⁰, Irina Kirillova³¹, Liora Kolska Horwitz³², Julka Kuzmanović-Cvetković³³, Yaroslav Kuzmin^{34,35}, Robert J. Losey³⁶, Daria Ložnjak Dizdžar³⁷, Marjan Mashkour^{19,38}, Mario Novak³⁹, Vedat Onar⁴⁰, David Orton⁴¹, Maja Pasarić⁴², Miljana Radivojević⁴³, Dragana Rajković⁴⁴, Benjamin Roberts⁴⁵, Hannah Ryan⁸, Mikhail Sablin⁴⁶, Fedor Shidlovskiy⁴⁷, Ivana Stojanović⁴⁷, Antonio Tagliacozzo²⁴, Katerina Trantalidou^{48,49}, Inga Ullén⁵⁰, Aritza Villaluenga⁵¹, Paula Wapnish⁵², Keith Dobney^{9,10,53,54}, Anders Götherström^{21,54}, Anna Linderholm⁵⁵, Love Dalén^{6,7}, Ron Pinhasi^{56*}, Greger Larson^{8*}, Pontus Skoglund^{1*}

Dogs were the first domestic animal, but little is known about their population history and to what extent it was linked to humans. We sequenced 27 ancient dog genomes and found that all dogs share a common ancestry distinct from present-day wolves, with limited gene flow from wolves since domestication but substantial dog-to-wolf gene flow. By 11,000 years ago, at least five major ancestry lineages had diversified, demonstrating a deep genetic history of dogs during the Paleolithic. Coanalysis with human genomes reveals aspects of dog population history that mirror humans, including Levant-related ancestry in Africa and early agricultural Europe. Other aspects differ, including the impacts of steppe pastoralist expansions in West and East Eurasia and a near-complete turnover of Neolithic European dog ancestry.

Wolves were the first animal with which humans formed a mutualistic relationship, eventually giving rise to dogs. Although there is little consensus regarding when (1–9), where (2, 8–13), and how many times (1, 8, 9, 14) domestication took place, the archaeological record (9, 15) attests to a long-term and close relationship to humans (9, 16–18). Modern dog genomes have revealed a complex population structure (5, 8, 10, 12, 19, 20), but because only six ancient dog and wolf genomes are currently available (4, 9, 14, 21), the process by which this structure emerged remains largely unknown.

Previous mitochondrial DNA (22–29) and genomic (9, 14, 21) studies have suggested an association between the genetic signatures of

dogs and their archeological context. However, dog and human genomes have not been quantitatively coanalyzed to assess the degree to which the population history of dogs was linked to that of humans—or may have been decoupled as a result of trade, human preference for particular types of dogs, variation in infectious disease susceptibility, or dogs moving between human groups.

To reconstruct dog population history, we sequenced 27 ancient dog genomes up to 10.9 thousand years (ka) old from Europe, the Near East, and Siberia (table S1) to a median of 1.5-fold coverage (range, 0.1- to 11-fold) (Fig. 1A and table S2) (30). To test the association with human population history, we compiled 17 sets of human genome-wide data (30) that matched the age, geographic location, and cultural con-

texts of the ancient dogs (table S4), and we directly compared genetic relationships within the two species.

Global dog population structure has its origins in the Pleistocene

To characterize the global population structure of ancient and modern dogs, we applied principal component analysis (PCA) to a matrix of all possible f_4 -statistics (30), alleviating differences in error rates and missing data. This approach recapitulates a major east–west axis of dog ancestry (PC1) (8, 9, 12), in which the western extreme comprises modern and ancient western Eurasian dogs and modern African dogs (Fig. 1B). The eastern extreme is represented by precontact North American dogs (21), three dogs from 7 ka ago from Lake Baikal in Siberia, and modern East Asian dogs, including New Guinea singing dogs and Australian dingoes. Similar results were obtained through standard model-based clustering (fig. S2).

All ancient and modern European dogs have greater affinity to eastern dog ancestry than ancient Near Eastern dogs have on the basis of f_4 tests (fig. S3), despite the overall east–west axis on PC1. Ancient European dogs are also distributed widely across a genetic cline between the East Eurasian and ancient Near Eastern dogs, which furthermore manifests as a linear cline along the diagonal when contrasting shared genetic drift with Baikal dogs and Levantine (Israel, 7 ka ago) dogs using outgroup f_3 -statistics (Fig. 1C). Simulations indicate that this linear, diagonal cline is difficult to explain with long-standing continuous gene flow or a tree-like history; instead, they suggest that the history of Mesolithic and Neolithic European dogs was marked by a major admixture episode (Fig. 1D) (30).

We modeled the genetic history underlying dog population structure for five populations that represent major ancestries and tested all 135,285 possible admixture graph models with up to two admixture events (30). Only one

¹Ancient Genomics Laboratory, The Francis Crick Institute, London, UK. ²School of Biological and Chemical Sciences, Queen Mary University of London, London, UK. ³Palaeogenomics Group, Department of Veterinary Sciences, Ludwig Maximilian University, Munich, Germany. ⁴School of Archaeology and Earth Institute, University College Dublin, Dublin, Ireland. ⁵CIBIO-InBIO, University of Porto, Campus de Vairão, Portugal. ⁶Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden. ⁷Centre for Palaeogenetics, Svante Arrhenius väg 18C, Stockholm, Sweden. ⁸The Palaeogenomics and Bio-Archaeology Research Network, Research Laboratory for Archaeology and History of Art, University of Oxford, Oxford, UK. ⁹Department of Archaeology, Classics and Egyptology, University of Liverpool, Liverpool, UK. ¹⁰Department of Archaeology, University of Aberdeen, Aberdeen, UK. ¹¹Liverpool John Moores University, Liverpool, UK. ¹²Department of Zoology, University of Oxford, Oxford, UK. ¹³Department of Anthropology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA. ¹⁴Stockholm University, Stockholm, Sweden. ¹⁵Gothenburg University, Gothenburg, Sweden. ¹⁶Hartwick College, Oneonta, NY, USA. ¹⁷Department of Human Evolutionary Biology, Harvard University, Cambridge, MA, USA. ¹⁸Institute of Archaeology of the Russian Academy of Sciences, Moscow, Russian Federation. ¹⁹Bioarchaeology Laboratory, Central Laboratory, University of Tehran, Tehran, Iran. ²⁰University of Haifa, Haifa, Israel. ²¹Irkutsk State University, Irkutsk, Russian Federation. ²²University of Belgrade, Belgrade, Serbia. ²³North-Eastern Federal University, Yakutsk, Russian Federation. ²⁴Bioarchaeology Service, Museo delle Civiltà, Rome, Italy. ²⁵Environmental and Evolutionary Biology Doctoral Program, Sapienza University of Rome, Rome, Italy. ²⁶Baylor University, Waco, TX, USA. ²⁷Royal Belgian Institute of Natural Sciences, Brussels, Belgium. ²⁸University of Copenhagen, Copenhagen, Denmark. ²⁹Lundbeck GeoGenetics Centre, The Globe Institute, Copenhagen, Denmark. ³⁰University of Leiden, Leiden, Netherlands. ³¹Ice Age Museum, Moscow, Russian Federation. ³²Hebrew University, Jerusalem, Israel. ³³Homeland Museum of Toplica, Prokuplje, Serbia. ³⁴Sobolev Institute of Geology and Mineralogy of the Siberian Branch of Russian Academy of Sciences, Novosibirsk, Russian Federation. ³⁵Tomsk State University, Tomsk, Russian Federation. ³⁶University of Alberta, Edmonton, AB, Canada. ³⁷Institute of Archaeology, Zagreb, Croatia. ³⁸Archéozoologie, Archéobotanique, Sociétés, Pratiques et Environnements, Centre National de la Recherche Scientifique, Muséum National d'Histoire Naturelle, Paris, France. ³⁹Centre for Applied Bioanthropology, Institute for Anthropological Research, Zagreb, Croatia. ⁴⁰Istanbul University—Cerrahpaşa, Istanbul, Turkey. ⁴¹University of York, York, UK. ⁴²Institute of Ethnology and Folklore Research, Zagreb, Croatia. ⁴³University College London, London, UK. ⁴⁴Archaeological Museum Osijek, Osijek, Croatia. ⁴⁵Durham University, Durham, UK. ⁴⁶Zoological Institute of the Russian Academy of Sciences, Saint Petersburg, Russian Federation. ⁴⁷Institute of Archaeology, Belgrade, Serbia. ⁴⁸Hellenic Ministry of Culture & Sports, Athens, Greece. ⁴⁹University of Thessaly, Argonafton & Philellinon, Volos, Greece. ⁵⁰National Historical Museums, Stockholm, Sweden. ⁵¹Consolidated Research Group on Prehistory (IT-1223-19), University of the Basque Country (UPV-EHU), Vitoria-Gasteiz, Spain. ⁵²Pennsylvania State University, University Park, PA, USA. ⁵³Department of Archaeology, Simon Fraser University, Burnaby, BC, Canada. ⁵⁴School of Philosophical and Historical Inquiry, Faculty of Arts and Social Sciences, University of Sydney, Sydney, NSW, Australia. ⁵⁵Texas A&M University, College Station, TX, USA. ⁵⁶Department of Evolutionary Anthropology, University of Vienna, Vienna, Austria.

*Corresponding author. Email: anders.bergstrom@crick.ac.uk (A.B.); laurent.frantz@gmail.com (L.F.); ron.pinhasi@univie.ac.at (R.P.); greger.larson@arch.ox.ac.uk (G.L.); pontus.skoglund@crick.ac.uk (P.S.)

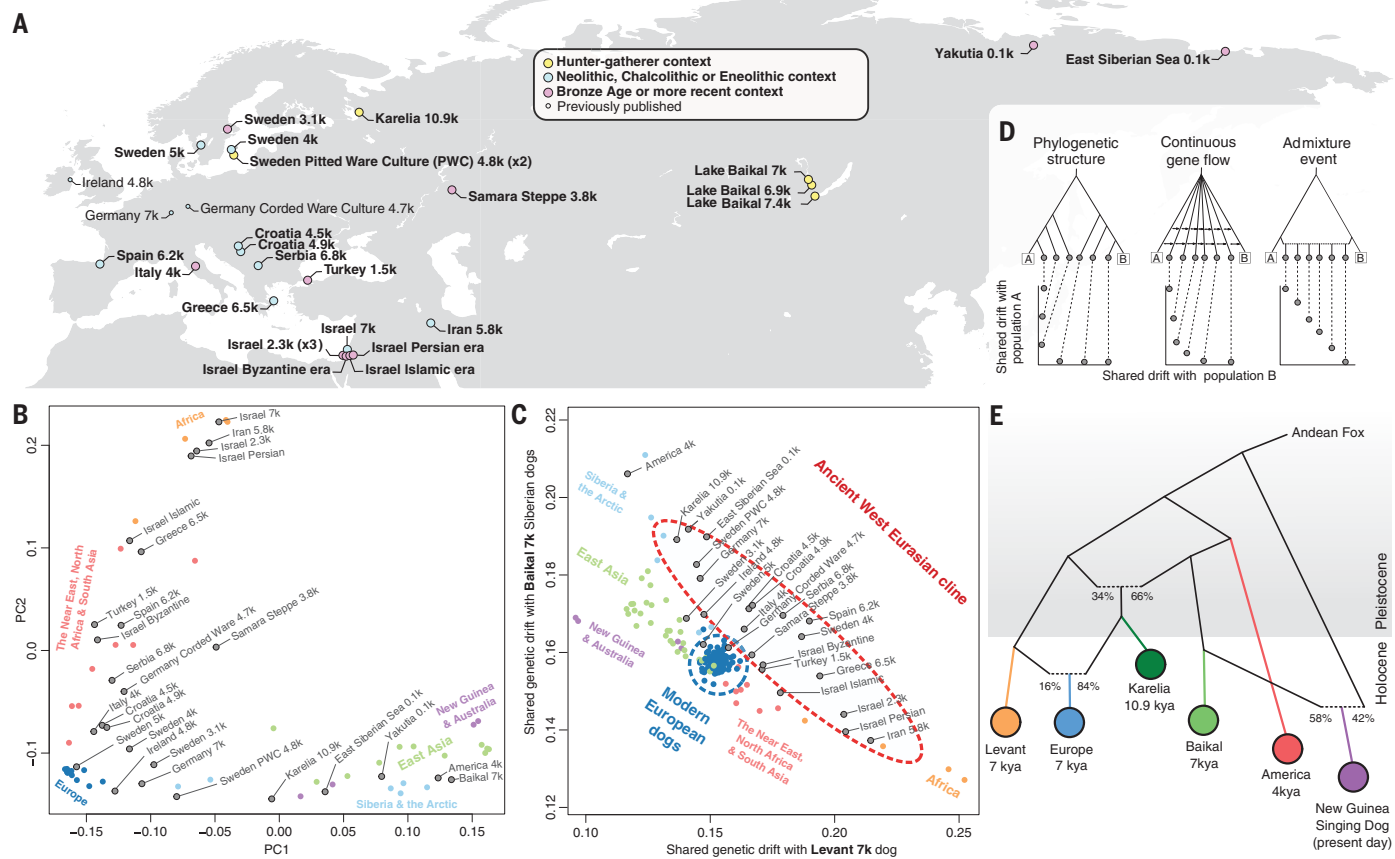


Fig. 1. Genomic structure of dogs dates to the Pleistocene. (A) Sampling locations of ancient dogs. k, 1000 years. (B) PCA results for all possible f_4 -statistics among ancient dogs (gray) and a selection of worldwide modern dogs. (C) Outgroup f_3 -statistics reveal a cline of Levant-related versus Baikal-related (horizontal and vertical axes, respectively) ancestry across ancient West Eurasian dogs, but not

among modern European dogs. (D) Coalescent simulations demonstrating that a diagonal f_3 cline as in (C) is consistent with an admixture event, but less so with continuous gene flow and not with phylogenetic structure alone. (E) An admixture graph that fits all f_4 -statistics between major dog lineages. The European dog was grafted onto the graph identified through exhaustive testing. kya, 1000 years ago.

model fits the data, and features the Mesolithic Karelian dog (10.9 ka ago) as having received part of its ancestry from a lineage related to eastern dogs and part from the Levantine lineage (Fig. 1E) [two highly similar models nearly fit (fig. S4)]. The model can be extended to feature the earliest Neolithic European dog (7 ka ago) (14) as a mixture of the Karelian and the Levantine branches without loss of fit (fig. S5), supporting the dual ancestry model for European dogs suggested by the ancient ancestry cline (Fig. 1C). The observed phylogenetic structure implies that all five ancestry lineages (Neolithic Levant, Mesolithic Karelia, Mesolithic Baikal, ancient America, and New Guinea singing dog) must have existed by 10.9 ka ago (the radiocarbon date of the Karelian dog) and thus most likely existed prior to the transition from the Pleistocene to the Holocene epoch ~11.6 ka ago.

No detectable evidence for multiple dog origins or extensive gene flow from wild canids

Studies have suggested that wolf populations in Europe (3, 11), the Middle East (12), Central

Asia (10), Siberia (31), and East Asia (2, 8), or more than one of these (9), contributed to early dog diversity. One study, however, demonstrated that modern wolves and dogs are reciprocally monophyletic and suggested bidirectional gene flow (5). We corroborated that gene flow must have occurred by identifying widespread asymmetries between dogs in their affinity to wolves (Fig. 2, A and B, and fig. S7). However, the gene flow was likely largely unidirectional from dogs into wolves, as we also identified some gray wolves that are symmetrically related to all modern and ancient dogs (Fig. 2C). Past gene flow from wolves into specific dog populations would have manifested as an affinity to any member of the modern gray wolf lineage in these tests, so our results suggest that persistent gene flow into dogs has been so limited as to be undetectable at the current resolution of the data. Furthermore, this result is consistent with a scenario in which all dogs derive from a single ancient, now-extinct wolf population, or possibly multiple closely related wolf populations. Although it is still possible that other, thus-far-unsampled

ancient wolf populations were independently involved in early domestication (3, 9, 31), our data indicate that they did not contribute substantially to later dogs.

In contrast to the lack of wolf admixture into dogs, we identified dog admixture into almost all analyzed present-day wolves (Fig. 2B), with the strongest signals typically coming from dogs into geographically proximate wolf populations in Europe, the Near East, and East Asia (fig. S7). We also replicated affinities between ancient American dogs and coyotes (21) and between African dogs and African golden wolves (32), although the direction of gene flow in both cases is unclear and the small magnitude is unlikely to impact most analyses of dog relationships (table S5). We did not find genome-wide evidence for gene flow from Tibetan wolves into Tibetan dogs, despite evidence for wolf ancestry locally around the *EPAS1* gene, which is associated with adaptation to altitude (33, 34). Dogs thus do not show similar evidence of wild introgression as has been found in pigs, goats, horses, sheep, and cattle (35–40).

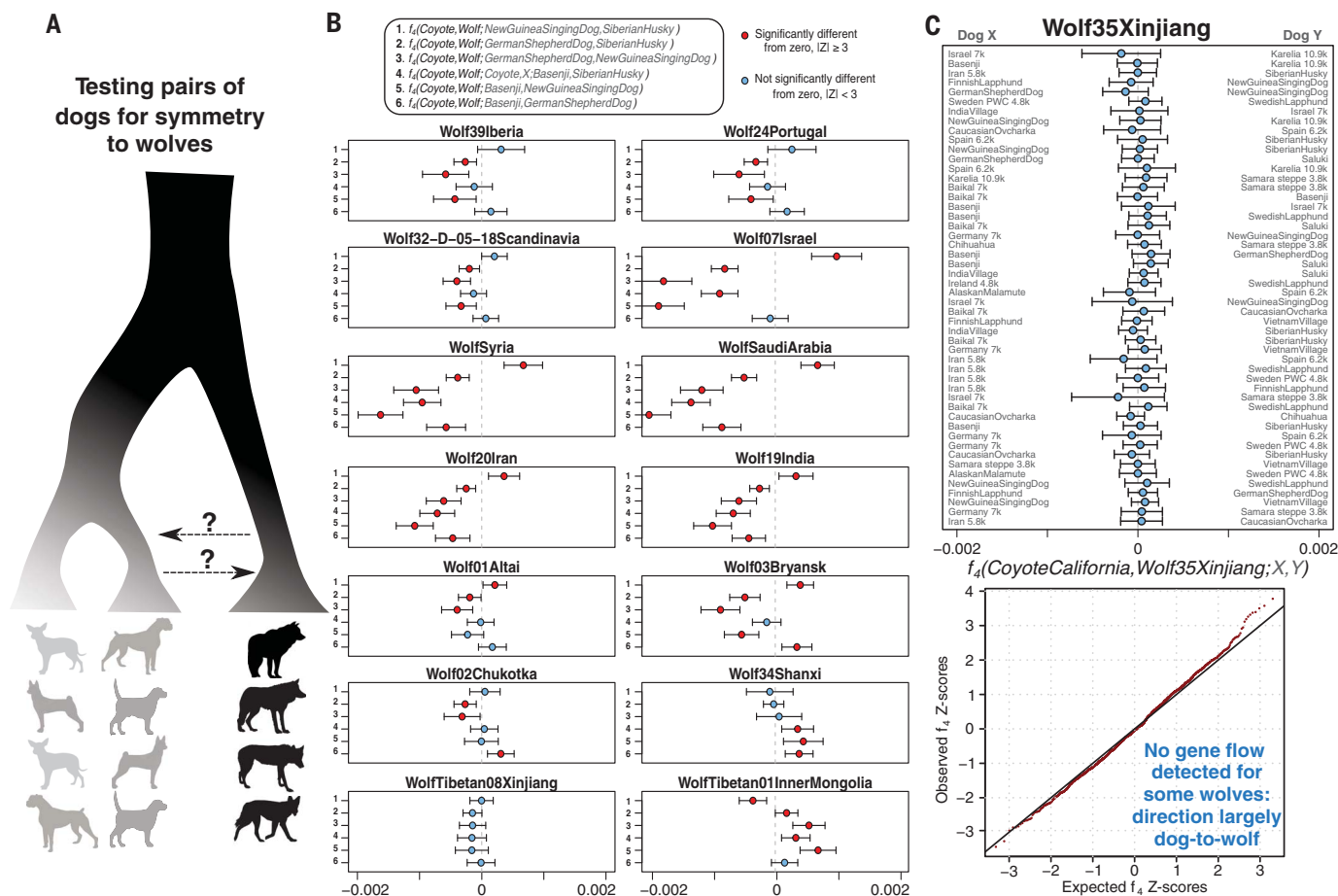


Fig. 2. All detectable gene flow is consistent with being unidirectional from dogs into wolf populations. (A) Illustration of asymmetry tests (f_4 -statistics) comparing 35 Eurasian gray wolves to all pairs of 66 ancient and modern dogs. (B) Selected results using coyotes as outgroup. (C) A wolf from Xinjiang, western China, is not closer to some dog populations

Assessing the relationship between dog and human population histories

We next quantitatively compared the population relationships observed in dogs with those of humans. First, using Procrustes rotation to align f_4 PCA results obtained on dog and human genomes matched in time and space (Fig. 3A) (30), we find that the population structures of the two species resemble each other (Procrustes correlation = 0.48, $P = 0.043$). However, there are also several cases where the matched dogs and humans cluster in different parts of the PCA space. The greatest differences (Fig. 3B) are observed for Chalcolithic Iran, in which the human population is different from the Neolithic Levant (41, 42) but the dogs in the two regions are similar. In Neolithic Germany and Ireland, the humans are more shifted toward the Levant (43, 44) but the dogs are shifted toward Northern European hunter-gatherer contexts. In the Bronze Age Steppe and in Corded Ware Germany, the humans are shifted away from the Neolithic Euro-

pean cluster (45, 46) in a manner not seen in dogs.

Second, we evaluated if the admixture graph topologies that best fit the data for one species could also explain population relationships of the other. Although we found no graphs that fit the data perfectly for both species, graphs that fit or nearly fit dogs ranked among the 0.8 to 2.8% top-scoring graphs in the human search, and graphs that fit humans ranked among the 0.007 to 1.2% top-scoring graphs in the dog search (Fig. 3C and fig. S9). However, this analysis did not take into account the different time depth of the two species' population histories: The >40-ka-ago divergence of human East and West Eurasian ancestries (47) is markedly older than the earliest appearance of dog morphology in the fossil record, conservatively dated to 14.5 ka ago (48), although older (3, 31), disputed (49, 50) specimens have been claimed.

Third, we found that the sign (positive or negative) of f_4 -statistics in dogs matched the

than to others, as the test statistics are consistent with being normally distributed around 0 (the quantile–quantile plot includes all 66 dogs). If there had been a substantial gene flow from some wolf population into some dog population, we would expect all wolf individuals to display asymmetric relationships.

sign in humans in 71% of 31,878 tests (null expectation is 50%) across 24 matched dog–human pairs, although this decreased to 58% when restricted to dogs and humans from Europe. We identified specific f_4 -statistics that exemplify both concordance and discrepancy between the species (Fig. 3D). Whereas it is not known what degree of concordance would be expected between the histories of two species on the basis of biogeographical factors alone, the results of these three analyses demonstrate that ancestry relationships in dogs and humans share overall features but are not identical over space and time, and there are several cases where they must have been decoupled.

Recurrent population histories

One notable example of concordance is that both humans and dogs in East Asia are closer to European than to Near Eastern populations, which in both humans (43) and our best-fitting graph (Fig. 1E) is best modeled by European ancestry being a mixture of

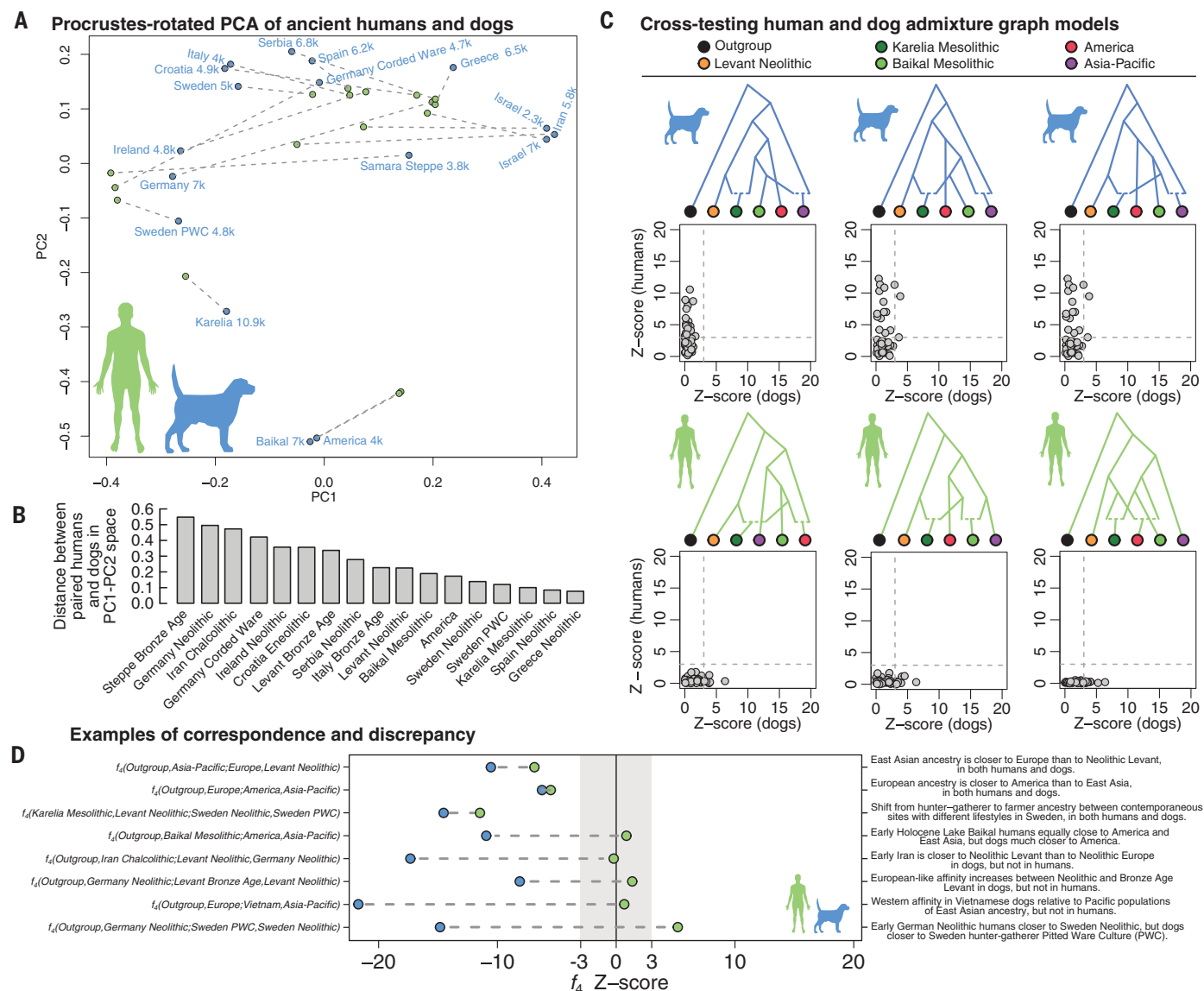


Fig. 3. Quantitative comparisons between dog and human population genomic structures. (A) PCA results for all possible f_4 -statistics on ancient dogs (blue), overlaid through Procrustes transformation by the corresponding analysis performed on ancient humans matched in time, space, and cultural context to the dogs (green). Dashed lines connect each matched pair. **(B)** Euclidean residuals

between the Procrustes-rotated human and dog coordinates. **(C)** The three admixture graphs that fit for one species and provide the smallest error for the other. Scatter plots show absolute Z-scores for the difference between observed and predicted f_4 -statistics. **(D)** Examples of f_4 -statistics that reveal similarities and differences between humans and dogs (far right text).

ancestry related to the Near East and East Asia. However, the divergence of Near Eastern “Basal Eurasian” ancestry in humans was likely >45 ka ago (43), suggesting that dog population dynamics may have mimicked earlier processes in humans. A second example is that all European dogs have a stronger affinity toward American and Siberian dogs than they have to New Guinea singing dogs, which likely represent a type of unadmixed East Asian dog ancestry, mirroring a circum-polar affinity between humans in Europe and the Americas (Fig. 3D) (51). Human groups at Lake Baikal 24 to 18 ka ago had western Eurasian affinities and contributed to Native

American ancestry (51) but were largely replaced by the Holocene (52). Although the dogs at Lake Baikal dated to 7 ka ago constitute a similar link between the Americas and Europe (Fig. 1, C and E), this link occurred >10 ka later (Fig. 3D). Thus, shared circumpolar ancestry through northern Eurasia is an important feature of both human and dog population structures, though this likely did not result from the same migration episodes.

Neolithic expansion into Europe

Ancient human genomes have revealed a major ancestry transformation associated with

the expansion of Neolithic agriculturalists from the Near East into Europe (43, 45, 53), and a study of ancient dog mitochondria suggested they were accompanied by dogs (27). We hypothesized that the genomic ancestry cline we observe across ancient European dogs (Fig. 1C) could be, at least in part, due to admixture between dogs associated with Mesolithic hunter-gatherers and incoming Neolithic farmers. Three observations support this: First, the hypothesized hunter-gatherer end of the cline is occupied by the 10.9-ka-old Mesolithic Karelian dog and dogs from a 4.8-ka-old hunter-gatherer Pitted Ware Culture site in Sweden. Second, relative to the Swedish

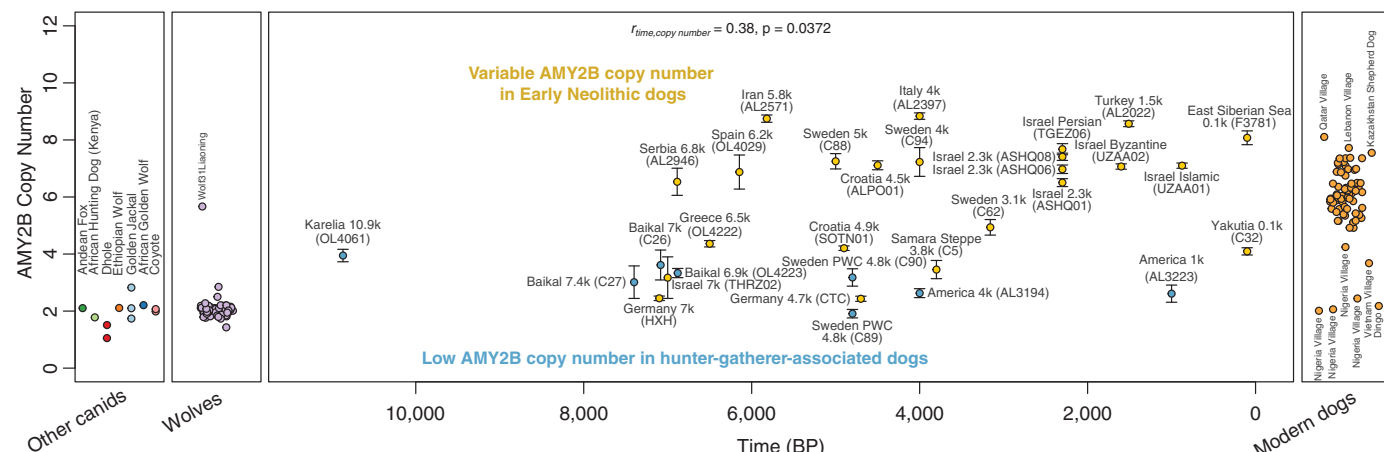


Fig. 4. Expansion of copy number in the *AMY2B* pancreatic amylase gene largely occurred after the transition to agriculture. Ancient dogs are plotted against their age, with blue color indicating dogs from likely hunter–gatherer human contexts. Bars denote 95% binomial confidence intervals around the ratio of the number of reads mapping to the copy number variable region to those mapping to control regions throughout the genome.

hunter–gatherer dogs, a contemporaneous dog from a Swedish Neolithic agricultural context is shifted toward the Levantine end of the cline, mirroring humans at the same sites (41, 53, 54) (Fig. 3, A and D, and fig. S10D). Third, Neolithic Levantine affinity increases toward the south ($P = 0.0196$, linear regression), consistent with a range expansion alongside Neolithic human groups. Whereas dogs clearly associated with Mesolithic continental “Western hunter–gatherer” (43) human groups have yet to be identified, our results suggest that such dogs would have strong affinity toward the Siberian end of the European cline. Overall, these results indicate that the Neolithic expansion of farmers into Europe was also associated with an ancestry transformation for dogs.

Increased copy number of the *AMY2B* gene, which is involved in starch digestion, has been linked to dietary adaptations of dogs during the agricultural transition (6, 55, 56). The paralogous *AMY1* gene has been under adaptive evolution in humans (57), though this does not seem clearly linked to agriculture (58). We observe low copy numbers in dogs from human hunter–gatherer contexts (Fig. 4), although the Mesolithic Karelian dog may already have possessed an elevated number relative to wolves. Several Neolithic dogs have as many copies as present-day dogs, as early as in 5.8-ka-old Iranian and 6.2-ka-old Spanish dogs, but others display low numbers (14, 56), e.g., the 7-ka-old Levantine individual. These results suggest that selection for increased *AMY2B* copy number did not take place during the early stages of domestication, and in contrast to humans (58) it was not advanced in Mesolithic hunter–gatherer contexts but was variable in early agricultural populations and did not become widespread until several thousand years after the first appearance of starch-rich agricultural lifestyles.

Africa and the Near East

The clustering of modern African dogs with ancient dogs from the Levant and Iran, especially the oldest individual, dating to 7 ka ago, suggests a Near Eastern origin (Fig. 1, B and C, and fig. S2). Western (Anatolia and the Levant) and eastern (Zagros mountains of Iran) human groups in the Fertile Crescent were highly genetically differentiated (41), and the western groups were the primary source of gene flow into Europe and Africa (41, 59) during the Neolithic. A source of African dog ancestry from the Levant (7 ka ago) is a better fit than Iran (5.8 ka ago) (Fig. 5A), mirroring the human history, as well as that of cattle (40). In contrast, we are unable to distinguish whether the Levant or Iran is the better source for Neolithic dog ancestry in Europe. Our results suggest a single origin of sub-Saharan African dogs from a Levant-related source (Fig. 5B), with limited gene flow from outside the continent until the past few hundred years.

In contrast to Africa, the 7-ka-old Neolithic Levantine population does not appear to have contributed much, if any, ancestry to present-day dogs in the Near East. Instead, 2.3-ka-old dogs in the Levant can be modeled as having 81% Iran-related and 19% Neolithic Europe-related ancestry (data file S1). By this time in the Levant, there was also human gene flow from Iran (41) and transient gene flow from Europe (60). However, our results suggest a more complete replacement of dog ancestry in the Levant by 2.3 ka ago (Fig. 5B). Later, modern Near Eastern dogs are best modeled as mixtures of the 2.3-ka-old Levantine and modern European sources (data file S1).

Steppe pastoralist expansions

Expansions of steppe pastoralists associated with the Yamnaya and Corded Ware cultures into Late Neolithic and Bronze Age Europe transformed the ancestry of human popula-

tions (43, 45, 46). To test whether dog ancestry was similarly affected, we analyzed a 3.8-ka-old dog from the eastern European steppe associated with the Bronze Age Srubnaya culture. Although its ancestry resembles that of western European dogs (Fig. 1C and fig. S10), it is an outlier in the center of PC1–PC2 space (Fig. 1B). A Corded Ware-associated dog (4.7 ka ago) from Germany, hypothesized to have steppe ancestry (14), can be modeled as deriving 51% of its ancestry from a source related to the Srubnaya steppe dog and the rest from a Neolithic European source (data file S1) (30). We obtain similar results for a Bronze Age Swedish dog (45%; 3.1 ka ago), but not a Bronze Age Italian dog (4 ka ago).

Despite this potential link between the steppe and the Corded Ware dog, most later European dogs display no particular affinity to the Srubnaya dog. Modern European dogs instead cluster with Neolithic European dogs (Fig. 1B) and do not mirror the lasting ancestry shift seen in humans after the pastoralist expansion (Fig. 3A). Earlier and additional steppe dog genomes are needed to better understand this process, but the relative continuity between Neolithic and present-day individuals suggests that the arrival of steppe pastoralists did not result in persistent large-scale shifts in the ancestry of European dogs.

Although steppe pastoralists also expanded east, they do not appear to have contributed much ancestry to present-day people in East Asia (46, 52). Many modern Chinese dogs display unambiguous evidence [negative f_3 tests (30)] of being the product of admixture between a population related to the New Guinea singing dog (and the Australian dingo) and a West Eurasian-related population (table S6). A recent study also found a mitochondrial turnover in Chinese dogs in the last few thousand years (61). The best-fitting models involve not only ancestry from modern European breeds

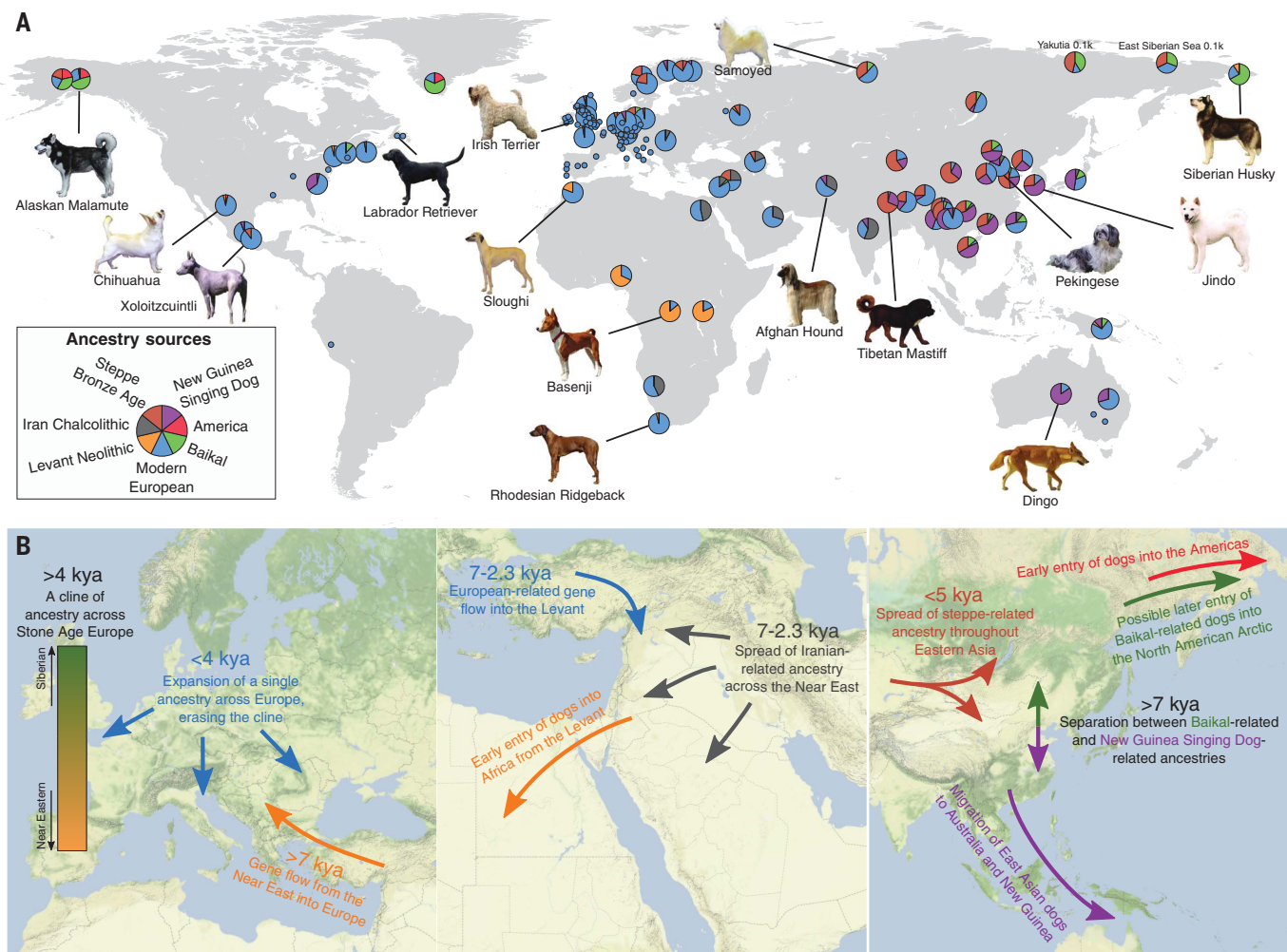


Fig. 5. Ancestry of global dogs today. (A) For each present-day population, the ancestry proportions estimated by the best-fitting qpAdm model, restricted to models containing up to four of seven selected sources, are displayed. Populations for which a single component accounts for $\geq 98\%$ of the ancestry

are collapsed to smaller circles. Dog pictures were obtained from Wikimedia under the CC BY-SA 3.0 license (<https://commons.wikimedia.org/wiki/Special:ListFiles/Desaix83>). (B) Illustrations of inferred population histories in three regions of the world.

but also substantial contributions from the 3.8-ka-old Srubnaya steppe dog (Fig. 5A and data file S1). Some populations, especially those in Siberia, additionally require a fourth source related to the 7-ka-old Lake Baikal dogs, but no or minimal New Guinea singing dog-related ancestry. Our results thus raise the possibility that the eastward migrations of steppe pastoralists had a more substantial impact on the ancestry of dogs than humans in East Asia (Fig. 5B).

Later homogenization of dog ancestry in Europe

The extensive range of ancestry diversity among early European dogs is not preserved today, as modern European dogs are all symmetrically related to the ancient dogs in our dataset (Fig. 1C, fig. S13, and data file S1) (30). This suggests little to no contribution of most local Mesolithic and Neolithic populations to present-day diversity in Europe. Instead,

we found that a single dog from a Neolithic megalithic context dated to 5 ka ago at the Frälsegården site in southwestern Sweden can be modeled as a single-source proxy for 90 to 100% of the ancestry of most modern European dogs, to the exclusion of all other ancient dogs (fig. S13 and data file S1). This implies that a population with ancestry similar to this individual, but not necessarily originating in Scandinavia, replaced other populations and erased the continent-wide genetic cline (Fig. 5B). This ancestry was in the middle of the cline (Fig. 1C), and so present-day European dogs can be modeled as having about-equal proportions of Karelian- and Levantine-related ancestries [54 and 46%, respectively, for German shepherd on the basis of the admixture graph (Fig. 1E)].

The Frälsegården dog is also favored as a partial ancestry source for a 4-ka-old Bronze Age dog from Italy, a 1.5-ka-old dog from Turkey and Byzantine and Medieval, but not earlier

dogs in the Levant (data file S1), providing some constraints on the timing of this ancestry expansion. However, the circumstances that initiated or facilitated the homogenization of dog ancestry in Europe from a narrow subset of that present in the European Neolithic, including the phenomenal phenotypic diversity and genetic differentiation of modern breeds (12, 19, 20) (Fig. 1C), remain unknown.

More recently, this modern European ancestry has dispersed globally and today is a major component of most dog populations worldwide (Fig. 5A). Our ancestry models, however, reveal that some precolonial ancestry does survive in breeds such as the Mexican chihuahua (~4%) and Xoloitzcuintli (~3%) and the South African Rhodesian ridgeback (~4%) (data file S1).

Discussion

The diversification of at least five dog ancestry lineages by the onset of the Holocene was

followed by a dynamic population history that in many ways tracked that of humans, likely reflecting how dogs migrated alongside human groups. However, in several instances, these histories do not align, suggesting that humans also dispersed without dogs, dogs moved between human groups, or that dogs were cultural and/or economic trade commodities.

Certain aspects of genetic relationships between dog populations, such as an east–west Eurasian differentiation, circumpolar connections, and possible basal lineages in the Near East, resemble features of human population history that were established before the earliest estimated dates of dog domestication. This superficial mirroring between the species may therefore instead point to recurrent population dynamics due to biogeographic or anthropological factors that remain to be understood. A key question is how dogs spread across Eurasia and the Americas by the Holocene, since no major human population movements have been identified after the initial out-of-Africa expansion that could have driven this global dispersal.

We find that the modern and ancient genomic data are consistent with a single origin for dogs, though a scenario involving multiple closely related wolf populations remains possible. However, in our view, the geographical origin of dogs remains unknown. Previously suggested points of origin based upon present-day patterns of genomic diversity (2, 8, 10) or affinities to modern wolf populations (12) are sensitive to the obscuring effects of more recent population dynamics and gene flow. Ultimately, integrating DNA from dogs and wolves even older than those analyzed here with archaeology, anthropology, ethology, and other disciplines is needed to determine where and in which environmental and cultural context the first dogs originated.

REFERENCES AND NOTES

- C. Vilà *et al.*, *Science* **276**, 1687–1689 (1997).
- P. Savolainen, Y. P. Zhang, J. Luo, J. Lundberg, T. Leitner, *Science* **298**, 1610–1613 (2002).
- M. Germonpré *et al.*, *J. Archaeol. Sci.* **36**, 473–490 (2009).
- P. Skoglund, E. Ersmark, E. Palkopoulou, L. Dalén, *Curr. Biol.* **25**, 1515–1519 (2015).
- A. H. Freedman *et al.*, *PLOS Genet.* **10**, e1004016 (2014).

- E. Axelsson *et al.*, *Nature* **495**, 360–364 (2013).
- P. Skoglund, A. Götherström, M. Jakobsson, *Mol. Biol. Evol.* **28**, 1505–1517 (2011).
- G.-D. Wang *et al.*, *Cell Res.* **26**, 21–33 (2016).
- L. A. F. Frantz *et al.*, *Science* **352**, 1228–1231 (2016).
- L. M. Shannon *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13639–13644 (2015).
- O. Thalmann *et al.*, *Science* **342**, 871–874 (2013).
- B. M. Vonholdt *et al.*, *Nature* **464**, 898–902 (2010).
- J.-F. Pang *et al.*, *Mol. Biol. Evol.* **26**, 2849–2864 (2009).
- L. R. Botigué *et al.*, *Nat. Commun.* **8**, 16082 (2017).
- D. F. Morey, *Am. Sci.* **82**, 336–347 (1994).
- J. Clutton-Brock, *Science* **197**, 1340–1342 (1977).
- S. J. M. Davis, F. R. Valla, *Nature* **276**, 608–610 (1978).
- M. Sablin, G. Khlopachev, *Curr. Anthropol.* **43**, 795–799 (2002).
- H. G. Parker *et al.*, *Science* **304**, 1160–1164 (2004).
- H. G. Parker *et al.*, *Cell Rep.* **19**, 697–708 (2017).
- M. Ni Leathlobhair *et al.*, *Science* **361**, 81–85 (2018).
- B. van Asch *et al.*, *Proc. Biol. Sci.* **280**, 1142 (2013).
- J. A. Leonard *et al.*, *Science* **298**, 1613–1616 (2002).
- S. Castroviejo-Fisher, P. Skoglund, R. Valadez, C. Vilà, J. A. Leonard, *BMC Evol. Biol.* **11**, 73 (2011).
- K. Greig *et al.*, *Sci. Rep.* **8**, 9130 (2018).
- P. Savolainen, T. Leitner, A. N. Wilton, E. Matisoo-Smith, J. Lundberg, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 12387–12390 (2004).
- M. Ollivier *et al.*, *Biol. Lett.* **14**, 20180286 (2018).
- C. Armeen *et al.*, *Proc. R. Soc. B Biol. Sci.* **286**, 1929 (2019).
- H. Malmström *et al.*, *BMC Evol. Biol.* **8**, 71 (2008).
- Materials and methods are available as supplementary materials.
- N. D. Ovodov *et al.*, *PLOS ONE* **6**, e22821 (2011).
- Y.-H. Liu *et al.*, *Mol. Biol. Evol.* **35**, 287–298 (2018).
- B. Miao, Z. Wang, Y. Li, *Mol. Biol. Evol.* **34**, 734–743 (2017).
- B. vonHoldt, Z. Fan, D. Ortega-Del Vecchyo, R. K. Wayne, *PeerJ* **5**, e3522 (2017).
- L. A. F. Frantz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 17231–17238 (2019).
- K. G. Daly *et al.*, *Science* **361**, 85–88 (2018).
- A. Fages *et al.*, *Cell* **177**, 1419–1435.e31 (2019).
- M. Barbato *et al.*, *Sci. Rep.* **7**, 7623 (2017).
- S. D. E. Park *et al.*, *Genome Biol.* **16**, 234 (2015).
- M. P. Verdugo *et al.*, *Science* **365**, 173–176 (2019).
- I. Lazaridis *et al.*, *Nature* **536**, 419–424 (2016).
- F. Broushaki *et al.*, *Science* **353**, 499–503 (2016).
- I. Lazaridis *et al.*, *Nature* **513**, 409–413 (2014).
- L. M. Cassidy *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 368–373 (2016).
- W. Haak *et al.*, *Nature* **522**, 207–211 (2015).
- M. E. Allentoft *et al.*, *Nature* **522**, 167–172 (2015).
- Q. Fu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2223–2227 (2013).
- L. Janssens *et al.*, *J. Archaeol. Sci.* **92**, 126–138 (2018).
- A. Perri, *J. Archaeol. Sci.* **68**, 1–4 (2016).
- D. F. Morey, *J. Archaeol. Sci.* **52**, 300–307 (2014).
- M. Raghavan *et al.*, *Nature* **505**, 87–91 (2014).
- P. B. Damgaard *et al.*, *Nature* **557**, 369–374 (2018).
- P. Skoglund *et al.*, *Science* **336**, 466–469 (2012).
- P. Skoglund *et al.*, *Science* **344**, 747–750 (2014).
- M. Arendt, K. M. Cairns, J. W. O. Ballard, P. Savolainen, E. Axelsson, *Heredity* **117**, 301–306 (2016).
- M. Ollivier *et al.*, *R. Soc. Open Sci.* **3**, 160449 (2016).

- G. H. Perry *et al.*, *Nat. Genet.* **39**, 1256–1260 (2007).
- S. Mathieson, I. Mathieson, *Mol. Biol. Evol.* **35**, 2957–2970 (2018).
- P. Skoglund *et al.*, *Cell* **171**, 59–71.e21 (2017).
- M. Feldman *et al.*, *Sci. Adv.* **5**, eaax0061 (2019).
- M. Zhang *et al.*, *Mol. Biol. Evol.* **37**, 1462–1469 (2020).

ACKNOWLEDGMENTS

We thank S. Charlton, I. Lazaridis, A. Manin, and I. Mathieson for comments on the manuscript, G.-D. Wang and C. Marsden for help with data access, and GORDAILUA (the Gipuzkoa Centre for Heritage Collections), S. San José, C. Olaetxea, M. Urteaga, A. Sampson, A. R. Sardari Zarchi, and M. Abdollahi (ICHTO, Iran) for facilitating sample access. **Funding:** Ancient genome sequencing was supported by SciLifeLab National Projects and the Erik Philip Sörensen Foundation (to P.S.). A.B., T.D., and P.S. were supported by the Francis Crick Institute core funding (FC001595) from Cancer Research UK, the UK Medical Research Council, and the Wellcome Trust. P.S. was also supported by the European Research Council (grant no. 852558), a Wellcome Trust Investigator award (217223/Z/19/Z) and the Vallee Foundation. R.J.L. was supported by the Social Sciences and Humanities Research Council of Canada (#SSHR IG 435-2014-0075). Y.K. was supported by State Assignment of the Sobolev Institute of Geology and Mineralogy. M.S. was supported by ZIN RAS (state assignment no. AAA-A19-119032590102-7). A.T.L. was supported by the Smithsonian's Peter Buck Postdoctoral Fellowship. Archaeological work in Serbia was supported by AHRC grant AH/J001406/1. Computations were supported by SNIC-UPPMAX (b2016004) and the UOXF ARC facility. L.F. was supported by the Wellcome Trust (grant 210119/Z/18/Z) and by Wolfson College (University of Oxford). G.L. was supported by the ERC (grant ERC-2013-StG-337574-UNDEAD). G.L. and K.D. were supported by the Natural Environmental Research Council (grants NE/K005243/1 and NE/K003259/1). Dating was supported by the NERC Radiocarbon Facility (NF/2016/2/4). **Author contributions:** G.L. and P.S. initiated the study. J.S., K.-G.S., D.A., E.A., S.A., G.B.-O., V.I.B., J.B., D.B., S.F., I.F., D.F., M.G., L.K.H., L.J., J.K.-C., Y.K., R.J.L., D.L.D., M.M., M.N., V.O., D.O., M.P., M.R., D.R., B.R., M.S., I.S., A.T., K.T., I.U., A.V., P.W., A.G., and L.D. contributed material and archaeological information. R.S., E.E., O.L., L.G.-F., J.H., A.J., H.R., and A.L. did ancient DNA molecular work, supervised by A.G., L.D., R.P., G.L., and P.S. A.B., L.F., A.C., T.D., E.K.I.-P., and P.S. processed the genome sequence data, supervised by L.F. and P.S. A.B. did population genomic analyses, supervised by P.S. A.T.L. did mitochondrial DNA analyses, supervised by G.L. A.B., L.F., G.L., and P.S. wrote the paper with input from R.P., K.D., and all other authors. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** The generated DNA sequencing data are available in the European Nucleotide Archive (ENA) under study accession PRJEB38079.

SUPPLEMENTARY MATERIAL

science.sciencemag.org/content/370/6516/557/suppl/DC1
Materials and Methods
Figs. S1 to S13
Tables S1 to S6
References (62–150)
Data File S1
MDAR Reproducibility Checklist

[View/request a protocol for this paper from Bio-protocol.](#)

22 January 2020; accepted 10 September 2020
10.1126/science.aba9572

CORONAVIRUS

The emergence of SARS-CoV-2 in Europe and North America

Michael Worobey^{1*}, Jonathan Pekar^{2,3}, Brendan B. Larsen¹, Martha I. Nelson⁴, Verity Hill⁵, Jeffrey B. Joy^{6,7,8}, Andrew Rambaut⁵, Marc A. Suchard^{9,10,11*}, Joel O. Wertheim^{12*}, Philippe Lemey^{13*}

Accurate understanding of the global spread of emerging viruses is critical for public health responses and for anticipating and preventing future outbreaks. Here we elucidate when, where, and how the earliest sustained severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) transmission networks became established in Europe and North America. Our results suggest that rapid early interventions successfully prevented early introductions of the virus from taking hold in Germany and the United States. Other, later introductions of the virus from China to both Italy and Washington state, United States, founded the earliest sustained European and North America transmission networks. Our analyses demonstrate the effectiveness of public health measures in preventing onward transmission and show that intensive testing and contact tracing could have prevented SARS-CoV-2 outbreaks from becoming established in these regions.

In late 2019, the emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which causes coronavirus disease 2019 (COVID-19), ignited a pandemic that has been associated with more than 500,000 deaths globally as of July 2020. As the original outbreak in Hubei province, China, spilled into other countries, containment strategies focused on travel restrictions, isolation, and contact tracing. Given the virus's exponential growth rate, delaying the onset of community transmission by even a few weeks likely bought government officials valuable time to establish diagnostic testing capacity and implement social distancing plans.

Viral genetic sequence data can provide critical information about whether viruses separated by time and space are likely to be

epidemiologically linked. Genomic data have suggested differences in the timing, spatial origins, and transmission dynamics of early SARS-CoV-2 outbreaks in multiple North American locations, including Washington state (1, 2); the East Coast of the United States (3, 4); California (5); and British Columbia (BC), Canada (5, 6). The first confirmed U.S. case was associated with a virus strain ("WA1") isolated in Washington state from a traveler who returned from Wuhan, China, on 15 January 2020 (7). No onward transmission was detected after extensive follow-up in what appeared to be successful containment of the country's first known incursion of the virus (8). However, subsequent identification of viruses that were genetically similar to WA1—first in Washington, then in Connecticut (3), California (5), BC (6), and elsewhere—raised the possibility that WA1 had actually established chains of cryptic transmission that started on 15 January and went undetected for several weeks (1, 2). If true, this introduction would predate early SARS-CoV-2 community transmission chains documented elsewhere on the continent (3–5) and would establish the Seattle area as the epicenter of the North American epidemic. Hence, it is necessary to resolve this question to determine where the virus first initiated substantial community outbreaks and whether the earliest coast-to-coast spread of the virus within the United States (3) was from west to east or from east to west.

In Europe, the first diagnosed case occurred in an employee of an automobile supplier who visited the company's headquarters in Bavaria, Germany, from Shanghai, China, on 20 January 2020 (9). She had been infected with SARS-CoV-2 in Shanghai (after her parents had visited from Wuhan) (10) and transmitted the virus to a German man who tested positive on 27 January

(11) and whose viral genome ("BavPat1") was sampled on 28 January (10). In total, this outbreak infected 16 employees but was apparently contained through rapid testing and isolation (9). Italy's first major outbreak in Lombardy, which was apparent by ~20 February 2020, was associated with viruses closely related to BavPat1 but of a separate lineage (designated "B.1"), which differs from BavPat (a lineage "B" virus) by just 1 nucleotide in the nearly 30,000-nucleotide genome. A narrative took hold that the virus from Germany had not been contained but had been transmitting undetected for weeks and had been carried to Italy by an infected German (9, 12). In addition to igniting a severe outbreak in Italy, this B.1 lineage subsequently spread widely across Europe and beyond, initiating outbreaks in many countries, including the intense U.S. outbreak in New York City (NYC) (13, 14). Greater clarity about the effectiveness of Germany's early contact tracing efforts has implications for the feasibility of controlling the virus through non-pharmaceutical interventions.

There are a number of limitations in phylogenetic and spatial inferences drawn from SARS-CoV-2 genomic data. SARS-CoV-2 has a relatively long (~29 kb) positive-sense single-stranded RNA genome that evolves at a rate of $<1 \times 10^{-3}$ substitutions per site per year, amounting to ~2 substitutions per genome per month. This rate is slower than that of most RNA viruses, owing to the "proofreading" activity encoded by the nonstructural gene *nsp14* (15). Consequently, the entire global population of SARS-CoV-2 through March 2020 differed by only 0 to 12 nucleotide substitutions from the inferred ancestor of the entire pandemic. Transmission clusters tend to be defined by 1 to 3 nucleotide differences across the entire viral genome. Phylogeographic inferences are further confounded by the relatively low availability of genomic sequence data from locations that experienced early outbreaks, including Italy, Iran, and the original epicenter in Hubei. The combination of the relatively slow rate of SARS-CoV-2 evolution, its rapid dissemination within and between locations, and unrepresentative sampling presents risks for serious misinterpretation.

In this study, we investigated fundamental questions about when, where, and how SARS-CoV-2 established itself globally. We developed phylogenetic inferences from multiple sources of information—including airline passenger flow data between potential sources and destinations of viral dispersals early in the pandemic, as well as disease incidence data in Hubei province and other locales that likely affected the probability of infected travelers moving the virus around the globe. By combining a genomic epidemiology approach, which aims to account for the effects of undersampling viral genetic diversity in the epicenter of the

¹Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA. ²Bioinformatics and Systems Biology Graduate Program, University of California San Diego, La Jolla, CA 92093, USA. ³Department of Biomedical Informatics, University of California San Diego, La Jolla, CA 92093, USA. ⁴Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA. ⁵Institute of Evolutionary Biology, University of Edinburgh, King's Buildings, Edinburgh EH9 3FL, UK. ⁶Department of Medicine, University of British Columbia, Vancouver, BC, Canada. ⁷BC Centre for Excellence in HIV/AIDS, Vancouver, BC, Canada. ⁸Bioinformatics Programme, University of British Columbia, Vancouver, BC, Canada. ⁹Department of Biomathematics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095, USA. ¹⁰Department of Biostatistics, Fielding School of Public Health, University of California Los Angeles, Los Angeles, CA 90095, USA. ¹¹Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095, USA. ¹²Department of Medicine, University of California San Diego, La Jolla, CA 92093, USA. ¹³KU Leuven Department of Microbiology, Immunology and Transplantation, Rega Institute, Laboratory of Clinical and Epidemiological Virology, Leuven, Belgium.

*Corresponding author. Email: worobey@arizona.edu (M.W.); msuchard@ucla.edu (M.A.S.); jwertheim@health.ucsd.edu (J.O.W.); philippe.lemey@kuleuven.be (P.L.)

pandemic, with consideration of expected evolutionary patterns for a novel pathogen with low diversity, we resolved key questions about how and when the SARS-CoV-2 pandemic unfolded in Europe and North America.

Emergence of SARS-CoV-2 in the United States

A key turning point in the U.S. outbreak occurred when researchers sequenced the first viral genome recovered from a putative case of community transmission in the United States (“WA2,” sampled in the Seattle area on 24 February 2020), reporting on 29 February that it was similar to WA1, the viral variant from the first-diagnosed COVID-19 patient (1). This finding led to the suggestion that WA1 might have established cryptic transmission in Washington state in mid-January (1). The researchers did, however, acknowledge the possibility of an independent introduction of WA2 separate from that of WA1. This finding fundamentally altered the picture of the SARS-CoV-2 situation in the United States, playing a decisive role in Washington state’s early adoption of intensive social distancing efforts. This, in turn, appeared to explain Washington state’s relative success in controlling the outbreak, compared with that of states that adopted a delayed approach, such as New York.

The availability of hundreds of SARS-CoV-2 genomes sampled in Washington state by mid-March revealed that WA2 belongs to a large, monophyletic clade of “A.1” lineage viruses that accounted for ~85% of cases in Washington state at that point, designated the “Washington state outbreak clade” (2) (hereafter “WA outbreak clade”). To investigate whether the WA outbreak clade was initiated in mid-January by WA1, we used these data to simulate the epidemic under the constraint that it had been established by WA1 and then compared the observed evolutionary patterns with those expected under that scenario. A range of phylogenetic patterns could have been observed in this large sample (e.g., Fig. 1, A to C) but were not (Fig. 1D).

To investigate whether the observed pattern of evolution reported in (1, 2) was consistent with the WA outbreak clade having descended from WA1, we used FAVITES (FrAmework for VIRal Transmission and Evolution Simulation) to simulate outbreaks (16) (fig. S1 and table S1). These simulated outbreaks had a median doubling time of 4.7 days (95% range across simulations: 4.2 to 5.1 days)—including those that originated from so-called “superspreading” events (fig. S2)—and a fixed evolutionary rate of 0.8×10^{-3} substitutions per site per year. A duration of 2 months (61 days) was chosen to reflect the time period between WA1 and the implementation of disease mitigation efforts that would affect the median doubling time.

We examined the phylogenetic structure of maximum likelihood trees inferred from subsampled simulated viral sequences to determine how frequently they matched the observed relationship between WA1 and the WA outbreak clade. Specifically, a simulation tree matching the observed tree must produce a single branch emanating from WA1 that experiences at least two mutations (C17747T and A17858G in the observed tree) before es-

tablishment of a single outbreak clade (Fig. 2A). Alternative patterns include: (i) a virus identical to WA1 (Fig. 2B), (ii) a virus that differs from WA1 by a single mutation (Fig. 2C), (iii) a viral lineage forming a basal polytomy with WA1 and the outbreak clade (Fig. 2D), and (iv) a viral lineage that is a “sibling” of the outbreak clade but experiences fewer than two mutations before divergence (Fig. 2E). The frequency of alternative phylogenetic patterns in

Fig. 1. Schematic showing a hypothetical path that the key mutations in the WA outbreak could have taken in a susceptible population, alongside the inferred phylogeny.

(A) Scenario in which a hypothetical mutation occurs from WA1-like genomes. (B) Hypothetical phylogeny in which A17747 and C17858

from the original WA1 virus are maintained in the population and sampled at the end. (C) Hypothetical scenario in which a virus that differs from WA1 by one mutation (A17858G) is maintained in the population. (D) Observed tree from the WA outbreak.

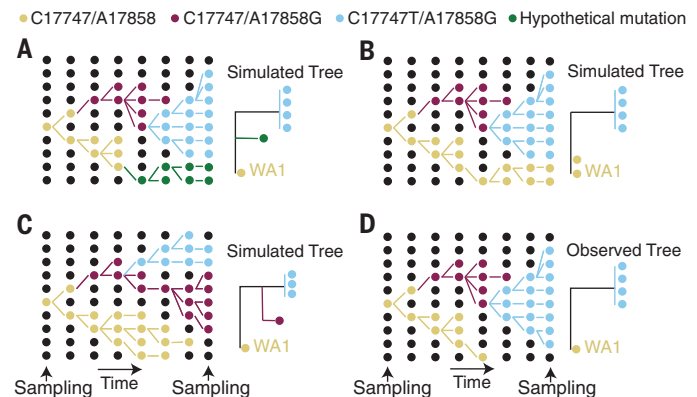
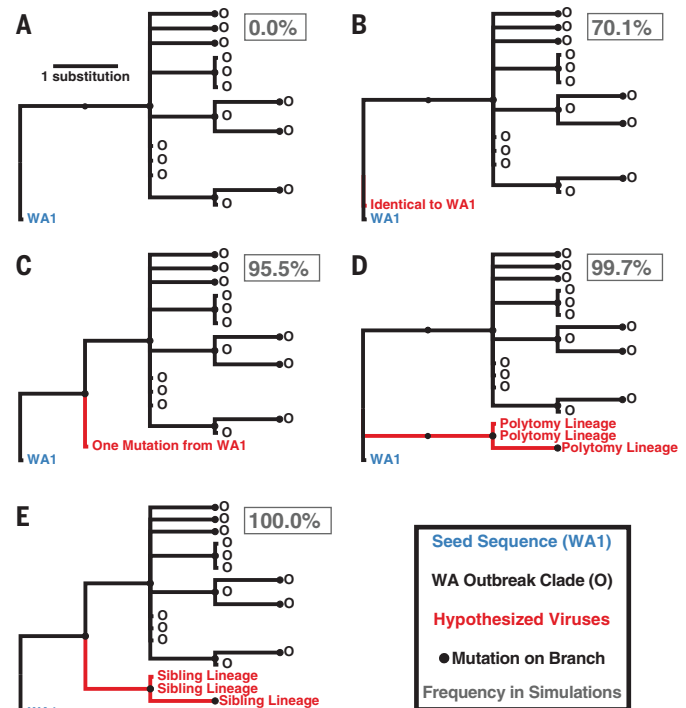


Fig. 2. Potential phylogenetic relationships between WA1 and the WA outbreak clade and their occurrence probabilities. (A) Observed pattern in which the WA1 genome is the direct ancestor of the outbreak clade, separated by at least two mutations. (B) Identical sequence to that of WA1. (C) Sequence that diverges from the WA1 sequence by one mutation. (D) Lineage forming a basal polytomy with WA1 and the outbreak clade. (E) Sibling lineage to the outbreak clade, with fewer than two mutations from WA1 before divergence. The frequency of each relationship across 1000 simulations is reported in the gray box.



the simulated epidemics represents the probability that the true topology (Fig. 2A) could not have occurred if the WA outbreak clade had been initiated by WA1.

In 70.1% of simulations, we observed at least one virus that is genetically identical to WA1, with a median of 12 identical viruses in each simulation (95% range: 0 to 85 identical viruses) (Fig. 2). Not observing a virus identical to WA1 in the real Washington data does not significantly differ from expectation ($P = 0.299$). However, viruses with one mutation from WA1 were observed in 95.5% of simulations, indicating a low probability of failing to detect even a single sequence from Washington within one mutation of WA1 ($P = 0.045$). Lineages forming a basal polytomy with WA1 and the epidemic clade were observed in 99.7% of populations ($P = 0.003$), and 100% of simulations had at least one sibling lineage that diverged before experiencing two mutations and the formation of the outbreak clade ($P < 0.001$). Therefore, even if C1774T and A1785G were linked—a possibility because they are both nonsynonymous mutations located in the nsp13 helicase gene—we would still expect to see descendants of their predecessors in Washington

before 15 March. In summary, when we simulated the Washington outbreak beginning with WA1 on 15 January 2020 and sampled 294 genomes in the first two months of this outbreak, we failed to observe a single simulated epidemic that had the characteristics of the real phylogeny (Fig. 2). These findings were robust to simulations that used a slower epidemic doubling time of 5.6 days (95% range: 5.2 to 5.9 days) or an accelerated substitution rate of 1.6×10^{-3} substitutions per site per year (16) (supplementary text).

Although WA outbreak-related genomes lacking one or the other of the clade-defining substitutions C1774T and A1785G (Fig. 2, C and E) were absent in this initial large sample from Washington state, such genomes have been reported to be very common in nearby BC, Canada (supplementary text). Genomes with the ancestral C1774T state constituted 16 of the first 27 WA outbreak-related genomes sequenced in BC and have been sampled occasionally at much lower frequency in several U.S. states (3). Such a high frequency of these viruses in BC but not in Washington state raises the possibility that BC, rather than Washington state, was the site of introduc-

tion of the founding virus of this key lineage. Another possibility is that these BC genomes are descendants of a separate A.1 lineage introduction from China. The first scenario seems unlikely because of epidemiological evidence that the outbreak was larger in Washington state than in BC in February and March; the second scenario is unlikely because it would necessitate both introduced lineages to independently acquire the C1774T mutation.

We therefore considered a third hypothesis: that these 16 BC viral genomes contain a sequencing error at position 17747 and, in reality, bear the derived C1774T mutation. We reasoned that if this were the case, some of these genomes might share additional derived mutations with C1774T and A1785G genomes sampled in the same location (i.e., they might be identical or highly similar except for a spurious C1774T base) (supplementary text). As shown in Fig. 3, this is indeed the case: Each of the six C1774T genomes from BC that contain one or more derived mutations at positions other than 17747 and 17858 shared one to four of these mutations with other locally sampled genomes. Such a pattern is virtually impossible to explain through homoplasy events. Observing even one such homoplasy in a genome with more than 29,000 bases is rare; the probability of observing more than one is infinitesimally small. Similarly, the hypothesis that the C1774T state in these genomes is due to multiple, independent T1774C reversions is untenable. Occasional C1774T genomes from California, Oregon, Wyoming, Minnesota, Washington state, and elsewhere also share derived mutations with viruses sampled in the same location (Fig. 3, table S2, and supplementary text). Most of these genomes were generated through the amplicon-based ARTIC protocol, and we speculate that mistaken incorporation of a primer sequence containing C17747 (“nCov2019_58_RIGHT”) may be the cause.

When we investigated an exhaustive collection of genomes sampled in Washington state, including those of viruses sampled after 15 March that are related to the WA outbreak clade (supplementary text), we detected a single virus lineage—“WA-S566,” sampled on 29 March 2020—that lacked the derived C1774T and A1785G mutations found in the rest of the WA outbreak clade. The phylogenetic position of this virus matches the pattern in Fig. 2D, though it differs from WA1 at seven additional sites. Hence, the observed pattern in this larger, and later, sample of ~1000 viral genomes reflects the scenario depicted in Fig. 1A rather than that in Fig. 1D. Consequently, we revisited our WA simulations, sampled 1000 genomes instead of the original 294, and looked for instances in which more than two lineages diverged before the formation of the outbreak clade. In 88.8% of the simulations, we observed

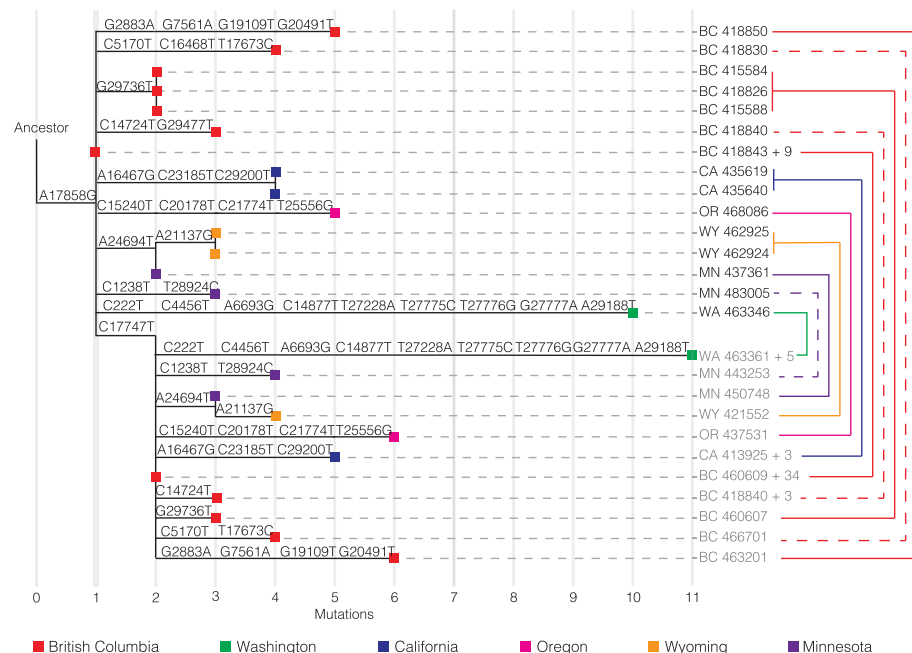
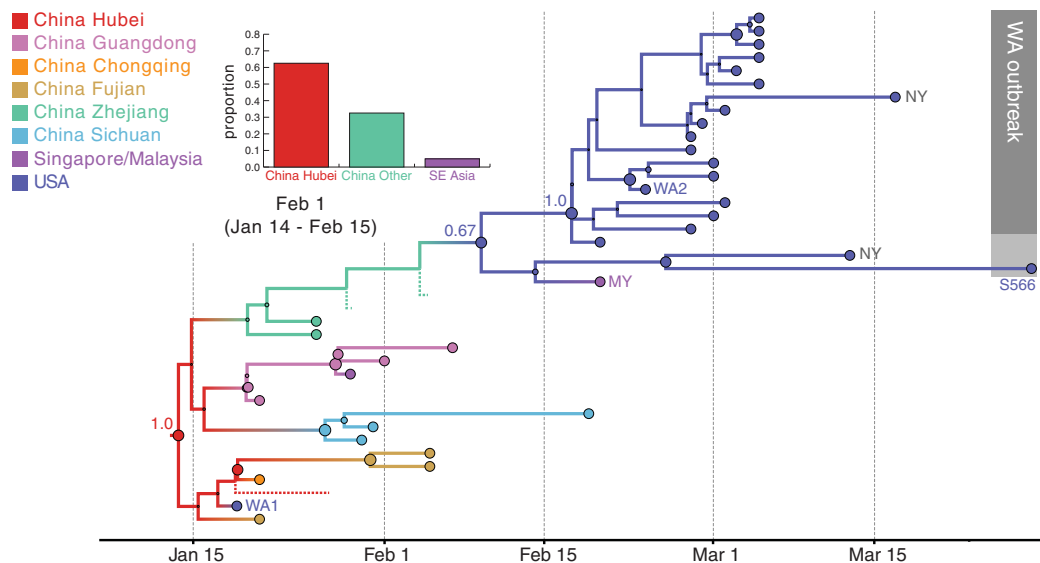


Fig. 3. Phylogeny of representative sequences, showing connections between sequences that share derived mutations despite differences at the key site 17747. Derived mutations from ancestral states (relative to the reference sequence hCoV-19/Wuhan/Hu-1/2019/EPI_ISL_402125) are shown above each branch, with position numbers indicated. Branches are connected to taxon names with horizontal dotted lines. The taxon names include a two-letter state or province code, as well as the GISAID accession number. In cases for which more than one sequence is represented, the total number of additional, identical sequences is indicated after the “+” symbol. Sequences that share derived mutations are connected with colored lines on the right, with colors indicating the locations where the connected sequences were sampled. Some lines on the right are dashed for clarity. Names of sequences that contain the derived nucleotide at site 17747 are shaded in gray.

Fig. 4. Hypothesis of SARS-CoV-2 entry into Washington state. A

subtree of the maximum clade credibility (MCC) tree is shown, depicting the evolutionary relationships inferred between (i) the first identified SARS-CoV-2 case in the United States (WA1); (ii) the clade associated with the Washington state outbreak (including WA2) and related viruses (WA-S566 and a virus from New York); and (iii) closely related viruses that were identified in multiple locations in Asia. Genome sequences sampled at the tips of the phylogeny are represented by circles shaded according to sampling location. Internal node circles, representing posterior clade support values, and branches are shaded similarly by location. Dotted lines represent branches associated with unsampled taxa assigned to Hubei and Zhejiang, China. Posterior location state probabilities are shown for three well-supported key nodes (circle color indicates inferred location state). The inset bar chart summarizes the probability by location for a second introduction giving rise to the WA outbreak clade. The mean date and 95% HPD intervals represent estimated time of introduction from Hubei.



two or fewer basally divergent lineages and, therefore, cannot reject a scenario in which WA1 gave rise to only two lineages that diverge as a basal polytomy ($P = 0.112$). However, in 99.0% of simulations, we observed three or more divergent lineages before two mutations (i.e., lineages that experienced zero or one mutation from WA1 before diverging; fig. S3). As a result, it is unlikely that, had it been the ancestral virus, WA1 would have given rise to only the S566 lineage and the WA outbreak clade ($P = 0.010$). Therefore, to explain the presence of S566 and the WA outbreak clade, we must seriously consider the possibility that there were multiple introductions of genetically similar viruses into the United States.

We thus turned to a distinct phylogeographic approach that explicitly considers the relatively late sampling time of WA-S566, along with other temporal, epidemiological, and geographic data. This method accounts for geographical gaps in sampling and integrates relevant covariates for global spatial spread in a Bayesian framework (16). We investigated how tree topologies were affected by the inclusion of unsampled viruses assigned to 12 of the most severely undersampled locations, both in China and globally, on the basis of COVID-19 incidence data (16). Realistic sampling time distributions were also inferred from COVID-19 incidence data. To better inform placement of unsampled viruses on the phylogeography, we adopted a generalized linear model formulation of the phylogenetic diffusion process (17). This approach estimates a significant contribution for both air passenger flow and asymmetric flow in

and out of Hubei (both with Bayes factors >8042 and positive log effect sizes; supplementary text).

The resulting phylogeny (Fig. 4) provides one reconstruction of the possible evolutionary relationships of WA outbreak viruses and their closest relatives that realistically accounts for major gaps in sequence data. For low-diversity data, a single phylogeny has a resolution that is largely not supported by the full posterior tree distribution containing several plausible phylogeographic scenarios that must be considered, all of which are compatible with the genetic data [e.g., the mutation trees in (2) and those available at nextstrain.org]. The posterior maximum clade credibility (MCC) tree (Fig. 4) suggests that the WA outbreak clade (plus S566 and a sibling virus sampled in New York, “NY”) resulted from an introduction from Zhejiang, China, as supported by the clustering of sampled and unsampled taxa from this location. Additionally, although an introduction from a Chinese location other than Hubei yields considerable posterior support (bar chart inset in Fig. 4), Hubei is preferred over Zhejiang for the entire posterior sample as the most likely source for this introduction. Notably, although the genome from NY (near S566 in Fig. 4) is identical to that of WA1, its much more recent sampling time separates it from WA1 (and, similarly, early Chinese sampling) in the time-calibrated phylogeographic reconstruction. The more recent collection date for both this NY sample and S566, as well as modest support [posterior probability (pp) = 0.67] indicating that they share a U.S. location with the WA outbreak viruses, results in a reconstruction with a

single introduction for these viruses. Using Markov jump estimates that account for phylogenetic uncertainty (18), we inferred 1 February 2020 [95% highest posterior density (HPD): 14 January to 15 February] as the time for this introduction, consistent with the observation that viruses from the WA outbreak clade were likely present during the voyage of the Grand Princess cruise ship to Mexico starting on 11 February (5).

Through a comparison with a time-inhomogeneous model, we show that our estimates are relatively robust to the assumption of constant covariate effect sizes through time (fig. S4). Although the time-inhomogeneous model was fitted to a dataset without unsampled viruses, it also provides strong support for an independent introduction from Hubei (fig. S5). Without unsampled taxa, we estimate a somewhat earlier date for the introduction of the ancestor of the WA outbreak clade plus S566 [26 January 2020 (95% HPD: 15 January to 7 February)], likely because the time-homogeneous analysis allows unsampled taxa from Hubei or other Chinese locations (as in the MCC tree in Fig. 5) to branch off near the WA outbreak clade. In the light of the travel restrictions, specifically those from Hubei, the earlier mean date obtained without unsampled taxa may be the more realistic estimate.

The MCC tree suggests that a Malaysian virus also descended from this introduction (i.e., that it resulted from a subsequent United States-to-Malaysia jump). It is, however, much more plausible that this virus was introduced directly from China to Malaysia, but both the sequence and covariate data in the phylogeographic model lack the information to strongly

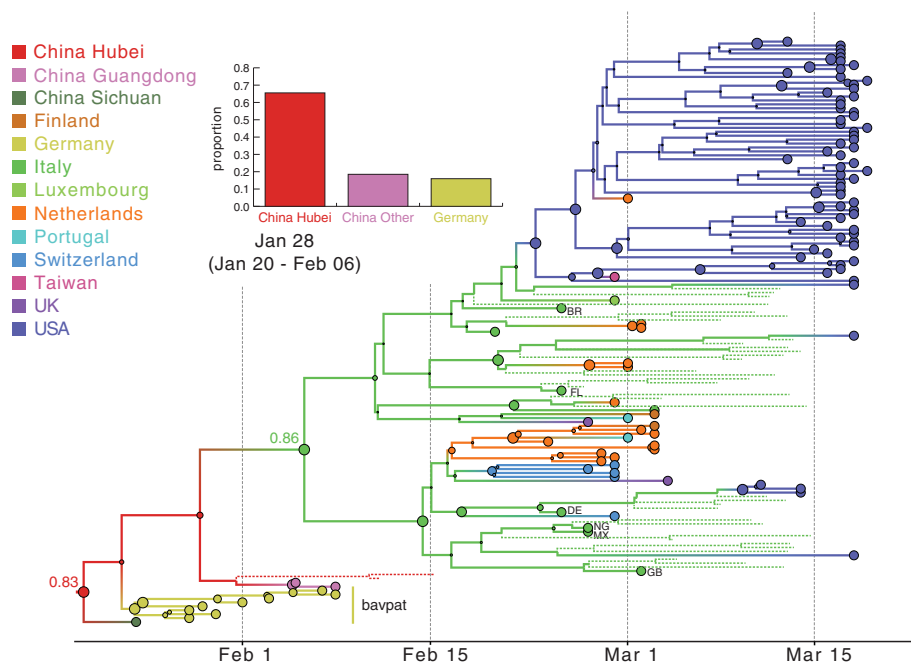


Fig. 5. MCC tree of SARS-CoV-2 entry into Europe. A subtree was inferred for viruses from (i) the first outbreak in Europe (Germany, BavPat) and identical viruses from China, (ii) outbreaks in Italy and New York, and (iii) other locations in Europe. Dotted lines represent branches associated with unsampled taxa assigned to Italy and Hubei, China. Country codes are shown at branch tips for genomes sampled from travelers returning from Italy (BR, Brazil; FL, Finland; DE, Germany; NG, Nigeria; MX, Mexico; GB, United Kingdom of Great Britain and Northern Ireland). The inset bar chart summarizes the probability distribution for the location state ancestral to the Italian clade. Other features as described in Fig. 4.

support this scenario. In light of the simulation results, there is a distinct possibility that S566 and the related NY virus may have descended from a separate introduction from Asia, with the site of arrival in the United States unresolved owing to the presence of both a West Coast and East Coast virus in the clade. Accordingly, an analysis that does not assign a known location to S566 and the related NY virus supports independent introductions from Hubei for these viruses and for the WA outbreak clade (fig. S6), with 7 February (95% HPD: 23 January to 18 February) as the date for the latter.

Consistent with estimates of the introduction date of this viral lineage into Washington state, the Seattle Flu Study tested 6908 archived samples from January and February, of which only 10, from the end of February, were positive (19). Our estimates of the introduction date of the WA outbreak clade into Washington state around the end of January or beginning of February 2020 are ~2 weeks later than they would be if the outbreak had originated with WA1's arrival on 15 January (2), implying that: (i) archived "self-swab" samples retrospectively detected the virus within a few weeks of its arrival (19), (ii) this Washington state outbreak may have been smaller than estimates based on the assumption of a 15 January arrival of WA1, and (iii) the individual who introduced

the founding virus likely arrived in the United States when entry to the country was suspended for non-U.S. residents from China (beginning on 2 February 2020) (20), perhaps during the concurrent period when ~40,000 U.S. residents were repatriated from China, with screening described as cursory or lax (27). These passengers were directed to a short list of airports, including those in Los Angeles, San Francisco, New York, Chicago, Newark, Detroit, and Seattle (27). The late-February timing of COVID-19 cases in Solano County and Santa Clara County in California (5) (supplementary text) suggests that self-limited outbreaks may have originated from returning U.S. residents during this period. So although our reconstructions incorporating unsampled lineages do not account for travel restrictions, the remaining influx likely provided an opportunity for a second introduction of virus (distinct from the WA1 lineage), or even multiple such introductions, into Washington state. Recent inferences that there have been >1000 independent introductions of SARS-CoV-2 into the United Kingdom (22) lend support to this idea.

Early establishment of SARS-CoV-2 in Europe

We used a similar approach to investigate whether the Northern Italy SARS-CoV-2 outbreak was introduced from the German outbreak

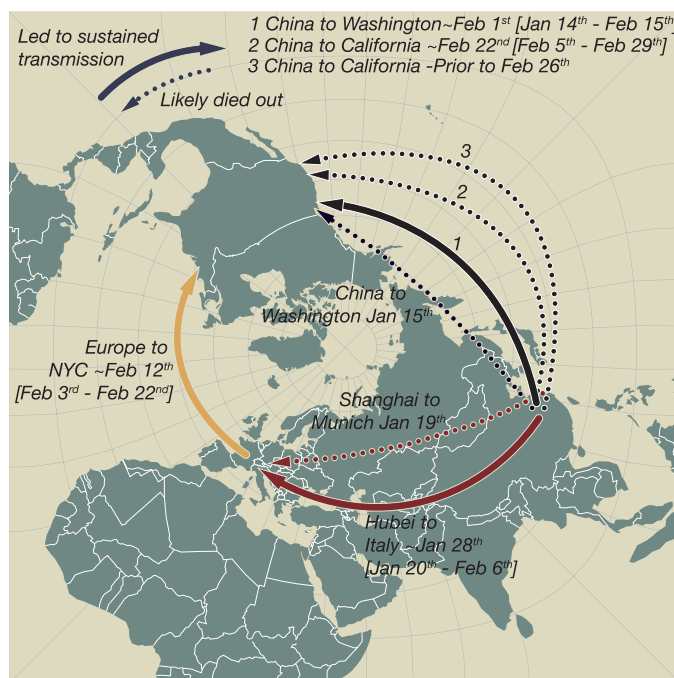
or independently from China: We simulated the Northern Italy outbreak under the hypothetical constraint that it was initiated by a virus imported from the German outbreak (fig. S7) and conducted phylogeographic analyses (Fig. 5). Our simulation framework suggested that the outbreak in Bavaria, Germany, was unlikely to be responsible for initiating the Italian outbreak (see fig. S7 and supplementary results for detailed phylogenetic scenarios). We again used realistic epidemiological parameters to simulate the origins of the Italian outbreak under the assumption that it was associated with viruses genetically related to the German virus BavPat1. Simulations with a median doubling time of 3.4 days (95% range: 2.9 to 4.4 days) resulted in a median epidemic size (including outbreaks that died out) of 725 infections (95% range: 140 to 2847 infections) after 36 days. In the observed phylogeny, the Italian outbreak is the sole descendant lineage from BavPat1. Within the Italian outbreak, no viruses are identical to BavPat1, and 4 of the 27 related viruses included in this analysis are separated from BavPat1 by a single mutation. In simulation, the distributions of identical and one-mutation-divergent viruses are not significantly different from expectation ($P = 0.156$ and 0.157 , respectively). However, the lack of at least one descendant lineage that forms a polytomy with BavPat1 and the Italian outbreak significantly differs from expectation ($P = 0.004$). Therefore, it is highly unlikely that BavPat1 or a virus identical to it initiated the Italian outbreak (fig. S7). As with the WA outbreak, these findings were robust to different infection rates and faster evolutionary rates (supplementary text). Notably, therefore, both a WA1-origin of the WA outbreak and a German origin of the Italian outbreak are rejected even by misspecified models of the epidemiological and evolutionary process.

An alternative scenario in which the outbreaks in both Germany and Italy were independently introduced from China is further supported by our phylogeographic inference (Fig. 5). The resulting reconstruction provides stronger support for independent viral introductions from China into Germany and into Italy ($pp = 0.84$) than for a direct connection between Germany and Italy ($pp = 0.16$) (Fig. 5). Similar support is obtained for this scenario by a time-inhomogeneous inference without unsampled taxa (fig. S8). These findings emphasize that epidemiological linkages inferred from genetically similar SARS-CoV-2 associated with outbreaks in different locations can be highly tenuous, given low levels of sampled viral genetic diversity and insufficient background data from key locations.

Our approach infers that the European B.1 clade (emanating from the green node labeled 0.86 in Fig. 5), which also dominates in NYC

Fig. 6. SARS-CoV-2 introductions to Europe and the United States.

Pierce projection mapping of early and apparently “dead-end” introductions of SARS-CoV-2 to Europe and the United States. Successful dispersals between late January and mid-February are shown with solid arrows: from Hubei Province, China, to Northern Italy; from China to Washington state; and later from Europe (as the Italian outbreak spread more widely) to NYC and from China to California. Dashed arrows indicate dead-end introductions.



(14) and Arizona (23), originated in Italy, as might be expected from the epidemiological evidence. Both travel history and unsampled diversity contribute to this inference. Although only two samples in our dataset are from Italy, five additional genomes were obtained from people who arrived from Italy (Fig. 5). The unsampled taxa from Italy further contributed to a reconstruction with stronger support for Italy at the origin of the entire clade (Fig. 5 versus fig. S8; also see fig. S9). The introduction from Hubei to Italy was dated to 28 January 2020 (95% HPD: 20 January to 6 February). This Italian-and-European cluster, in turn, was the source of multiple introductions to NYC (14). Using the same approach, we dated the introduction that led to the largest NYC transmission cluster to 12 February 2020 (95% HPD: 3 February to 22 February). This is consistent with the finding that the earliest seropositive samples in NYC were from the week of 17 February through 23 February (24).

Hence, even though a second introduction into Washington state (independent of WA1) implies that the Washington transmission cluster had a more recent origin date than under the WA1-origin scenario (~1 February versus 15 January, if it had originated with WA1), the WA outbreak clade still predates the earliest genomically identified transmission clusters elsewhere in the United States: the large one in NYC (4) and two smaller, apparently self-limited clusters in California (in Solano County and Santa Clara County) that appear to have been introduced from China (5). Of these, the transmission cluster from Santa Clara County appears older, dating to before 22 February 2020

(95% HPD: 5 February to 29 February) (supplementary text).

Discussion

Despite the early successes in containment, SARS-CoV-2 eventually took hold in both Europe and North America during the first 2 months of 2020—first in Italy around the end of January, then in Washington state around the beginning of February, and followed by NYC later that month. Our analyses therefore delineate when widespread community transmission was first established on both continents (Fig. 6) and clarify the period before SARS-CoV-2 establishment when contact tracing and isolation might have been most effective.

Our findings highlight the potential value of establishing intensive, community-level respiratory virus surveillance architectures, such as the Seattle Flu Study, during a pre-pandemic period. The value of detecting cases early, before they have bloomed into an outbreak, cannot be overstated in a pandemic situation (25). Given that every delay in case detection reduces the feasibility of containment, it is also worth assessing the impact of lengthy delays in the U.S. Food and Drug Administration's approval of testing the Seattle Flu Study's stored samples for SARS-CoV-2.

The public health response to the WA1 case in Washington state and the particularly impressive response to an early outbreak in Germany delayed local COVID-19 outbreaks by a few weeks and bought crucial time for U.S. and European cities, as well as those in other countries, to prepare for the virus when it finally did arrive. Surveillance efforts and

genomic analyses subsequently helped close the gap between the onset of sustained community transmission and mitigation measures in Washington state, relative to other locales such as NYC. However, our evidence suggests that the period between the founding of the outbreak and the initiation of mitigation measures in Washington state was not as long as supposed under the WA1-origin hypothesis and that the outbreak may therefore have been smaller than some estimates based on that hypothesis.

Because the evolutionary rate of SARS-CoV-2 is slower than its transmission rate, many identical genomes are rapidly spreading. This genetic similarity places limitations on some inferences, such as calculating the ratio of imported cases to local transmissions in a given area. Nevertheless we have shown that, precisely because of this slow rate, when viral genomes are separated by as few as one mutation, this difference can provide enough information for hypothesis testing when appropriate methods are employed. Bearing this in mind will put us in a better position to understand SARS-CoV-2 in the coming years.

REFERENCES AND NOTES

1. Trevor Bedford (@trvr), "The team at the @seattleflu study have sequenced the genome the #COVID19 community case reported yesterday from Snohomish County, WA, and have posted the sequence publicly to gisaid.org. There are some enormous implications here. 1/9" Twitter, 29 February 2020, 11:20 p.m.; <https://twitter.com/trvr/status/1233970271318503426>.
2. T. Bedford et al., *Science* **370**, 571–575 (2020).
3. J. R. Fauver et al., *Cell* **181**, 990–996.e5 (2020).
4. A. S. Gonzalez-Reiche et al., *Science* **369**, 297–301 (2020).
5. X. Deng et al., *Science* **369**, 582–587 (2020).
6. Trevor Bedford (@trvr), "This separate introduction may have been to British Columbia or may have been elsewhere. Better resolving this introduction geographically would benefit from additional sequencing of samples collected closer in time to the introduction event. 14/18." Twitter, 25 May 2020, 7:35 p.m.; <https://twitter.com/trvr/status/1265063937663328256>.
7. M. L. Holshue et al., *N. Engl. J. Med.* **382**, 929–936 (2020).
8. A. Harmon, "Inside the Race to Contain America's First Coronavirus Case." *The New York Times*, 5 February 2020; www.nytimes.com/2020/02/05/us/coronavirus-washington-state.html.
9. D. A. Bolduc, "Webasto disputes link to Italy coronavirus outbreak." *Automotive News*, 9 March 2020; www.autonews.com/suppliers/webasto-disputes-link-italy-coronavirus-outbreak.
10. M. M. Böhrer et al., *Lancet Infect. Dis.* **20**, 920–928 (2020).
11. C. Rothe et al., *N. Engl. J. Med.* **382**, 970–971 (2020).
12. P. Forster, L. Forster, C. Renfrew, M. Forster, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 9241–9243 (2020).
13. A. Rambaut et al., *bioRxiv* 2020.04.17.046086 [Preprint], 19 April 2020. <https://doi.org/10.1101/2020.04.17.046086>
14. M. T. Maurano et al., *medRxiv* 2020.04.15.20064931 [Preprint], 19 August 2020. [doi:10.1101/2020.04.15.20064931](https://doi.org/10.1101/2020.04.15.20064931)
15. E. Minskaia et al., *Proc. Natl. Acad. Sci. U.S.A.* **103**, 5108–5113 (2006).
16. Materials and methods are available as supplementary materials.
17. P. Lemey et al., *PLOS Pathog.* **10**, e1003932 (2014).
18. V. N. Minin, M. A. Suchard, *Philos. Trans. R. Soc. B* **363**, 3985–3995 (2008).

19. H. Y. Chu *et al.*, *N. Engl. J. Med.* **383**, 185–187 (2020).
20. The White House, “Proclamation on Suspension of Entry as Immigrants and Nonimmigrants of Persons who Pose a Risk of Transmitting 2019 Novel Coronavirus” (2020); www.whitehouse.gov/presidential-actions/proclamation-suspension-entry-immigrants-nonimmigrants-persons-pose-risk-transmitting-2019-novel-coronavirus/.
21. S. Eder, H. Fountain, M. H. Keller, M. Xiao, A. Stevenson, “430,000 People Have Traveled From China to U.S. Since Coronavirus Surfaced.” *The New York Times*, 15 April 2020; www.nytimes.com/2020/04/04/us/coronavirus-china-travel-restrictions.html.
22. O. Pybus *et al.*, “Preliminary analysis of SARS-CoV-2 importation and establishment of UK transmission lineages.” *Virological*, 8 June 2020; <https://virological.org/t/preliminary-analysis-of-sars-cov-2-importation-establishment-of-uk-transmission-lineages/507>.
23. J. T. Ladner *et al.*, *medRxiv* 2020.05.08.20095935 [Preprint]. 13 May 2020. doi:10.1101/2020.05.08.20095935.
24. D. Stadlbauer *et al.*, *medRxiv* 2020.06.28.20142190 [Preprint]. 29 June 2020. doi:10.1101/2020.06.28.20142190.
25. M. Worobey, *Nature* **546**, 355–357 (2017).
26. Worobeylab, Worobeylab/SC2_outbreak: Release with published paper (Version v1.0), Zenodo (2020); <http://doi.org/10.5281/zenodo.3979896>.

ACKNOWLEDGMENTS

We thank the patients and healthcare workers who made the collection of this global viral dataset possible and all those who made viral genomic data available for analysis. We thank N. Moshiri for guidance on FAVITES, T. Bedford for insights into how viral genomic inferences influenced public health responses in Washington state, and L. du Plessis for insights into the timing of the origin of the WA outbreak clade

on the basis of Grand Princess voyage dates. **Funding:** M.W. was supported by the David and Lucile Packard Foundation as well as the University of Arizona College of Science. This work was supported by the Multinational Influenza Seasonal Mortality Study (MISMS), an ongoing international collaborative effort to understand influenza epidemiology and evolution, led by the Fogarty International Center, NIH. The research leading to these results has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement 725422-ReservoirDOCS) and from the European Union's Horizon 2020 project MOOD (grant agreement 874850). The Artic Network receives funding from the Wellcome Trust through project 206298/Z/17/Z. J.O.W. acknowledges funding from the NIH (K01AI110181, AI135992, and AI136056). P.L. acknowledges support by the Research Foundation–Flanders (“Fonds voor Wetenschappelijk Onderzoek–Vlaanderen,” G066215N, G0D5117N, and G0B9317N). M.A.S. acknowledges support from NIH U19 AI135995. J.B.J. is thankful for support from the Canadian Institutes of Health Research Coronavirus Rapid Response Programme 440371 and Genome Canada Bioinformatics and Computational Biology Programme grant 287PHY. J.P. acknowledges funding from the NIH (T15LM011271). V.H. acknowledges funding from the Biotechnology and Biological Sciences Research Council (BBSRC) (grant BB/M010996/1). The content of this paper is solely the responsibility of the authors and does not necessarily represent official views of the NIH. We gratefully acknowledge support from NVIDIA Corporation, with the donation of parallel computing resources used for this research. **Author contributions:** Conceptualization: M.W. Methodology: M.W., J.P., M.A.S., P.L., and J.O.W. Software: J.P., M.A.S., P.L., and J.O.W. Validation: J.P., M.A.S., and P.L. Formal analysis: M.W., J.P., P.L., and M.A.S. Investigation: M.W., J.P., B.B.L., J.B.J., A.R., M.I.N., and V.H. Resources: M.W., P.L., and M.A.S. Data curation: B.B.L., J.B.J., and V.H. Writing – original draft: M.W. and M.I.N. Writing – review and editing: M.W., B.B.L., M.A.S., J.O.W., J.B.J., and A.R. Visualization:

B.B.L., J.O.W., and A.R. Supervision: M.W. and J.O.W. Project administration: M.W. Funding acquisition: M.W., M.A.S., and J.O.W. **Competing interests:** J.O.W. has received funding from Gilead Sciences, LLC (completed) and the CDC (ongoing) via grants and contracts to his institution that are unrelated to this research. M.A.S. receives funding from Janssen Research & Development, IQVIA and Private Health Management via contracts unrelated to this research. **Data and materials availability:** All data used in this analysis are free to access: a BEAST .xml file example, FAVITES simulated phylogenies, the GISAID accession numbers for all sequences used in the analysis, and alignments are hosted at GitHub (https://github.com/Worobeylab/SC2_outbreak) and Zenodo (26). This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/564/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S9
Tables S1 to S3
References (27–40)
MDAR Reproducibility Checklist

18 May 2020; accepted 3 September 2020
Published online 10 September 2020
10.1126/science.abc8169

REPORT

CORONAVIRUS

Cryptic transmission of SARS-CoV-2 in Washington state

Trevor Bedford^{1,2,3,*†}, Alexander L. Greninger^{1,4†}, Pavitra Roychoudhury^{1,4†}, Lea M. Starita^{2,3,†}, Michael Famulare^{5†}, Meei-Li Huang^{1,4}, Arun Nalla⁴, Gregory Pepper⁴, Adam Reinhardt⁴, Hong Xie⁴, Lasata Shrestha⁴, Truong N. Nguyen⁴, Amanda Adler⁶, Elisabeth Brandstetter⁷, Shari Cho^{2,3}, Danielle Giroux³, Peter D. Han^{2,3}, Kairsten Fay¹, Chris D. Frazer³, Misja Ilcisin¹, Kirsten Lacombe⁶, Jover Lee¹, Anahita Kiavand^{2,3}, Matthew Richardson³, Thomas R. Sibley¹, Melissa Truong^{2,3}, Caitlin R. Wolf⁷, Deborah A. Nickerson^{2,3}, Mark J. Rieder^{2,3}, Janet A. Englund^{2,6,8}, The Seattle Flu Study Investigators[†], James Hadfield¹, Emma B. Hodcroft^{9,10}, John Huddleston^{1,11}, Louise H. Moncla¹, Nicola F. Müller¹, Richard A. Neher^{9,10}, Xianding Deng¹², Wei Gu¹², Scot Federman¹², Charles Chiu¹², Jeffrey S. Duchin^{7,13}, Romesh Gautam¹⁴, Geoff Melly¹⁴, Brian Hiatt¹⁴, Philip Dykema¹⁴, Scott Lindquist¹⁴, Krista Queen¹⁵, Ying Tao¹⁵, Anna Uehara¹⁵, Suxiang Tong¹⁵, Duncan MacCannell¹⁶, Gregory L. Armstrong¹⁶, Geoffrey S. Baird⁴, Helen Y. Chu^{2,7,§}, Jay Shendure^{2,3,17,§}, Keith R. Jerome^{1,4,§}

After its emergence in Wuhan, China, in late November or early December 2019, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus rapidly spread globally. Genome sequencing of SARS-CoV-2 allows the reconstruction of its transmission history, although this is contingent on sampling. We analyzed 453 SARS-CoV-2 genomes collected between 20 February and 15 March 2020 from infected patients in Washington state in the United States. We find that most SARS-CoV-2 infections sampled during this time derive from a single introduction in late January or early February 2020, which subsequently spread locally before active community surveillance was implemented.

The novel coronavirus, referred to alternately as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (1) or human coronavirus 2019 (hCoV-19) (2), emerged in Wuhan, Hubei, China, in late November or early December 2019 (3). As of 18 May 2020, there have been >4 million

confirmed cases of coronavirus disease 2019 (COVID-19)—the disease caused by SARS-CoV-2—that have resulted in >300,000 deaths (4). After its initial emergence in China, travel-associated cases with travel histories related to Wuhan appeared in other parts of the world (5). The first confirmed case in the United States was travel associated and was detected in Snohomish County, Washington state, on 19 January 2020. Until 27 February 2020, the U.S. Centers for Disease Control and Prevention (CDC) guidance recommended prioritizing testing for COVID-19 on persons with direct travel history from an affected area or with exposure to a known case. Cases of respiratory disease with no known risk factors were not routinely tested. In the 6 weeks between 19 January and 27 February, 59 confirmed cases were reported in the United States (6), all outside of Washington state and with either direct travel history or exposure to a known, confirmed case. On 28 February 2020, a community case was identified in Snohomish County (7). One month later, on 25 March, as a result of increased testing and ongoing transmission, Washington state reported 2580 confirmed cases and 132 deaths (8). Here, we report on the putative history of early community transmission in Washington state as revealed by genomic epidemiology. We conclude that SARS-CoV-2 was circulating for several weeks undetected by the surveillance appara-

tus in Washington state from late January to early February 2020.

Although publicly available SARS-CoV-2 genomes (9, 10) are not sampled in strict proportion to the burden of infections through time and across geography, their genetic relationships can still shed light on underlying patterns of spread. SARS-CoV-2 genomes sampled between December 2019 and 15 March 2020 appear to be closely related, with between 0 and 12 mutations relative to a common ancestor estimated to exist in Wuhan between late November and early December 2019 (Fig. 1). This pattern is consistent with a reported rate of molecular evolution of $\sim 0.8 \times 10^{-3}$ substitutions per site per year or approximately two substitutions per genome per month (3). After its initial zoonotic emergence in Wuhan (11), SARS-CoV-2 viral genomes began to accumulate substitutions and spread from Wuhan to other regions in the world (3). During December 2019, the Wuhan outbreak was too small to seed many introductions outside of China, but by January 2020, it had grown large enough to begin seeding cases elsewhere (12).

Sequencing of viruses from the Washington state outbreak began on 28 February 2020 and has continued since then. We analyzed the sequences of 455 SARS-CoV-2 viruses from this outbreak collected between 19 January and 15 March 2020 (Fig. 1). Virus sequences from Washington state are closely related to those from viruses collected elsewhere. Clusters of closely related viruses indicate separate introduction events followed by local spread. The majority ($n = 384$; 84%) of these viruses fall into a closely related clade (marked by the larger arrow in Fig. 1), and these viruses have single-nucleotide polymorphisms (SNPs) C8782T, C17747T, A17858G, C18060T, and T28144C relative to the basal virus at the root of the phylogeny, which is equivalent to the reference virus Wuhan/Hu-1/2019. This clade derives from viruses circulating in China (Fig. 1, in blue), is closely related to viruses sampled in British Columbia (Fig. 1, in orange), and is labeled as Pangolin lineage A.1 (13). Going forward, we refer to this clade as the Washington state outbreak clade. Other viruses ($n = 39$; 9%) fall into a separate, smaller clade (marked by the smaller arrow in Fig. 1) and derive from viruses circulating in Europe. The remaining 33 viruses (7%) from Washington state are distributed across the phylogeny. Thus, we conclude that most early cases descend from a single introduction event followed by local amplification.

The Washington state outbreak clade has a highly comb-like structure (Fig. 2A), which is indicative of rapid exponential growth (14). This clade has a C17747T change relative to viruses sampled in British Columbia and a A17858G change relative to viruses sampled in

¹Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. ²Brotman Baty Institute for Precision Medicine, Seattle, WA, USA. ³Department of Genome Sciences, University of Washington, Seattle, WA, USA. ⁴Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA, USA. ⁵Institute for Disease Modeling, Bellevue, WA, USA. ⁶Division of Infectious Disease, Seattle Children's Hospital, Seattle, WA, USA. ⁷Department of Medicine, Division of Allergy and Infectious Diseases, University of Washington, Seattle, WA, USA. ⁸Department of Pediatrics, University of Washington, Seattle, WA, USA. ⁹Biozentrum, University of Basel, Basel, Switzerland. ¹⁰Swiss Institute of Bioinformatics, Lausanne, Switzerland. ¹¹Molecular and Cellular Biology Program, University of Washington, Seattle, WA, USA. ¹²Department of Laboratory Medicine, University of California San Francisco, San Francisco, CA, USA. ¹³Public Health — Seattle & King County, Seattle, WA, USA. ¹⁴Washington State Department of Health, Shoreline, WA, USA. ¹⁵Division of Viral Diseases, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, GA, USA. ¹⁶Office of Advanced Molecular Detection, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA, USA. ¹⁷Howard Hughes Medical Institute, Seattle, WA, USA.

*Corresponding author. Email: tbedford@fredhutch.org

†These authors contributed equally to this work.

‡The Seattle Flu Study Investigators and affiliations are listed in the supplementary materials.

§These authors contributed equally to this work.

Fig. 1. Maximum-likelihood phylogeny of 455 SARS-CoV-2 viruses collected from Washington state on a background of 493 globally collected viruses. Viruses collected from Washington state are shown as red circles. Tips and branches are colored on the basis of location, branch lengths are proportional to the number of mutations along a branch, and the x axis is labeled with the number of substitutions relative to the root of the phylogeny—here equivalent to basal Wuhan outbreak viruses. The clustering of related viruses indicates community transmission after an introduction event. Branch locations are estimated on the basis of a discrete traits model. We observe a single introduction leading to a large outbreak clade of 384 sampled viruses from Washington state (marked by the larger arrow), and we observe a second introduction leading to a smaller outbreak clade of 39 viruses (marked by the smaller arrow). An interactive version of this figure is available at <https://nextstrain.org/community/blab/ncov-cryptic-transmission/introductions>.

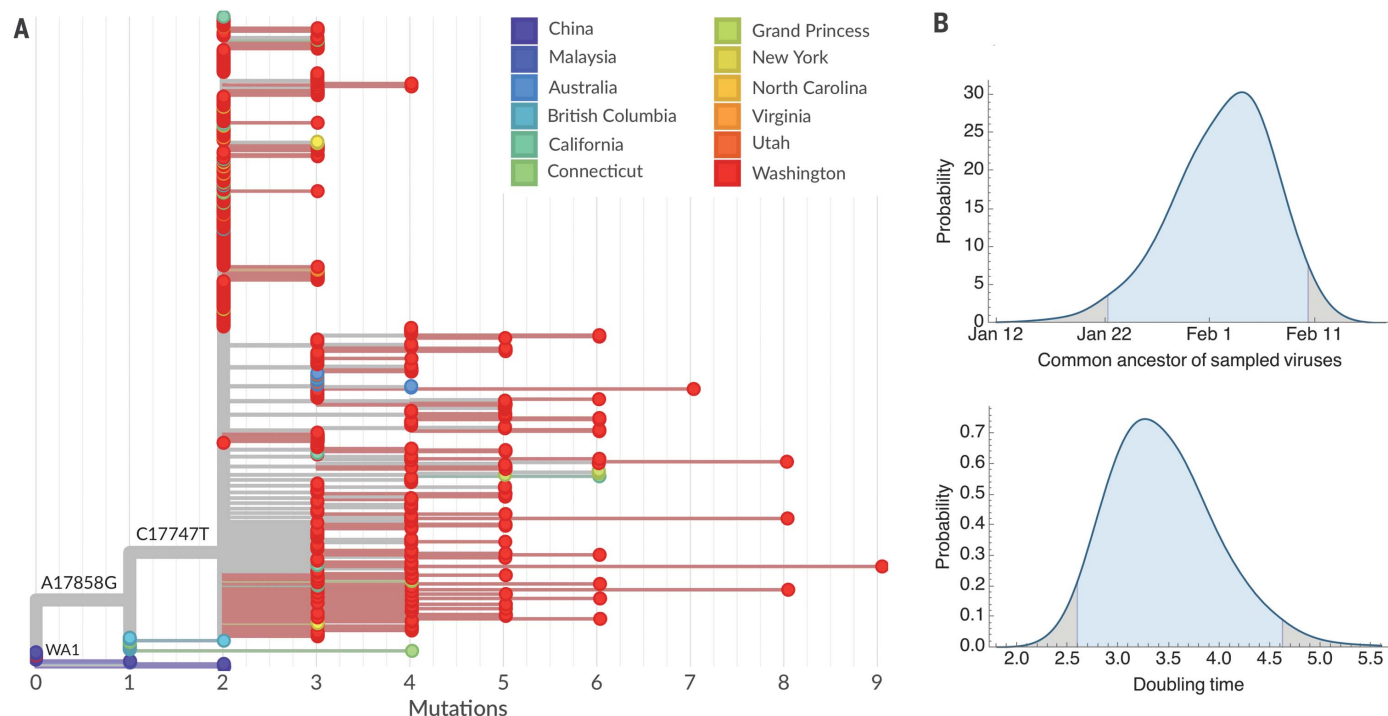
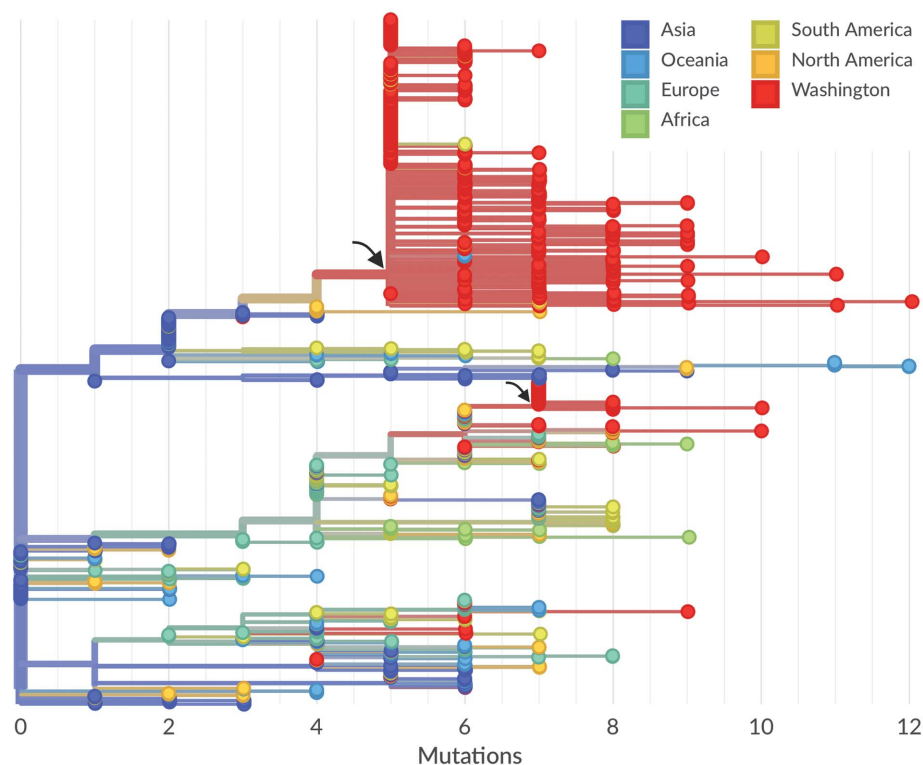


Fig. 2. Maximum-likelihood phylogeny of the Washington state outbreak clade and immediately ancestral variants containing 448 SARS-CoV-2 viruses and Bayesian estimates of the date of the outbreak common ancestor and outbreak doubling time. (A) Maximum-likelihood phylogeny. Tips are colored on the basis of location, branch lengths are proportional to the number of mutations between viruses, and the x axis is labeled with the number of substitutions relative to the root of the phylogeny—here equivalent to the

WA1 haplotype. This comb-like phylogenetic structure of the Washington state outbreak clade is consistent with rapid exponential growth of the virus population. An interactive version of this figure is available at <https://nextstrain.org/community/blab/ncov-cryptic-transmission/wa-clade>. (B) Highest posterior density estimates for the date of the common ancestor of viruses from the Washington state outbreak clade (top) as well as the doubling time in days of the growth of this clade (bottom).

Fujian, Chongqing, Hangzhou, and Guangdong. Given the limited and nonrepresentative sampling of viruses for sequencing, along with the rate of molecular evolution, it is difficult to make detailed assessments of geographic origins. However, we can be confident that this clade represents an introduction from China followed by local spread within the United States and Canada. British Columbia may have been the entry point or the location at which the first virus was sampled.

We analyzed the Washington state outbreak clade in a coalescent analysis to estimate evolutionary dynamics. Here, we assume a prior on evolutionary rate based on analysis of viruses sampled globally between December 2019 and July 2020 (see materials and methods). This analysis uses the degree and pattern of genetic diversity of sampled genomes to estimate the date of a common ancestor and the exponential growth rate of the virus population. We obtained a median estimate for the date of the clade's common ancestor of 2 February 2020,

with a 95% Bayesian credible interval of 22 January to 10 February 2020 (Fig. 2B). We note that the initiation of a transmission chain may slightly predate the common ancestor belonging to this chain in sampled viruses, as initial transmission events after introduction may not result in branching of the transmission tree. We calculated a rate of exponential growth from the coalescent analysis for this clade and found a median doubling time of 3.4 days, with a 95% Bayesian credible interval of 2.6 to 4.6 days (Fig. 2B).

In addition to the 384 viruses from Washington state identified in the Washington state outbreak clade, we observed 12 viruses from elsewhere, including from California, Connecticut, Minnesota, New York, North Carolina, Virginia, Utah, Australia, and the Grand Princess cruise ship (Fig. 2A). Viruses from outside Washington state nest within the diversity found in Washington state. In the case of the Grand Princess, the genetic relationship among these viruses is consistent with a single introduction onto the

cruise ship of the basal outbreak variant—having C17747T and A17858G changes—and subsequent transmission and evolution on the ship.

The first confirmed case recorded in the United States was a travel-associated case from an individual returning from Wuhan on 15 January 2020, who presented for care at an outpatient clinic in Snohomish County on 19 January 2020 and tested positive (15). This infection is recorded as strain USA/WA1/2020 (referred to here as WA1 and annotated in Fig. 2A), and it appears to be closely related to viruses from infections in China (Fujian, Hangzhou, and Guangdong provinces). Viruses from the Washington state outbreak clade group together as direct descendants of WA1 and its identical relatives (Fig. 2A). This tree structure is consistent with the WA1 strain transmitting locally after arrival into the United States. The rarity of the C8782T, T28144C, and C18060T mutations—characteristic of WA1—in viruses sampled from China (found in 6 of 224 or 3% of sequenced viruses) indicates that this is a parsimonious explanation for the origin of the Washington state outbreak clade. However, because the evolution rate for SARS-CoV-2 (one mutation per ~15 days) is slower than the transmission rate (one transmission event every 4 to 8 days) (16, 17), it is possible that WA1 sits on a side branch of the underlying transmission tree even if it appears as a direct ancestor in the maximum-likelihood tree. The fact that viruses sampled from British Columbia interdigitate between WA1 and the Washington state outbreak clade indicates that this clade may have been introduced into North America by a closely related infection to—but one distinct from—WA1 (Fig. 2A). Additionally, it remains possible that multiple viruses with the basal Washington state outbreak clade genotype were introduced, which resulted in the local amplification of this clade; however, this is markedly less likely than a single introduction of the virus.

Given that community transmission was first detected on 28 February 2020 from a transmission chain originating between 22 January and 10 February 2020, we sought to address community prevalence during this period. Here, we analyzed 10,382 acute respiratory specimens collected as part of the Seattle Flu Study between 1 January and 15 March 2020 (Fig. 3A). These specimens represented a mix of residual samples collected as part of routine clinical testing and samples collected as part of prospective community enrollment of individuals with acute respiratory illness. In total, 5270 samples collected between 1 January and 20 February tested negative. The first positive sample was collected on 21 February (Fig. 3B). From 21 February to 15 March, of 5112 samples collected, 65 samples tested positive. On

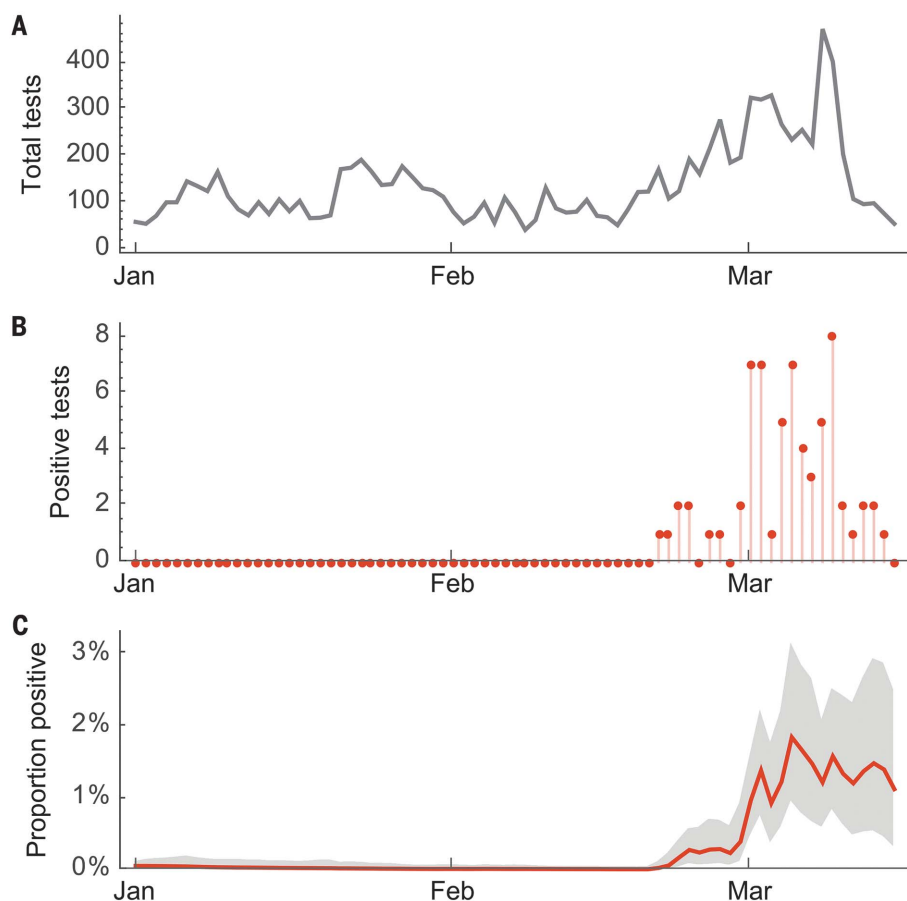


Fig. 3. Acute respiratory samples tested for SARS-CoV-2 collected as part of the Seattle Flu Study between 1 January and 15 March 2020. (A) Total samples tested per day. In total, 10,382 samples collected between 1 January and 15 March were tested. (B) Number of samples testing positive per day. (C) Estimated proportion positive using a sequential Monte Carlo model to provide day-to-day smoothing. The solid red line is the mean estimate of proportion positive, and the gray shaded region is the 95% credible interval. All dates are those of sample collection, not dates of testing.

1 March, a sequential Monte Carlo procedure estimated the proportion of acute respiratory specimens positive for SARS-CoV-2 as 1.1% with a 95% credible interval of 0.5 to 2.0% (Fig. 3C). It is challenging to directly convert this value into population prevalence of SARS-CoV-2; however, U.S. Health Weather data show a 4.5% prevalence of influenza-like illness on 1 March (18), from which we estimated a 0.05% population prevalence of SARS-CoV-2.

In January and February 2020, screening for SARS-CoV-2 in the United States was directed at travelers with fever, cough, and shortness of breath, with the point of origin broadening as new outbreaks were identified but continuing to solely specify travel to China up until 24 February 2020 (19, 20). Our analysis indicates that at least one clade of SARS-CoV-2 had been circulating in the Seattle area for 3 to 6 weeks by the time the virus was first detected in a non-traveler on 28 Feb 2020. By then, variants within this clade constituted the majority of confirmed infections in the region (384 of 455; 84%). Several factors could have contributed to the delayed detection of presumptive community spread, including limited testing among nontravelers or the presence of asymptomatic or mild illnesses.

Both the WA1 strain sampled in Snohomish County, Washington, on 19 January as well as viruses sampled from British Columbia in early March appear to be phylogenetically ancestral to viruses from the Washington state outbreak clade (Fig. 2A), which suggests a possible route of introduction. However, in both of these cases, a lack of comprehensive geographic sampling makes it difficult to rely on phylogenetic structure for transmission inference. Viruses sampled from British Columbia may derive from local spread after a direct introduction event, or they may be offshoots of an introduction elsewhere that subsequently spread to British Columbia. Refining the time and geographic origin of the introduction into Washington state will require a combination of earlier samples and samples from other geographic locations. Other states in the United States have shown different genetic histories from that seen in Washington state, with most SARS-CoV-2 sequences from New York (27) and Connecticut (22) clustering with European lineages, which indicates repeated introductions from Europe. We also observed a second cluster of Washington state viruses related to a later introduction from Europe.

Our results highlight the critical need for widespread surveillance for community transmission of SARS-CoV-2 throughout the United States and the rest of the world, even after the current pandemic is brought under control. The broad spectrum of disease severity (23) makes such surveillance challenging (24).

The combination of traditional public health surveillance and genomic epidemiology can provide actionable insights, as happened in this instance: Upon sequencing the initial community case on 29 February 2020, results were immediately shared with national, state, and local public health agencies, which resulted in the rapid rollout of social distancing policies as Seattle and Washington state came to grips with the extent of existing COVID-19 spread. The confirmation of local transmission in Seattle prompted a change in testing criteria to emphasize individuals with no travel history. From 29 February onward, genomic data were immediately posted to the GISAID EpiCoV sequence database (9, 10) and analyzed alongside other public SARS-CoV-2 genomes by means of the Nextstrain online platform (25) to provide immediate and public situational awareness. We see the combination of community surveillance, genomic analysis, and public real-time sharing of results as a pathway to empower infectious disease surveillance systems.

REFERENCES AND NOTES

1. A. E. Gorbalenya, S. C. Baker, R. S. Baric, R. J. de Groot, C. Drosten, A. A. Gulyaeva, B. L. Haagmans, C. Lauber, A. M. Leontovich, B. W. Neuman, D. Penzar, S. Perlman, L. L. M. Poon, D. Samborskiy, I. A. Sidorov, I. Sola, J. Ziebuhr, Severe acute respiratory syndrome-related coronavirus: The species and its viruses – a statement of the Coronavirus Study Group. *bioRxiv* 2020.02.07.937862 [Preprint]. 11 February 2020. <https://doi.org/10.1101/2020.02.07.937862>.
2. S. Jiang *et al.*, *Lancet* **395**, 949 (2020).
3. A. Rambaut, Phylogenetic analysis of nCoV-2019 genomes (Virological, 2020); <https://virological.org/t/phylogenetic-analysis-176-genomes-6-mar-2020/356>.
4. World Health Organization (WHO), “Coronavirus disease 2019 (COVID-19): Situation report – 55” (WHO, 2020); www.who.int/docs/default-source/coronavirus/situation-reports/20200315-sitrep-55-covid-19.pdf?sfvrsn=33daa5cb_8.
5. F. Schlosser, B. F. Maier, O. Baranov, D. Brockmann, C. Jengen, A. Zachariae, A. Rose, Coronavirus COVID-19 Global Risk Assessment, Event Horizon - COVID-19 (2020); <http://rocs.hu-berlin.de/corona/>.
6. World Health Organization (WHO), “Coronavirus disease 2019 (COVID-19): Situation report – 38” (WHO, 2020); www.who.int/docs/default-source/coronavirus/situation-reports/20200227-sitrep-38-covid-19.pdf?sfvrsn=2db7a09b_4.
7. K. Bray, “Coronavirus update: Addressing questions about a presumptive positive case in an adolescent” (Snohomish Health District Public Health Essentials, 2020); www.snohd.org/Blog.aspx?ID=13.
8. Washington State Department of Health, “2019 Novel Coronavirus Outbreak (COVID-19)” (2020); www.doh.wa.gov/Emergencies/Coronavirus.
9. Y. Shu, J. McCauley, *Euro Surveill.* **22**, 30494 (2017).
10. S. Elbe, G. Buckland-Merrett, *Glob Chall* **1**, 33–46 (2017).
11. K. G. Andersen, A. Rambaut, W. I. Lipkin, E. C. Holmes, R. F. Garry, *Nat. Med.* **26**, 450–452 (2020).
12. N. Imai, I. Dorigatti, A. Cori, C. Donnelly, S. Riley, N. M. Ferguson, “Report 2: Estimating the potential total number of novel Coronavirus cases in Wuhan City, China” (Imperial College London, 2020); <http://hdl.handle.net/10044/1/77150>.
13. A. Rambaut *et al.*, *Nat. Microbiol.* **10**, 1038/s41564-020-0770-5 (2020).
14. E. M. Volz, K. Koelle, T. Bedford, *PLOS Comput. Biol.* **9**, e1002947 (2013).
15. M. L. Holshue *et al.*, *N. Engl. J. Med.* **382**, 929–936 (2020).
16. Q. Li *et al.*, *N. Engl. J. Med.* **382**, 1199–1207 (2020).
17. H. Nishiura, N. M. Linton, A. R. Akhmetzhanov, *Int. J. Infect. Dis.* **93**, 284–286 (2020).

18. A. C. Miller, I. Singh, E. Koehler, P. M. Polgreen, *Clin. Infect. Dis.* **67**, 388–397 (2018).
19. U.S. Centers for Disease Control and Prevention (CDC), “Update and Interim Guidance on Outbreak of 2019 Novel Coronavirus (2019-nCoV)” (CDC Health Alert Network, 2020); <https://emergency.cdc.gov/han/han00427.asp>.
20. U.S. Centers for Disease Control and Prevention (CDC), “Criteria to Guide Evaluation of Persons Under Investigation (PUI) for 2019-nCoV” (CDC, 2020); <https://web.archive.org/web/2020022215422/https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-criteria.html>.
21. A. S. Gonzalez-Reiche *et al.*, *Science* **369**, 297–301 (2020).
22. J. R. Fauver *et al.*, *Cell* **181**, 990–996.e5 (2020).
23. W.-J. Guan *et al.*, *N. Engl. J. Med.* **382**, 1708–1720 (2020).
24. R. Li *et al.*, *Science* **368**, 489–493 (2020).
25. J. Hadfield *et al.*, *Bioinformatics* **34**, 4121–4123 (2018).
26. T. Bedford, blab/ncov-cryptic-transmission: Release 2020-05-31, version 2020-05-31, Zenodo (2020); <https://doi.org/10.5281/zenodo.3871089>.

ACKNOWLEDGMENTS

We gratefully acknowledge the authors and the originating and submitting laboratories of the sequences from GISAID’s EpiFlu Database, on which this research is based. A full acknowledgments table is available as supplementary materials. We have tried our best to avoid any direct analysis of genomic data not submitted as part of this paper and use these genomic data as background. We particularly thank R. Harrigan, N. Prystajek, M. Kraiden, G. Lee, K. Kamelian, H. Lapointe, J. Choi, L. Hoang, I. Sekirov, P. Levett, J. Tyson, T. Snutch, N. Loman, J. Quick, K. Li, and J. Gilmour, who shared virus genomes from British Columbia collected by the British Columbia Centre for Disease Control Public Health Laboratory. We thank N. Thakkar, J. Felsenstein, and C. Spitters for helpful input and discussion. **Funding:** The Seattle Flu Study is run through the Brotman Baty Institute for Precision Medicine and funded by Gates Ventures, the private office of Bill Gates. The funder was not involved in the design of the study and does not have any ownership over the management and conduct of the study, the data, or the rights to publish. J.S. is an Investigator of the Howard Hughes Medical Institute. T.B. is a Pew Biomedical Scholar and is supported by NIH R35 GM119774-01. E.B.H. and R.A.N. are supported by University of Basel core funding. Sequencing analyses of SARS-CoV-2 genomes from California were supported by an NIH grant R33-AI129455 and the Charles and Helen Schwab Foundation to C.C. and by an NIH grant K08-CA230156 and the Burroughs-Wellcome CAMS Award to W.G. **Author contributions:** M.-L.H., A.N., G.P., A.R., H.X., L.S., T.N.N., A.L.G., P.R., G.S.B., and K.R.J. generated sequence and diagnostic data from University of Washington Virology samples. A.A., E.B., S.C., D.G., P.D.H., K.F., C.D.F., M.I., K.L., J.L., A.K., M.R., T.R.S., M.T., C.R.W., D.A.N., M.J.R., J.A.E., T.B., L.M.S., M.F., H.Y.C., and J.S. collected Seattle Flu Study specimens and generated sequence and diagnostic data. T.B., J.H.a., E.B.H., J.Hu., L.H.M., N.F.M., and R.A.N. wrote bioinformatic analysis software and performed phylogenetic analyses. X.D., W.G., S.F., and C.C. generated sequence data from University of California San Francisco samples. R.G., G.M., B.H., P.D., and S.L. collected Washington State Department of Health specimens and generated diagnostic data. K.Q., Y.T., A.U., S.T., D.M., and G.L.A. generated sequence data for the WA1 specimen. T.B., A.L.G., P.R., L.M.S., M.F., J.S.D., G.L.A., H.Y.C., J.S., and K.R.J. interpreted the data and wrote the paper. **Competing interests:** J.A.E. is a consultant for Sanofi Pasteur and Meissa Vaccines, Inc., and she receives research support from GlaxoSmithKline, AstraZeneca, and Novavax. H.Y.C. is a consultant for Merck and GlaxoSmithKline. J.S. is a consultant with Guardant Health, Maze Therapeutics, Camp4 Therapeutics, Nanostring, Phase Genomics, Adaptive Biotechnologies, and Stratos Genomics, and he has a research collaboration with Illumina. G.S.B. is a consultant for Avalon Healthcare Solutions. All other authors declare no competing interests. **Data and materials availability:** Sequencing and analysis of samples from the Seattle Flu Study was approved by the institutional review board at the University of Washington (protocol STUDY00006181). Informed consent was obtained for all community participant samples and survey data. Informed consent for residual sample and clinical data collection was waived. For the University of Washington Virology Laboratory, use of residual clinical specimens was approved by the institutional review board at the University of Washington (protocol STUDY00000408), with a waiver of informed consent. This manuscript represents the opinions of the authors and does not necessarily reflect the position of the U.S. Centers for Disease Control and Prevention. Data and code associated with this

work are available at <https://github.com/blab/ncov-cryptic-transmission> (26). SARS-CoV-2 consensus genome sequences associated with this work have been uploaded to the GISAID EpiFlu database, and accession numbers are available in the supplementary materials. Sequencing reads have been deposited to NCBI SRA (Bioproject PRJNA610428). This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a

copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/571/suppl/DC1
Materials and Methods

Fig. S1
Seattle Flu Study Investigators List
References (27–40)
MDAR Reproducibility Checklist
Data S1

3 April 2020; accepted 8 September 2020
Published online 10 September 2020
10.1126/science.abc0523

PANDEMIC PAUSE

Singing in a silent spring: Birds respond to a half-century soundscape reversion during the COVID-19 shutdown

Elizabeth P. Derryberry^{1*†}, Jennifer N. Phillips^{2,3†}, Graham E. Derryberry¹, Michael J. Blum¹, David Luther⁴

Actions taken to control the coronavirus disease 2019 (COVID-19) pandemic have conspicuously reduced motor vehicle traffic, potentially alleviating auditory pressures on animals that rely on sound for survival and reproduction. Here, by comparing soundscapes and songs across the San Francisco Bay Area before and during the recent statewide shutdown, we evaluated whether a common songbird responsively exploited newly emptied acoustic space. We show that noise levels in urban areas were substantially lower during the shutdown, characteristic of traffic in the mid-1950s. We also show that birds responded by producing higher performance songs at lower amplitudes, effectively maximizing communication distance and salience. These findings illustrate that behavioral traits can change rapidly in response to newly favorable conditions, indicating an inherent resilience to long-standing anthropogenic pressures such as noise pollution.

Actions taken to mitigate the threats of coronavirus disease 2019 (COVID-19) to human life and welfare have inadvertently resulted in a natural experiment offering unanticipated insight into how human behavior affects animal behavior (1). Worldwide, elective quarantine and stay-at-home orders have reduced the use of public spaces and transportation networks, especially in cities. Anecdotal media accounts suggest that restricted movement has elicited rarely observed behaviors in commensal and peri-urban animals (2). Although not all of the reports have proven to be accurate (3), widely publicized observations such as coyotes crossing the normally heavily trafficked Golden Gate Bridge in the San Francisco Bay Area have provoked widespread fascination with the prospect that animals rapidly move back into landscapes recently vacated by humans.

Reports also indicate that animals have been exploiting newly emptied soundscapes. Media outlets have noted people becoming newly aware of more conspicuous animal sounds such as bird songs, particularly in normally noisy areas (4). Although people staying at home may simply be paying closer attention to the animals around them, it is possible that restricted human movement has reduced the use of motorized vehicles, effectively unmasking bird songs otherwise obscured by the associated noise pollution. Theory also suggests

that animals should respond to reduced background noise by altering their acoustic signals to optimize the transmission of information (5, 6). Resolving this uncertainty presents an unprecedented opportunity to address enduring questions about how human behavior alters soundscapes and animal acoustic behaviors (7) while offering vital insight into biotic resilience to long-standing anthropogenic pressures.

Our prior work on soundscapes and bird song across the San Francisco Bay Area provides a strong predictive framework for testing the hypothesis that birds altered their acoustic signaling in response to reduced noise pollution during the recent statewide COVID-19 shutdown. We quantified variation in the soundscape across urban areas (San Francisco and Contra Costa County; hereafter, simply “urban”) and nearby, more rural areas (Marin County; hereafter, simply “rural”; Fig. 1A), focusing on breeding territories of white-crowned sparrows (*Zonotrichia leucophrys*; Fig. 1B), a common songbird in the area (8). We also characterized songs produced by males at a subset of sites within urban and rural areas, circumscribing four independent song populations known as dialects (Fig. 1C and fig. S1). Prepandemic data collected in April to June of 2015 and 2016 were compared with parallel data derived from recordings made at the same sites in April and May of 2020 (i.e., shortly after the execution of regional and statewide shelter-at-home mandates). To provide further context, we have drawn additional comparisons with data collected from the same urban and rural areas in prior decades, leveraging long-term studies of white-crowned sparrow song in the region (9). This approach enabled us to determine (i) whether movement restriction resulted in a sound-

scape with lower anthropogenic noise levels and (ii) if birds responded by adjusting their communication behaviors to improve signal transmission (i.e., efficacy) and functional signal value (i.e., salience).

We have previously shown that white-crowned sparrow songs vary predictably according to transmission properties of urban and rural soundscapes around the San Francisco Bay Area (9, 10). Although urban and rural areas both exhibit spatial variation in noise levels (10), urban soundscapes exhibit more sound energy, particularly at low frequencies, which occurs with greater traffic flow. Birds holding breeding territories in areas with higher noise levels sing higher-amplitude songs (11), a common response to noise known as the Lombard effect (12). Males also produce songs with higher minimum frequencies (13) in areas with high-energy, low-frequency noise typical of traffic in cities (10). Consistent with signal detection theory, this improves signal efficacy by increasing communication distance (6), but it can come at a cost (14). Males that sing at higher minimum frequency often have lower vocal performance (15), which here is the ability to produce rapid trills at wide frequency bandwidths (Fig. 1, D and E) (16). Songs of males with lower vocal performance are less salient in male-male competitive interactions that occur when males defend territories (17, 18). Accordingly, if noise levels decline, then males should produce songs at lower amplitudes (19) and lower minimum frequencies (20), thereby increasing communication distance while also allowing for higher vocal performance.

Comparison of recordings before and during the recent COVID-19 shutdown demonstrated that movement restriction resulted in significantly lower noise levels across more urban areas of the study region, effectively reversing more than a half-century rise in noise pollution (Fig. 2). This is well illustrated in comparisons of background noise (LAF90: the maximum noise level experienced $\geq 90\%$ of the time), which is biologically relevant to songbirds and humans (21). Before the shutdown, breeding territories of urban white-crowned sparrows were on average nearly three times as loud as rural territories ($\beta = 9.2 \text{ dB} \pm 2.1$; $t_{10} = 4.4$, $P < 0.002$, where 9.5 dB is a tripling of sound pressure levels; Fig. 2A and table S1). During the shutdown, background noise was substantially lower across urban territories ($\beta = -7 \text{ dB} \pm 0.8$; $t_{457} = -8.8$, $P < 0.0001$) but not so across rural territories ($\beta = -1.4 \text{ dB} \pm 0.9$; $t_{457} = -1.6$, $P < 0.11$). This is consistent with the observation that traffic is a primary source of background noise across urban San Francisco, whereas the ocean and wind generate background noise in nearby rural Marin County (10). Notably, urban territories no longer exhibited higher noise levels than rural territories ($\beta = 3.6 \text{ dB} \pm 2.2$; $t_{10} = 1.7$, $P < 0.13$). The

¹Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA. ²Department of Biological Sciences, California Polytechnic State University, San Luis Obispo, CA 93407, USA. ³Department of Science and Mathematics, Texas A&M University–San Antonio, San Antonio, TX 78224, USA. ⁴Department of Biology, George Mason University, Fairfax, VA 22030, USA.

*Corresponding author. Email: liz@utk.edu

†These authors contributed equally to this work.

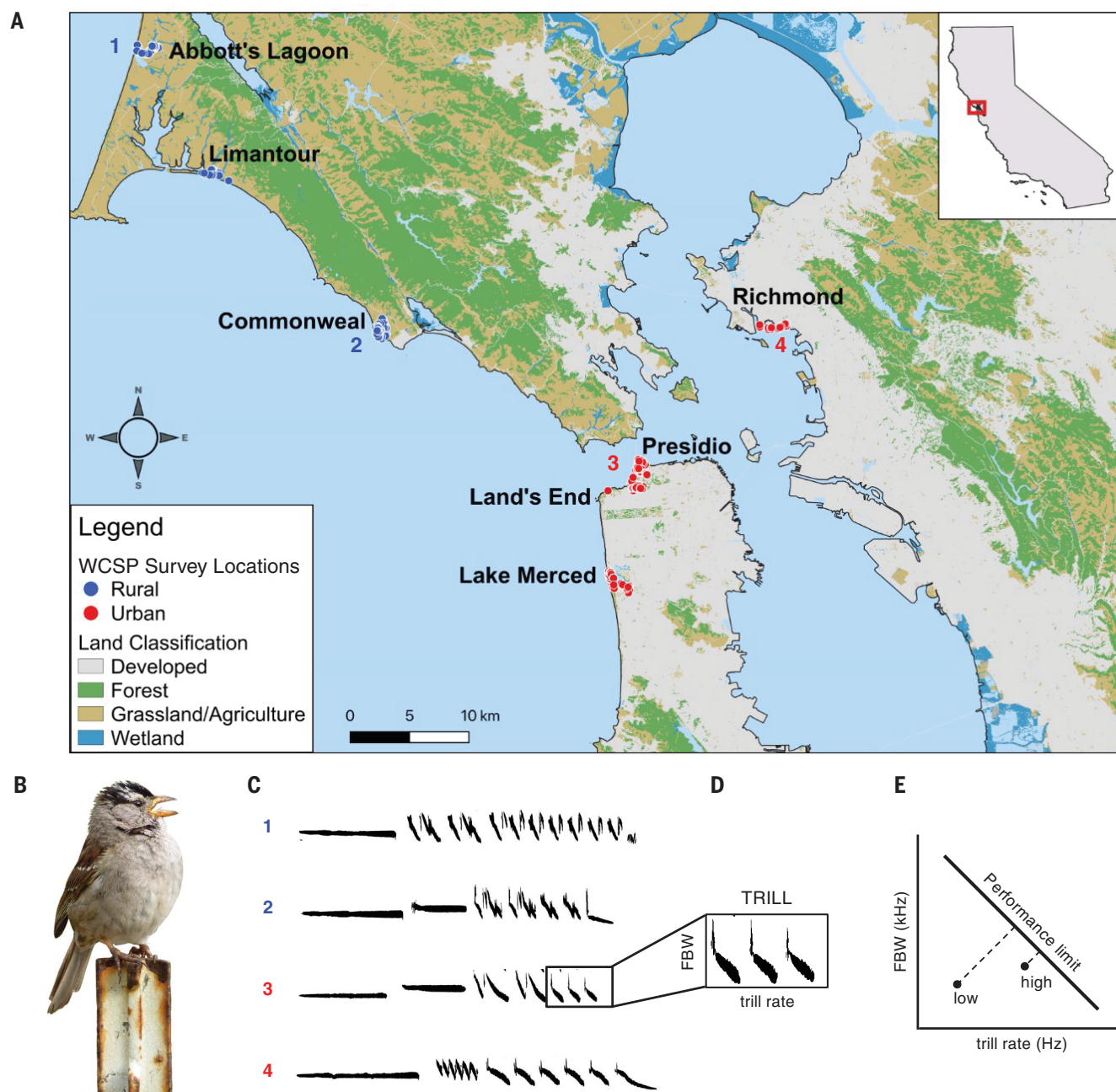


Fig. 1. Study system and song traits of interest. (A) Map illustrating locations where noise was recorded in urban areas (red; Presidio, Land's End, and Lake Merced) of San Francisco (San Francisco County) and Richmond (Contra Costa County), as well as in nearby, more rural areas [blue; Abbott's Lagoon, Limantour, and Commonweal (Marin County)] before (2016) and during (2020) the recent statewide COVID-19 shutdown. Numbers 1 to 4 denote locations where songs were also recorded. (B) Photograph of a male white-crowned sparrow (*Z. leucophrys*) singing in his territory (photo by J.N.P.). (C) Spectrograms of the four song dialects: (1) Drake, (2) Clear, (3) San Francisco, and (4) Berkeley, with numbers corresponding to their respective occurrence at sampling locations depicted in (A). (D) Trilled portion of song boxed on San Francisco dialect to illustrate measurements of trill rate (number of notes/s) and frequency bandwidth (FBW; i.e., the difference between trill maximum and minimum frequency). (E) Males face a physiological limit on producing fast trills at wide frequency bandwidths, resulting in a triangular distribution of songs with an upper bound performance limit. Songs closer to the limit are denoted "high" vocal performance compared with songs further from the limit ("low").

and (4) Berkeley, with numbers corresponding to their respective occurrence at sampling locations depicted in (A). (D) Trilled portion of song boxed on San Francisco dialect to illustrate measurements of trill rate (number of notes/s) and frequency bandwidth (FBW; i.e., the difference between trill maximum and minimum frequency). (E) Males face a physiological limit on producing fast trills at wide frequency bandwidths, resulting in a triangular distribution of songs with an upper bound performance limit. Songs closer to the limit are denoted "high" vocal performance compared with songs further from the limit ("low").

spectral profile of noise on urban territories also converged on that of rural territories (Fig. 2B). The inference that the observed shifts are due to a reduction in the high-energy, low-frequency sound generated by motor vehicles

is supported by traffic flow data from the Golden Gate Bridge. Although vehicle crossings have progressively increased since the bridge opened in 1937, vehicle crossings in April and May of 2020 returned to levels not

seen since 1954 (Fig. 2C). Although noise recordings are not available from the 1950s, this benchmark indicates that the relatively brief but large changes in human behavior effectively erased more than a half-century of urban

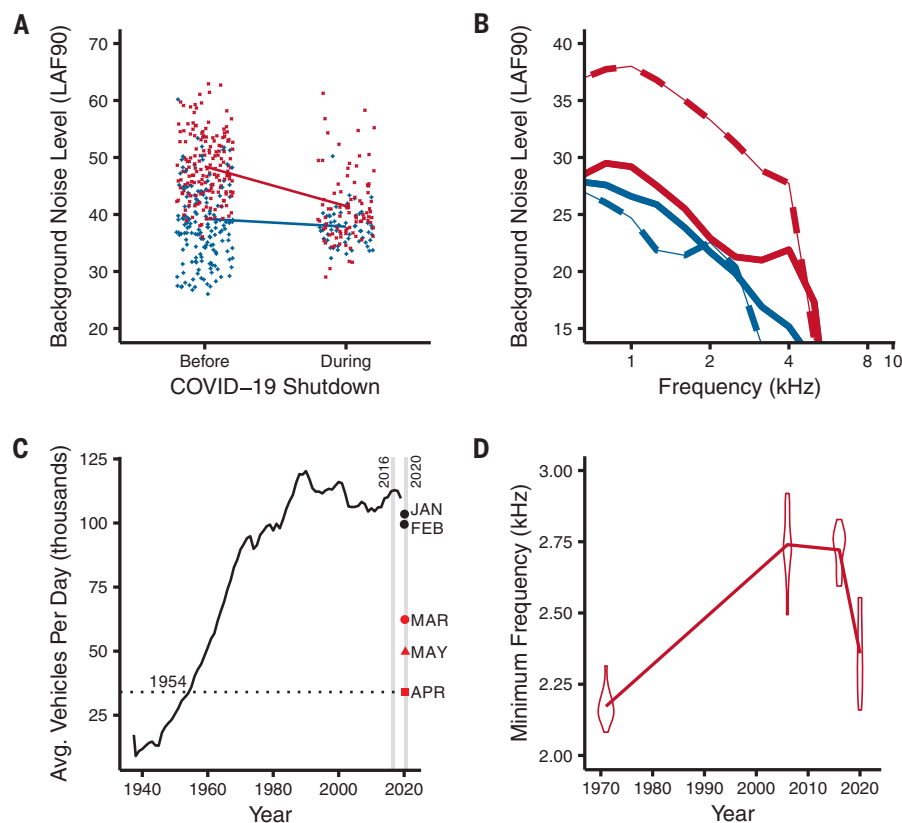


Fig. 2. Background noise levels. (A) Background noise levels recorded in urban (red) and rural (blue) areas, with regression lines representing model-predicted changes in background noise levels before (2016) versus during (2020) the COVID-19 shutdown. (B) Power spectra of background noise levels recorded before (dashed lines) versus during (solid lines) the shutdown. Urban soundscapes during the shutdown converge on the spectral profile of rural soundscapes (note that solid lines are closer together than dashed lines). (C) Toll data showing traffic flow on the Golden Gate Bridge in San Francisco between 1937 and 2020. Vertical gray bars mark the years before (2016) and during (2020) the shutdown, when background noise levels were recorded to illustrate the concurrent decline in the average number of vehicles crossing the bridge per day in March to May 2020, returning to levels not seen since the 1950s in April (dashed line). (D) Violin plots of trill minimum frequency for the Berkeley song dialect recorded in Richmond in 1971, 2006, 2016, and 2020. Line connects the mean for each time point.

noise pollution and the concomitant soundscape divergence between urban and nearby rural areas. In other words, the COVID-19 shutdown created a proverbial “silent spring” across the San Francisco Bay Area.

Movement restriction also resulted in significantly lower ambient noise levels (LAeq), which correspond to the short-term, loud events occurring $\leq 10\%$ of the time (e.g., planes flying overhead or dogs barking). Both urban and rural territories exhibited significantly lower ambient noise levels during the COVID-19 shutdown (urban: $\beta = -7.4 \text{ dB} \pm 0.74$; $t_{466} = -10$, $P < 0.00001$; rural: $\beta = -3.6 \text{ dB} \pm 0.8$; $t_{466} = -4.5$, $P < 0.00001$; fig. S2 and table S2). The drop in ambient noise levels in urban areas was greater than that in rural areas, again resulting in urban territory noise levels converging on those of rural territories.

We found clear evidence that birds responded to the reduction in noise pollution

during the COVID-19 shutdown. Consistent with prior studies (11, 22), we found that birds sang more softly when noise levels were lower ($\beta = 0.27 \text{ dB} \pm 0.04$; $t_{281} = 7.0$, $P < 0.0001$), i.e., the Lombard effect, and at shorter recording distances ($\beta = 0.43 \text{ dB/m} \pm 0.08$; $t_{281} = 5.3$, $P < 0.0001$) before and during the shutdown. Notably, birds produced songs at even lower amplitudes during the shutdown ($\beta = -4.08 \text{ dB} \pm 1.4$; $t_{87} = -3$, $P < 0.004$; Fig. 3, fig. S3, and table S3), well beyond what would be expected from the Lombard effect alone. This departure reveals that prevailing theories of animal communication do not capture the potential magnitude of vocal responses to noise abatement beyond the Lombard effect. Despite a reduction in song amplitude, communication distance more than doubled during the shutdown ($\beta = 8.4 \text{ dB} \pm 1.9$; $t_{87} = 4.4$, $P < 0.0001$; fig. S4 and table S4), further indicating the impact of noise pollution on communication during normal conditions.

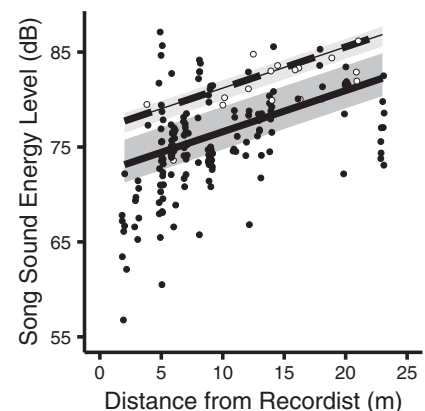


Fig. 3. Scatterplot of song amplitude versus distance between the recordist and bird. Regression lines represent model-predicted changes in song energy levels for songs recorded before (open circles, dashed line) versus during (closed circles, solid line) the COVID-19 shutdown from a slice of immediate ambient noise conditions (43 to 53 dB). The full range of noise levels is illustrated in fig. S3. Gray regions illustrate the middle 50% of the bird random effect (i.e., individual variation).

This doubling in communication distance could elevate fitness by reducing territorial conflicts (23) and increasing mating potential. In addition, the signal-to-noise ratio doubled in relative energy ($\beta = 6.5 \text{ dB} \pm 2$; $t_{95} = 3.3$, $P < 0.002$; table S5), which helps to explain media reports suggesting that bird songs sounded louder during the shutdown (4). A doubling would allow people to hear birds at twice the previous distance, or effectively four times more birds than usual (21).

Birds also exhibited greater vocal performance in response to being released from masking by high-energy, low-frequency noise. We found that birds sang at lower minimum frequencies, achieving greater bandwidth songs in newly open acoustic space (Fig. 4 and tables S6 to S10). An increase in frequency bandwidth results in the transmission of more information and greater vocal performance. Greater vocal performance could also have been achieved through an increase in trill rate, but temporal features of song are not predicted to change with acoustic noise levels. Consistent with this prediction, we found no change in trill rate (top model = null model). Observed changes in performance and related song attributes were much greater in urban than in rural areas, which corresponds to a greater decline in noise levels in urban areas. For example, songs in urban areas exhibited a fourfold greater decrease in minimum frequency ($\beta = -162 \text{ Hz} \pm 26$; $t_{181} = -6.3$, $P < 0.00001$; Fig. 4A) compared with songs in rural areas ($\beta = -40 \text{ Hz} \pm 35$; $t_{181} = -1.2$, $P > 0.26$). This translated into a substantially

greater increase in vocal performance in urban songs ($\beta = 11.1 \pm 0.6$; $t_{181} = 17.3$, $P < 0.00001$; Fig. 4C) compared with rural songs ($\beta = 2.1 \pm 0.8$; $t_{181} = 2.5$, $P < 0.02$). As suggested by our prior experiments (13, 15, 17, 18), the 11-point increase in vocal performance observed in urban areas far exceeds the four-point threshold of change that results in significantly greater signal salience during male-male competition in white-crowned sparrows (17, 18).

Because the same individuals were not sampled at each time point [the mean longevity of white-crowned sparrows is 13 months (24)], we cannot determine whether the observed shift in vocal performance was due to immediate flexibility (25) or if it was because males with higher performance (but typically more masked) songs outcompeted males with lower performance (but less masked) songs for breeding territories during the COVID-19 shutdown. It is nonetheless possible to infer that, on average, birds in urban areas exhibited much greater capacity to compete for breeding territories. This highlights the intriguing possibility that more juveniles preferentially copied higher performance songs during the shutdown. If so, then the shutdown may have altered the trajectory of cultural evolution within and among populations in the study region. Reevaluating the same birds after the resumption of human activity would clarify what behavior(s) gave rise to the observed population-level shift in vocal performance and potential evolutionary outcomes of the COVID-19 shutdown.

Like the half-century soundscape reversion that occurred in more urban areas of the study region, some bird songs exhibited traits during the shutdown that have not been heard in decades, such as trill minimum frequency (fig. S5). Comparisons of historical recordings illustrate that minimum frequencies have tracked a progressive half-century rise in background noise levels in urban songs. Notably, at the Richmond site in Contra Costa County (Fig. 1), the minimum frequency of the Berkeley dialect recorded during the COVID-19 shutdown approached lows not recorded since the spring of 1971 (Fig. 2D) (26).

Although it is well understood that noise pollution can constrain and alter animal communication (7, 19, 20), we show here that alleviation of acoustic pressures can elicit rapid responses that increase acoustic signal efficacy and salience (14). Our findings indicate that songbirds such as white-crowned sparrows have an outstanding capacity to exploit newly empty soundscapes after acute but ephemeral amelioration of noise pollution, suggesting that lasting remediation might engender even more promising outcomes such as demographic recovery and higher species diversity (27) in urban areas. Our findings also show that rapid recovery is possible despite long-standing legacies

of chronic exposure, akin to responses that have been observed after the cessation of chemical pollution that persists in the environment (28). Determining the pace and tempo of responses to the resurgence of noise pollution after the resumption of human activity would help to

shed further light on behavioral resilience. It would also afford opportunities to understand broader dynamics arising from movement restriction, including the possibility that elevated threats to animal welfare lead to complex trade-offs (29). Similarly minded assessments of organismal responses to the amelioration of other forms of pollution during the COVID-19 shutdown, such as reduced CO₂ emissions (30), would also provide exciting opportunities to develop a more integrated understanding of how animals respond to reduced human activity (1), including how and why animals move back into otherwise occupied landscapes and soundscapes.

REFERENCES AND NOTES

1. C. Rutz et al., *Nat. Ecol. Evol.* **4**, 1156–1159 (2020).
2. S. E. Garcia, "When humans are sheltered in place, wild animals will play," *New York Times*, 1 April 2020; <https://www.nytimes.com/2020/04/01/science/coronavirus-animals-wildlife-goats.html>.
3. N. Daly, "Fake animal news abounds on social media as coronavirus upends life," *National Geographic Magazine*, 22 March 2020; <https://www.nationalgeographic.co.uk/animals/2020/03/fake-animal-news-abounds-social-media-coronavirus-upends-life>.
4. D. Greene, "Do those birds sound louder to you? An ornithologist says you're just hearing things," *National Public Radio*, 6 May 2020; <https://www.npr.org/sections/coronavirus-live-updates/2020/05/06/843271787/do-those-birds-sound-louder-to-you-an-ornithologist-says-youre-just-hearing-thin>.
5. J. A. Endler, *Philos. Trans. R. Soc. London B Biol. Sci.* **340**, 215–225 (1993).
6. R. H. Wiley, in *Animal Communication and Noise*, H. Brumm, Ed. (Springer, 2013), vol. 2, pp. 7–30; https://doi.org/10.1007/978-3-642-41494-7_2.
7. J. P. Swaddle et al., *Trends Ecol. Evol.* **30**, 550–560 (2015).
8. Materials and methods are available as supplementary materials.
9. D. Luther, L. Baptista, *Proc. Biol. Sci.* **277**, 469–473 (2010).
10. E. P. Derryberry et al., *PLOS ONE* **11**, e0154456 (2016).
11. E. P. Derryberry et al., *Ecol. Evol.* **7**, 4991–5001 (2017).
12. S. A. Zollinger, H. Brumm, *Curr. Biol.* **21**, R614–R615 (2011).
13. D. A. Luther, E. P. Derryberry, *Anim. Behav.* **83**, 1059–1066 (2012).
14. R. H. Wiley, *Noise Matters: The Evolution of Communication* (Harvard Univ. Press, 2015); <https://doi.org/10.1016/j.anbehav.2016.04.010>.
15. D. A. Luther, J. Phillips, E. P. Derryberry, *Behav. Ecol.* **27**, 332–340 (2016).
16. J. Podos, *Evolution* **51**, 537–551 (1997).
17. J. N. Phillips, E. P. Derryberry, *Anim. Behav.* **132**, 209–215 (2017).
18. J. N. Phillips, E. P. Derryberry, *The Auk* **134**, 564–574 (2017).
19. H. Brumm, S. A. Zollinger, in *Animal Communication and Noise*, H. Brumm, Ed. (Springer, 2013), vol. 2, pp. 187–227; https://doi.org/10.1007/978-3-642-41494-7_7.
20. H. Slabbekoorn, A. den Boer-Visser, *Curr. Biol.* **16**, 2326–2331 (2006).
21. R. J. Dooling, A. N. Popper, "The effects of highway noise on birds," Report prepared for the California Department of Transportation Division of Environmental Analysis (2007); https://www.researchgate.net/publication/228381219_The_Effects_of_Highway_Noise_on_Birds.
22. H. Brumm, *J. Anim. Ecol.* **73**, 434–440 (2004).
23. J. N. Phillips, E. P. Derryberry, *Sci. Rep.* **8**, 7505 (2018).
24. M. C. Baker, L. R. Mewaldt, R. M. Stewart, *Ecology* **62**, 636–644 (1981).
25. G. L. Patricelli, J. L. Blickley, *The Auk* **123**, 639–649 (2006).
26. L. F. Baptista, *Song Dialects and Demes in Sedentary Populations of the White-Crowned Sparrow (Zonotrichia leucophrys nuttalli)* (Univ. of California Press, 1975).
27. L. Manzanares Mena, C. Macías García, *J. Urban Ecol.* **4**, juy022 (2018).
28. T. J. Mozdzer, S. E. Drew, J. S. Caplan, P. E. Weber, L. A. Deegan, *Sci. Total Environ.* **750**, 140927 (2020).
29. A. E. Bates, R. B. Primack, P. Moraga, C. M. Duarte, *Biol. Conserv.* **248**, 108665 (2020).
30. C. Le Quéré et al., *Nat. Clim. Chang.* **10**, 647–653 (2020).

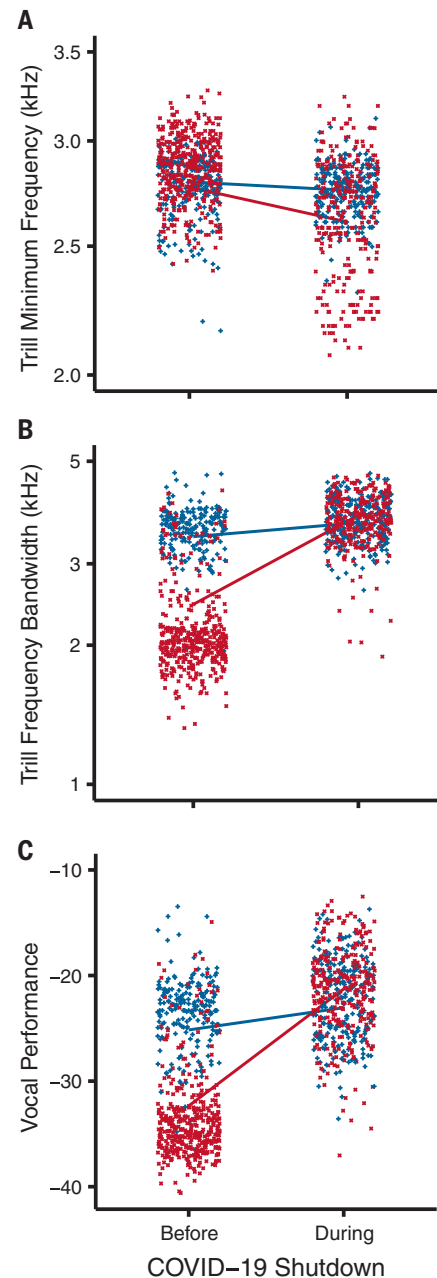


Fig. 4. Scatterplots of raw data. Shown are data recorded in urban (red) and rural (blue) areas, with regression lines representing model-predicted changes in traits before versus during the recent statewide COVID-19 shutdown for trill (A) minimum frequency, (B) frequency bandwidth, and (C) vocal performance. Frequency measures were plotted on a log scale because pitch perception functions on a log scale.

31. E. P. Derryberry, J. N. Phillips, G. E. Derryberry, M. J. Blum, D. Luther, Data for: Singing in a silent spring: Birds respond to a half-century soundscape reversion during the COVID-19 shutdown, Zenodo (2020); <https://doi.org/10.5281/zenodo.4012167>.
32. E. P. Derryberry, J. N. Phillips, G. E. Derryberry, M. J. Blum, D. Luther, Code for: Singing in a silent spring: Birds respond to a half-century soundscape reversion during the COVID-19 shutdown, Zenodo (2020); <https://doi.org/10.5281/zenodo.4012143>.

ACKNOWLEDGMENTS

This work was approved by Tulane University Institutional Animal Care and Use Committee (IACUC) protocol 0427-R and University of Tennessee IACUC protocol 2569, Bird Banding Laboratory Permit 23900, California State Collecting Permit 6799, Golden Gate National Recreation Area (GGNRA) Scientific Research and

Collecting Permit GOGA-00079, San Francisco Parks and Recreation Permit 032014, and Point Reyes National Park (PRNS) Scientific Research and Collecting Permit PORE-0014. We thank M. Berlow and L. Norden for assisting with fieldwork; J. Paschal for assisting with song analyses; J. Cooper for assisting with making the map; Point Blue Palomarin Field Station staff and interns; B. Merkle and M. Chasse at GGNRA; B. Becker at PRNS; D. Bell at East Bay Regional Parks; and L. Wayne at San Francisco Parks and Recreation. **Funding:** This work was supported by the U.S. National Science Foundation (grants 1354763, 1354756, and 1827290) and by an NSF Postdoctoral Research Fellowship in Biology (1812280). **Author contributions:** E.P.D., J.N.P., M.J.B., and D.L. conceived of the study; E.P.D., J.N.P., and D.L. developed the methods and J.N.P. collected the data; E.P.D. and G.E.D. conducted the analyses; and E.P.D. wrote the paper with assistance from J.N.P., M.J.B., and D.L. The manuscript reflects the

contributions and ideas of all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data (31) and code (32) are available at Zenodo.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/575/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S4
Tables S1 to S10
References (33–54)
MDAR Reproducibility Checklist

29 June 2020; accepted 9 September 2020
Published online 24 September 2020
10.1126/science.abd5777

HUMAN EVOLUTION

Denisovan ancestry and population history of early East Asians

Diyyendo Massilani^{1*}, Laurits Skov¹, Mateja Hajdinjak^{1,2}, Byambaa Gunchinsuren³, Damdinsuren Tseveendorj³, Seonbok Yi⁴, Jungeun Lee⁴, Sarah Nagel¹, Birgit Nickel¹, Thibaut Deviese⁵, Tom Higham⁵, Matthias Meyer¹, Janet Kelso¹, Benjamin M. Peter¹, Svante Pääbo^{1*}

We present analyses of the genome of a ~34,000-year-old hominin skull cap discovered in the Salkhit Valley in northeastern Mongolia. We show that this individual was a female member of a modern human population that, following the split between East and West Eurasians, experienced substantial gene flow from West Eurasians. Both she and a 40,000-year-old individual from Tianyuan outside Beijing carried genomic segments of Denisovan ancestry. These segments derive from the same Denisovan admixture event(s) that contributed to present-day mainland Asians but are distinct from the Denisovan DNA segments in present-day Papuans and Aboriginal Australians.

Modern humans may have been present in East Asia as early as 80,000 years before the present (BP) (1, 2), but how they eventually settled in the region remains largely unknown (3–5). To date, genomic data from the first half of the Upper Paleolithic in East Asia exist only for a single human, a ~40,000-year-old individual from Tianyuan Cave in the Beijing area in China (6). This individual was more closely related to present-day East Asians than to ancient Europeans but, surprisingly, shared more alleles with a ~35,000-year-old individual from Belgium (Goyet Q116-1) than with other ancient Europeans (7). In Siberia, which neighbors East Asia to the north, four modern human individuals older than 20,000 years BP have been studied: a ~45,000-year-old individual from Ust'-Ishim in West Siberia who did not contribute ancestry to present-day populations (8); a ~24,000-year-old individual from Mal'ta 1 in South Central Siberia who was more related to Western Europeans than to East Asians and was part of a population that contributed approximately one-third of the ancestry of present-day Native Americans (9); and two ~31,000-year-old individuals from the Yana Rhinoceros Horn Site in northeastern Siberia who show affinities to early modern humans in both West and East Eurasia (10).

In 2006, a hominin skull cap was discovered during mining operations in the Salkhit Valley in northeastern Mongolia (48°16'17.9" N, 112°21'37.9" E) (11) (Fig. 1A). Its unusual morphology led to it being referred to as

Mongolanthropus (11) and later to suggestions that it was affiliated with Neanderthals or *Homo erectus* (12–14). Recently, it was radiocarbon-dated to 33,900 to 34,950 calibrated years BP (95% probability interval), and its mitochondrial (mt) DNA was shown to belong to a basal lineage of the N-haplogroup of modern human mtDNAs that are widespread in Eurasia today (15).

To study the Salkhit individual's nuclear genome, we generated shotgun sequence data from six DNA libraries prepared from bone powder sampled from the Salkhit skull cap (fig. S1) (16). Between 0.6% and 5.6% of the DNA fragments in the libraries mapped uniquely to the human reference genome (hg19) (table S2). Apparent cytosine (C) to thymine (T) substitutions, which are common at the ends of ancient DNA molecules as a result of deamination of cytosine residues (17), affect 23% to 40% of the 5'-ends and 13% to 25% of the 3'-ends in the six libraries, indicating the presence of ancient hominin DNA (table S3 and figs. S2 and S3). Using C to T substitution patterns (18), we estimate the extent of contamination by present-day human DNA to vary between ~5% and ~50% among the libraries (table S3). Because of the high level of human contamination, subsequent analyses (unless specified otherwise) were performed using only DNA sequences showing evidence of cytosine deamination at their first or last base (referred to as “deaminated fragments”), among which contamination is estimated at 1 to 3% (table S4).

The average coverage of the autosomes is similar to that of the X chromosome, indicating that despite the robust morphology of the skull, the Salkhit individual was female (table S8 and fig. S7). To determine to which major group of hominins she belonged, we estimated the percentage of derived alleles shared with the genomes of a present-day human (Mbuti, HGDP00982) (19), a Neanderthal (Denisova 5) (20), and a Denisovan (Denisova 3) (21). We found that 32% of informative posi-

tions covered by deaminated fragments carry alleles seen in the present-day human, whereas 5% and 7% carry alleles seen in the Neanderthal and Denisovan genomes, respectively. This falls within the range seen for present-day Eurasian individuals (table S9 and fig. S8), indicating that the Salkhit individual was a modern human, in agreement with more recent morphological analyses (13, 14).

To investigate the relationship of the Salkhit individual to ancient and present-day modern humans, we enriched the libraries for human DNA fragments by hybridization capture using oligonucleotide probes targeting ~2.2 million single-nucleotide polymorphisms (SNPs) selected to be informative about modern human population history (22–24). Of these, 28% were covered by deaminated DNA fragments in the Salkhit libraries.

We inferred the extent of genetic similarity [using “outgroup” f_3 statistics and D statistics (25)] between the Salkhit individual, modern human individuals older than 20,000 years (table S1), and 131 present-day populations (19). The Salkhit individual, similar to the ~40,000-year-old Tianyuan individual from China, is more related to present-day East Eurasians and Native Americans than to West Eurasians (Fig. 1B, table S10, and fig. S10). Both early East Asians are equally related to most present-day East Eurasians and Native Americans (Fig. 1C and table S12) but differ in their affinity to West Eurasians; present-day West Eurasians share more alleles with the Salkhit individual than with the Tianyuan individual (Fig. 1C). Additionally, the Salkhit individual shares as many alleles with the Tianyuan individual as with the ~31,000-year-old Yana individuals from northeastern Siberia (tables S12 and S14), yet the Tianyuan and Yana individuals share fewer alleles with each other than with the Salkhit individual (Fig. 1C and tables S12 to S14). Those observations suggest that gene flow occurred between populations ancestral to the Salkhit individual and the Yana individuals before ~34,000 years BP—that is, between early populations in East Asia and in Siberia following the divergence of East and West Eurasians. The ~35,000-year-old Goyet Q116-1 individual from Belgium shares more alleles with the Salkhit and Tianyuan individuals (7) than do other Europeans analyzed to date (tables S15 and S16). The fact that the Salkhit individual shares even more alleles with the Goyet Q116-1 individual than does the Tianyuan individual (Fig. 1C) is probably due to gene flow bringing West Eurasian ancestry into the ancestors of the Salkhit individual.

Population admixture models that are compatible with genomic data from modern human individuals older than 20,000 years were evaluated using qpGraph (25) (Fig. 2 and fig. S14). Our models suggest that the Tianyuan individual and the ~37,000-year-old

¹Max Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany. ²Francis Crick Institute, London NW1 1AT, UK. ³Institute of Archaeology, Mongolian Academy of Sciences, Ulaanbaatar 13343, Mongolia. ⁴Department of Archaeology, Seoul National University, Gwanak-gu, Seoul 08826, Korea. ⁵Oxford Radiocarbon Accelerator Unit, Research Laboratory for Archaeology and the History of Art, University of Oxford, Oxford OX1 3QY, UK.

*Corresponding author. Email: diyyendo_massilani@eva.mpg.de (D.M.); paabo@eva.mpg.de (S.P.)

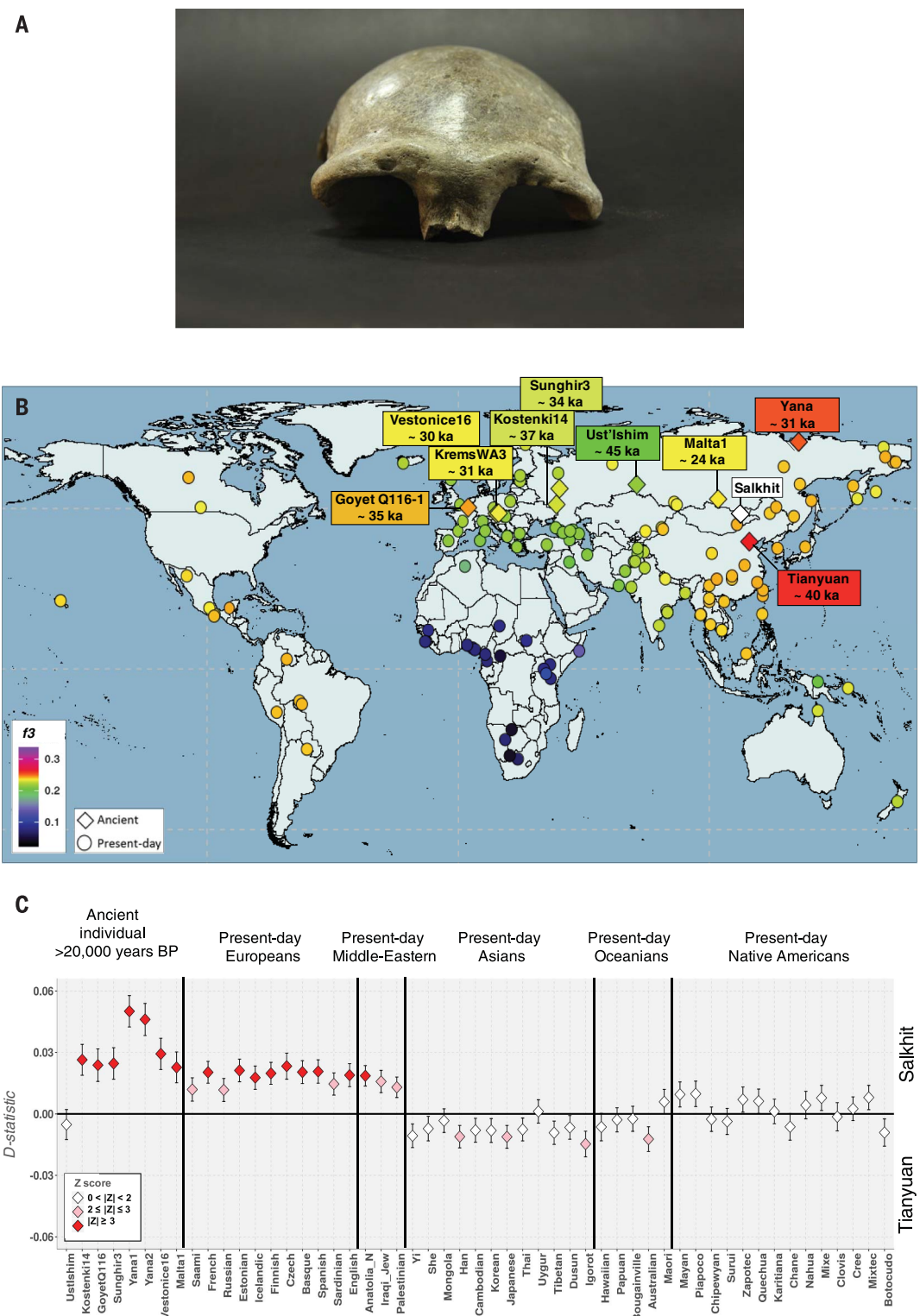
Kostenki14 individual from Russia are unadmixed representatives of early East and West Eurasian populations, respectively. The Salkhit individual, who lived in Mongolia about 6000 years after the Tianyuan individual, carries ~75% of its ancestry from a Tianyuan-related East Eurasian population and the re-

maining ~25% from a population related to the Siberian Yana individuals, who lived some 3000 years later than the Salkhit individual. In agreement with previous results (10), the Yana individuals are estimated to have about one-third of their ancestry from early East Eurasians and the remaining two-thirds from the early

West Eurasians. Their relationship to the Salkhit individual is complex: Models without bidirectional gene flow between an East Asian population ancestral to the Salkhit and a population related to the Yana individuals do not fit the data (fig. S14). Thus, some time before 34,000 years ago, gene flow from West to East Eurasia

Fig. 1. The Salkhit individual and her relationship to ancient and present-day humans.

(A) The Salkhit skull cap. [Image © Institute of Archaeology, Mongolian Academy of Sciences (Mongolia)] **(B)** Heat map illustrating the genetic similarity between the Salkhit individual and modern humans from Eurasia older than 20,000 years (20 ka) (diamonds) as well as present-day human populations (circles) determined by f_3 statistic of the form $f_3(\text{Salkhit}, X; \text{Mbuti})$. The warmer the color, the higher the genetic similarity between the Salkhit individual and a population or individual. **(C)** Relative amounts of allele sharing between the Salkhit and Tianyuan genomes and ancient and present-day humans determined by D statistics of the form $D(\text{Salkhit}, \text{Tianyuan}, X, \text{Mbuti})$. The D statistic is positive when the individual/population shares more alleles with the Salkhit individual than with the Tianyuan individual. The colors of the diamonds indicate whether the Z-score is significant (red), weakly significant (pink), or not significant (white).



occurred, probably mediated by ancestors of the early colonizers of Siberia as represented by the Yana individuals.

We estimate the proportion of Neanderthal ancestry in the Salkhit genome to be ~1.7% (table S17 and fig. S15), similar to other early Eurasians.

As is the case for other Eurasian individuals, the Neanderthal ancestry in the Salkhit individual is equally related to the two Neanderthals from Vindija Cave in Croatia and Chagyrskaya Cave in Siberia who are ~50,000 and 80,000 years old, respectively, and less related to the ~120,000-year-

old Siberian “Altai” Neanderthal from Denisova Cave (20, 26, 27) (fig. S16).

In addition to the Neanderthal ancestry, present-day individuals in East Asia carry ancestry from Denisovans, although in mainland Asia the amount of Denisovan ancestry in present-day populations is less than the amount of Neanderthal ancestry by a factor of 10 (28–30). This has hitherto made it impossible to determine whether ancient genomes from Asia, which are of lower quality than present-day genomes, carry Denisovan ancestry. We applied a novel hidden Markov approach (31) that is able to identify introgressed Neanderthal and Denisovan genomic segments in low-coverage ancient genomes; this method uses a genotype likelihood approach that incorporates contamination, so that all fragments can be used for this analysis. Using data from ~1.7 million SNPs where Neanderthal and/or Denisovan genomes differ from present-day African genomes, we detected 18 segments of Denisovan ancestry longer than 0.2 cM in the Salkhit genome (Fig. 3, table S18, and figs. S17 and S28) and 20 such segments in the Tianyuan genome (table S18 and figs. S19 and S28). We detected about one-third as many

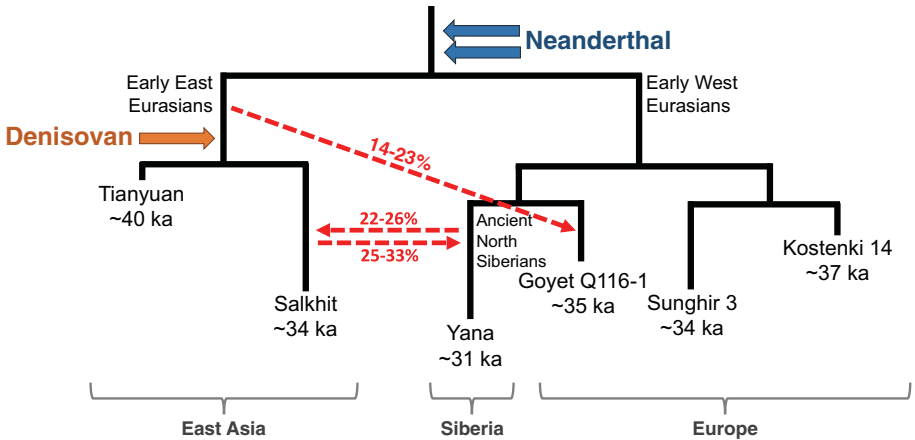
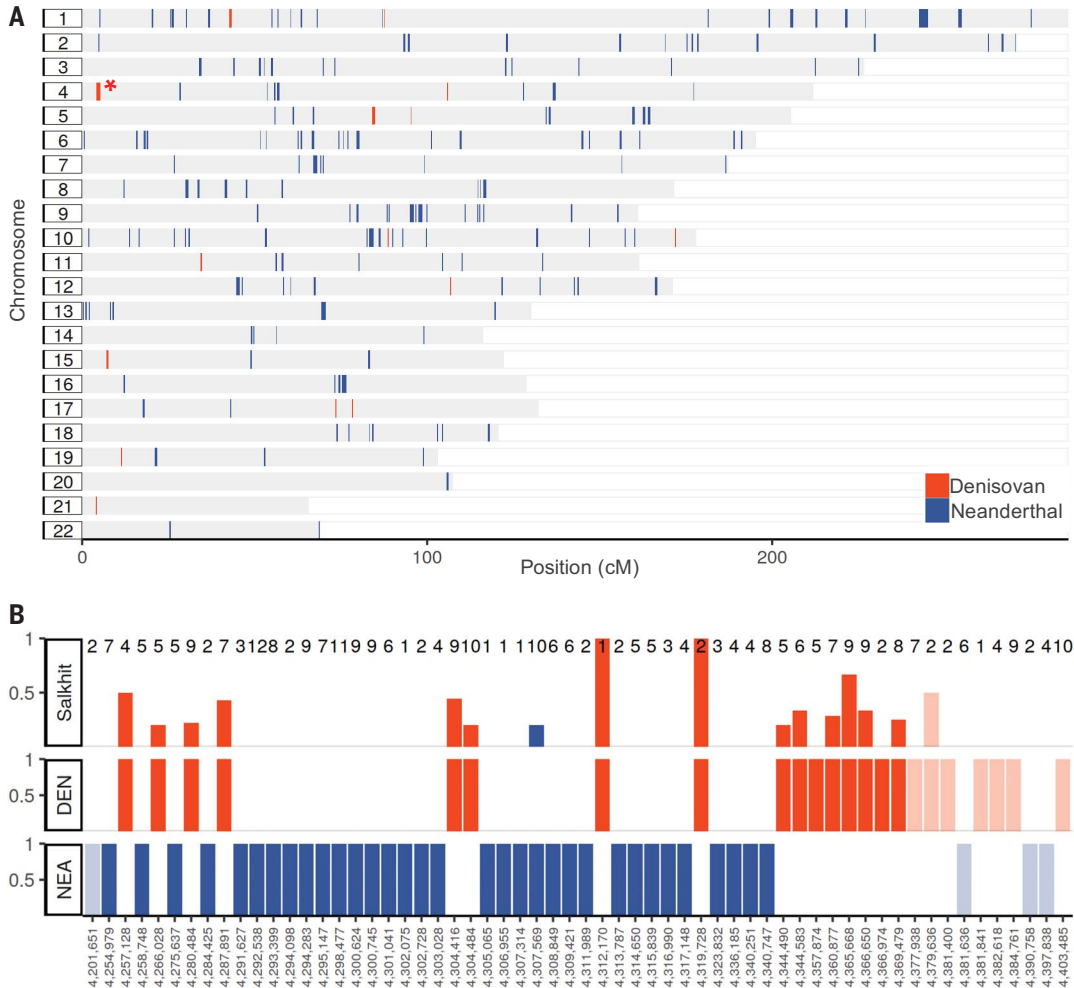


Fig. 2. Simplified demographic model including the Salkhit individual and other Eurasians older than 30,000 years. Admixtures between East and West Eurasians are represented by red arrows. Neanderthal and Denisovan admixtures are indicated by blue and orange arrows, respectively.

Fig. 3. Archaic ancestry in the Salkhit genome.

(A) Genomic distribution of Denisovan (orange) and Neanderthal (blue) DNA segments in the Salkhit genome. (B) Allele frequencies in the longest Denisovan ancestry segment [chr4: 4.16 to 5.31 cM, marked by an asterisk in (A)]. The bars in the top panel give the proportions of Salkhit DNA fragments carrying archaic alleles at sites where alleles are fixed between Africans and the Denisovan genome (red) or between Africans and two Neanderthal genomes (blue). Total numbers of fragments are shown at the top. SNP bars outside the inferred Denisovan segment are faded. Note that in the called region, all Denisovan-like alleles except two occur in the Salkhit genome.



segments of Denisovan DNA in the genomes of the ancient Siberians Yana 1, Yana 2, and Mal'ta 1 (table S18, figs. S19 to S21, and fig. S28), consistent with the proportion of their East Asian ancestry. In contrast, no Denisovan ancestry was detected in the genome of the ~45,000-year-old Siberian individual from Ust'Ishim in West Siberia, nor in any European individual older than 20,000 years (table S18, figs. S22 to S25, and fig. S28). Thus, the Salkhit and Tianyuan genomes provide direct evidence that ancestors of modern humans who lived in

East Asia 40,000 years ago had met and mixed with Denisovans. The small number of these segments does not provide enough power to date the introgression event. However, given their relatively short length (≤ 1.3 cM), the Denisovan introgression is likely to have happened at least 10,000 years before these individuals lived.

One of the risks of inferring ancestry fragments from ancient genomes is that genome quality may affect the ability to detect introgressed segments. Under the assumption that

many of the factors that affect the detection of Denisovan DNA will similarly affect the detection of Neanderthal DNA, the ratio of Denisovan to Neanderthal ancestry segments may be a reasonably robust metric of the relative amount of Denisovan ancestry. In the Salkhit and Tianyuan genomes, these ratios are about 7.5% and 8.1%, respectively. For the genomes of the North Siberians Yana 1 and Yana 2, the ratios are about 3.9% and 4.7%. Because there is no substantial difference in the amounts of Neanderthal DNA in the two

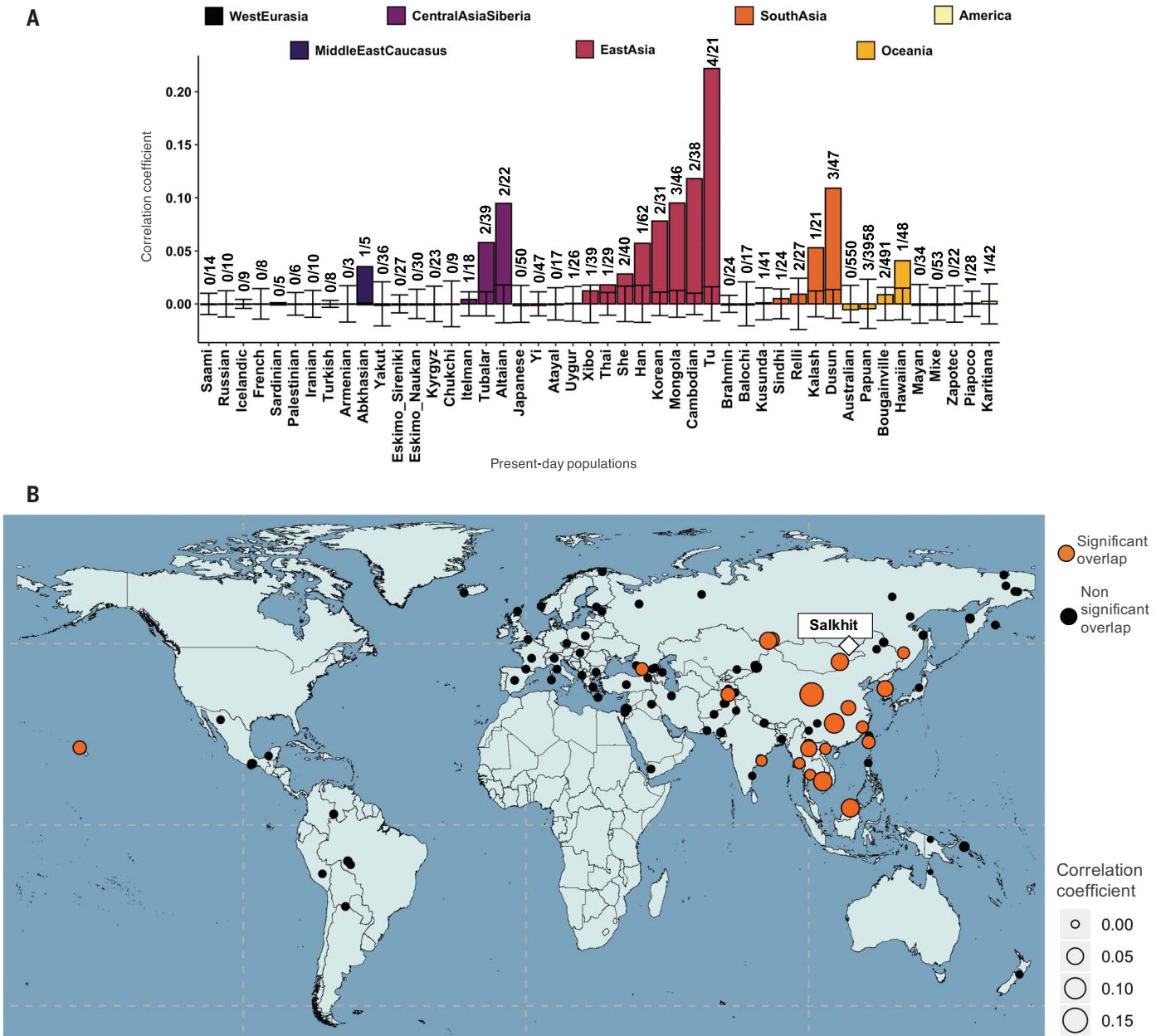


Fig. 4. Overlap of Denisovan segments in the Salkhit genome and present-day non-African populations. (A) Correlation coefficient of the overlap between the Denisovan segments larger than 0.2 cM in the Salkhit genome and Denisovan segments larger than 0.05 cM in 45 present-day Eurasian populations (see fig. S28 for the same with 111 present-day populations). Numbers above the bars

give the number of overlapping segments and the number of segments in the present-day population. The range of correlation coefficients generated by 500 bootstraps is indicated. (B) Geographic locations of present-day populations for which Denisovan ancestry segments overlap significantly with the Salkhit individual (orange circles). Sizes of circles are proportional to the correlation coefficients.

early East Asian genomes and the two Yana genomes (fig. S15), this observation indicates that the ancient Siberian individuals carry less Denisovan DNA than the Salkhit and Tianyuan individuals.

We compared the Denisovan segments in the Salkhit and Tianyuan genomes to those in present-day people to estimate whether introgressed segments between genomes overlap more often than expected by chance. Significance was assessed using 500 bootstrap reshuffles, where segments are randomly relocated across the analyzed genomes (supplementary text). The Denisovan DNA segments in the ancient East Asian genomes overlap more than expected with Denisovan segments detected in the genomes of several present-day populations in Asia and in populations with some Asian ancestry, such as Hawaiians (Fig. 4, table S19, and figs. S29 to S31). In contrast, we find no significant overlap with Denisovan segments detected in Papuans or Aboriginal Australians, although these groups carry on the order of 20 times as much Denisovan DNA as mainland Asians.

It has been shown that at least two Denisovan populations contributed ancestry to present-day East Asian populations, and that Denisovan ancestry in populations in Oceania derived from only one of these sources (29). The overlap of Denisovan DNA segments (Fig. 4, table S19, and figs. S29 to S31) is in agreement with this and suggests that the ancestral population of the Tianyuan and Salkhit individuals that mixed with Denisovans contributed ancestry to populations in large parts of Asia today. In contrast, the lack of any significant overlap with Aboriginal Australians and Papuans suggests that these Oceanian populations received most of their Denisovan ancestry from a different source.

Our results show that the 34,000-year-old Salkhit individual carried more West Eurasian ancestry than the 40,000-year-old Tianyuan individual, indicating that after the major West/East Eurasia split, gene flow from West Eurasia to East Asia occurred earlier than 34,000 years ago, probably mediated by populations related to the Siberian Yana individuals. We also show that these early East Asians carried segments of Denisovan DNA that come from admixture events that also contributed Denisovan DNA to populations across mainland Asia today, but not to Papuans and Aboriginal Australians.

REFERENCES AND NOTES

1. W. Liu *et al.*, *Nature* **526**, 696–699 (2015).
2. V. Michel *et al.*, *J. Hum. Evol.* **101**, 101–104 (2016).
3. N. Zwyns, B. Viola, *L'Asie centrale et la Sibérie durant la pléistocène supérieur in Neandertal/Cro-Magnon: La Rencontre*, M. Otte, Ed. (Errances, 2013), pp. 267–294.
4. M. M. Lahr, R. Foley, *Evol. Anthropol.* **3**, 48–60 (2005).
5. P. Mellars, *Science* **313**, 796–800 (2006).
6. Q. Fu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2223–2227 (2013).
7. M. A. Yang *et al.*, *Curr. Biol.* **27**, 3202–3208.e9 (2017).
8. Q. Fu *et al.*, *Nature* **514**, 445–449 (2014).
9. M. Raghavan *et al.*, *Nature* **505**, 87–91 (2014).
10. M. Sikora *et al.*, *Nature* **570**, 182–188 (2019).
11. D. Tseveendorj, N. Batbold, T. Amgalantugs, *Stud. Archaeol. Inst. Hist. Acad. Sci. Mongolici* **3**, 20 (2007).
12. Y. Coppens, D. Tseveendorj, F. Demeter, T. Turbat, P.-H. Giscard, C. R. *Palevol* **7**, 51–60 (2008).
13. S.-H. Lee, *Homo* **66**, 287–298 (2015).
14. D. Tseveendorj, B. Gunchinsuren, E. Gelegdorj, S. Yi, S. H. Lee, *Quat. Int.* **400**, 175–179 (2016).
15. T. Deviese *et al.*, *Nat. Commun.* **10**, 274 (2019).
16. See supplementary materials.
17. A. W. Briggs *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 14616–14621 (2007).
18. M. Meyer *et al.*, *Nature* **505**, 403–406 (2014).
19. S. Mallick *et al.*, *Nature* **538**, 201–206 (2016).
20. K. Prüfer *et al.*, *Nature* **505**, 43–49 (2014).
21. M. Meyer *et al.*, *Science* **338**, 222–226 (2012).
22. Q. Fu *et al.*, *Nature* **524**, 216–219 (2015).

23. W. Haak *et al.*, *Nature* **522**, 207–211 (2015).
24. Q. Fu *et al.*, *Nature* **534**, 200–205 (2016).
25. N. Patterson *et al.*, *Genetics* **192**, 1065–1093 (2012).
26. K. Prüfer *et al.*, *Science* **358**, 655–658 (2017).
27. F. Mafessoni *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 15132–15136 (2020).
28. D. Reich *et al.*, *Nature* **468**, 1053–1060 (2010).
29. S. R. Browning, B. L. Browning, Y. Zhou, S. Tucci, J. M. Akey, *Cell* **173**, 53–61.e9 (2018).
30. B. Vernot *et al.*, *Science* **352**, 235–239 (2016).
31. B. M. Peter, 100,000 years of gene flow between Neandertals and Denisovans in the Altai mountains. *bioRxiv* [preprint], 15 March 2020.

ACKNOWLEDGMENTS

We thank E. Gelegdorj (Institute of History and Archaeology, Mongolian Academy of Sciences), M. Slatkin (UC Berkeley), P. Skoglund (Francis Crick Institute), Q. Fu (IVPP, Chinese Academy of Sciences), D. Reich (Harvard Medical School), and B. Viola (University of Toronto) for helpful discussions. **Funding:** Supported by the Max Planck Society and the European Research Council through ERC grant 694707 (100 Archaic Genomes) (S.P.) and grant 324139 (PalaeoChron) (T.H. and T.D.). **Author contributions:** D.M., S.N., and B.N. performed the laboratory work. D.M., L.S., M.H., M.M., J.K., B.M.P., and S.P. generated, analyzed, and interpreted the data. D.M., L.S., M.H., and B.M.P. performed the computational work. D.T., B.G., and S.Y. carried out morphological analyses of the fossil. T.D. and T.H. carried out radiocarbon dating. S.Y., D.T., B.G., and J.L. analyzed all archaeological data. D.M. wrote the manuscript with input from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The sequence data are available in http://cdna.eva.mpg.de/modern_human/salkhit and in the European Nucleotide Archive (ENA) under accession number PRJEB37670. The mtDNA sequence is available on GenBank under accession number MT561166. All other data needed to replicate and extend the study are present in the paper and/or the supplementary materials. The Salkhit skull cap is housed at the Institute of History and Archaeology, Mongolian Academy of Sciences, Ulaanbaatar 13343, Mongolia, and can be accessed under request to B.G.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/579/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S31
Tables S1 to S19
References (32–46)
MDAR Reproducibility Checklist

14 May 2020; accepted 10 September 2020
10.1126/science.abc1166

PALEOGENOMICS

Denisovan DNA in Late Pleistocene sediments from Baishiya Karst Cave on the Tibetan Plateau

Dongju Zhang^{1,2,3*}, Huan Xia¹, Fahu Chen^{2,1}, Bo Li^{4,5*}, Viviane Slon⁶, Ting Cheng¹, Ruowei Yang^{7,8}, Zenobia Jacobs^{4,5}, Qingyan Dai⁷, Diyendo Massilani⁶, Xuke Shen¹, Jian Wang^{1,9}, Xiaotian Feng⁷, Peng Cao⁷, Melinda A. Yang¹⁰, Juanting Yao¹, Jishuai Yang¹, David B. Madsen^{1,11}, Yuanyuan Han¹, Wanjing Ping⁶, Feng Liu⁶, Charles Perreault^{12,13}, Xiaoshan Chen¹, Matthias Meyer⁶, Janet Kelso⁶, Svante Pääbo^{6*}, Qiaomei Fu^{7,8*}

A late Middle Pleistocene mandible from Baishiya Karst Cave (BKC) on the Tibetan Plateau has been inferred to be from a Denisovan, an Asian hominin related to Neanderthals, on the basis of an amino acid substitution in its collagen. Here we describe the stratigraphy, chronology, and mitochondrial DNA extracted from the sediments in BKC. We recover Denisovan mitochondrial DNA from sediments deposited ~100 thousand and ~60 thousand years ago (ka) and possibly as recently as ~45 ka. The long-term occupation of BKC by Denisovans suggests that they may have adapted to life at high altitudes and may have contributed such adaptations to modern humans on the Tibetan Plateau.

Denisovans are an extinct hominin group initially identified from a genome sequence determined from a fragment of a phalanx found at Denisova Cave in the Altai Mountains in southern Siberia (1–3). Subsequent analyses of the genome have shown that Denisovans diverged from Neanderthals ~400 thousand years ago (ka) (4) and that at least two distinct Denisovan populations mixed with ancestors of present-day Asians (2–9). Thus, they are assumed to have been widely dispersed across Asia. However, physical remains of Denisovans in Siberia have been restricted to a fragmentary phalanx (1), three teeth (2, 10, 11), and a cranial fragment (12), all of which were found at Denisova Cave.

Recently, half of a mandible from the Baishiya Karst Cave (BKC), Xiahe County, Gansu, China, dated to at least 160 ka, was identified to be of Denisovan origin (13). However, this identification of the Xiahe mandible as Denisovan is based on a single amino acid position and is therefore tenuous. Here we report the results of ongoing archaeological and chronological investigations and sedimentary DNA analyses from BKC. We find evidence for the long-term presence of Denisovans in BKC and provide stratigraphic and chronological context for their occupation in the cave.

BKC (35.45°N, 102.57°E, 3280 m above sea level) is a limestone cave located in the north-eastern margin of the Tibetan Plateau (Fig. 1A and fig. S1A). In 2018 and 2019, three units that measured 1 m by 2 m (T1, T2, and T3) were plotted for excavation in the entrance chamber, which is about 60 m long, 8 m wide, and 5 m high (Fig. 1B and fig. S1, B and C) (14). The second unit (T2) exposed intact cultural strata that are truncated in the southeastern part of the trench by a large pit (H1) dug during the historical period (14), 780 to 700

calibrated years before present (where present is 1950 CE) (figs. S2 and S3 and table S3). Ten stratigraphic layers were identified mainly on the basis of sedimentary characteristics (figs. S2B and S3) (14). Most layers are poorly sorted, composed of a silt matrix with abundant angular clasts of autogenic limestone gravels. The latter originates from the reworking of eroded parent bedrock, sediments by colluviation, or spalling of material from the cave walls and roof (see details in the supplementary materials). Stone artifacts and animal fossils were recovered from all layers (figs. S11 and S12) (14). A total of 1310 stone artifacts and 579 animal bone fragments were recorded and collected. Preliminary analysis of the stone artifact assemblage suggests that they were made mostly from local metamorphic quartz sandstone and hornstone stream cobbles using a simple core and flake technology (fig. S11). Remains of small and medium-size animals, including gazelles, marmots, and foxes, dominate the fossil assemblage in layers 6 through 1, whereas large animals, such as rhinoceros, large bovids, and hyenas, dominate layers 10 through 7 (fig. S12).

We constructed a numerical chronology for the T2 sequence from optical dating of 12 sediment samples and radiocarbon dating of 14 bone fragments (Fig. 2, fig. S3, and tables S3, S9, and S10). The age estimates were used to develop a Bayesian model for the depositional chronology of the site and to provide an age framework for hominin occupation (Fig. 2 and table S12). Details of sample locations and collection, preparation, measurement, and data analysis procedures are provided, together with the measured and modeled ages and related data (14). The deposits in layers 10 to 4 have a stratigraphically coherent chronology, limited age variation within layers, and equivalent dose (D_e) distributions that show minimal evidence for mixing. Layer 10 accumulated between 190 ± 34 and 129 ± 20 ka (here and below, we give modeled age estimates and total uncertainties at 95.4% probability), followed by relatively fast accumulation of layers

¹Key Laboratory of Western China's Environmental Systems (Ministry of Education), College of Earth and Environmental Sciences, Lanzhou University, Lanzhou 730000, China.

²Key Laboratory of Alpine Ecology (LAE), CAS Center for Excellence in Tibetan Plateau Earth Sciences and Institute of Tibetan Plateau Research, Chinese Academy of Sciences (CAS), Beijing 100101, China. ³Frontier Center for Eco-environment and Climate Change in Pan-third Pole Regions, Lanzhou University, Lanzhou 730000, China. ⁴Centre for Archaeological Science, School of Earth, Atmospheric and Life Sciences, University of Wollongong, Wollongong, New South Wales 2522, Australia. ⁵Australian Research Council (ARC) Centre of Excellence for Australian Biodiversity and Heritage, University of Wollongong, Wollongong, New South Wales 2522, Australia. ⁶Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig 04103, Germany. ⁷Key Laboratory of Vertebrate Evolution and Human Origins of Chinese Academy of Sciences, Institute of Vertebrate Paleontology and Paleoanthropology, CAS, Beijing 100044, China. ⁸Center for Excellence in Life and Palaeoenvironment, Chinese Academy of Sciences, Beijing 100044, China. ⁹School of Earth Sciences, Lanzhou University, Lanzhou 730000, China. ¹⁰Department of Biology, University of Richmond, Richmond, VA 23173, USA. ¹¹Department of Anthropology, University of Nevada–Reno, Reno, NV 89557, USA. ¹²School of Human Evolution and Social Change, Arizona State University, Tempe, AZ 85281, USA. ¹³Institute of Human Origins, Arizona State University, Tempe, AZ 85281, USA.

*Corresponding author. Email: djzhang@lzu.edu.cn (D.Z.); bli@uow.edu.au (B.L.); paabo@eva.mpg.de (S.P.); fuqiaomei@ivpp.ac.cn (Q.F.)

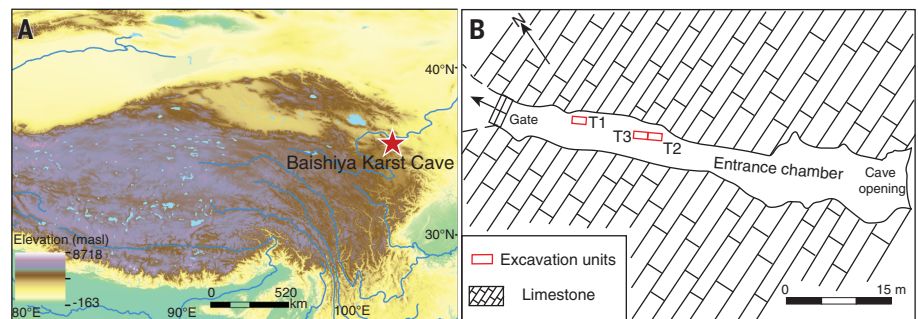


Fig. 1. Location of Baishiya Karst Cave on the Tibetan Plateau. (A) Regional map showing the location of the site. masl, meters above sea level. (B) Plan view of the entrance chamber and locations of excavation units (T1, T2, and T3). The gate separates the entrance chamber from other chambers farther inside the cave.

9 to 6 until 96 ± 5 ka. No age was obtained for layer 5. We modeled a time interval with a duration of 24 to 39 ka for layer 5. The sedimentary features of layer 5 (table S1) are indicative of a fluvial environment within BKC and may represent an erosional event that removed deposits dated to between ~ 100 and ~ 60 ka. Layer 4 was deposited from 66 ± 6 to 47 ± 2 ka. A depositional hiatus with a duration of ~ 7 to 18 ka was identified in the middle of layer 4 between ~ 60 and 50 ka, suggesting that sediments in layer 4 may have been deposited in two broad pulses. Layer 3 accumulated from 46 ± 2 ka 33 ± 1 ka, followed by layer 2 until 17 ± 12 ka. Layers 2 and 3 are more complex; radiocarbon ages vary considerably within a layer (table S3), and single-grain D_e values are broadly distributed (figs. S17 to S19 and S26).

To test whether ancient DNA was preserved in the cave, we extracted DNA (15) from eight sediment samples (100 to 250 mg each) collected from the middle of each layer (except layers 1 and 5) (Fig. 2A, fig. S3, and table S18). Aliquots of each extract were converted to DNA libraries and enriched for mammalian and human mitochondrial DNA (mtDNA) using probes for 242 mammalian mtDNAs (15) and for human mtDNA (16). For each library, the number of DNA fragments sequenced ranged from 0.07 to 1.7 million. From these, we obtained between 10 and 27,150 unique fragments mapping to mammalian mitochondrial genomes. All sampled layers, except layers 8 and 9, contained mammalian mtDNA. In layers 4, 6, 7, and 10, we detected DNA from animal species that have not been present in the area since ~ 10 ka, including extinct hyenas and rhinoceros (17), species for which bones were also identified in layer 10 (figs. S12 and S27). For all mammal sequences from all libraries corresponding to layers 2, 3, 4, 6, and 7 and four libraries from layer 10, the frequency of apparent terminal substitutions of cytosine (C) to thymine (T) at the 5' end ranges from 10 to 65% (table S18), which is typical for ancient DNA. These results confirm that ancient DNA is preserved in the cave.

We then assessed whether ancient hominin DNA was present in each library by determining the frequency of apparent C \rightarrow T substitutions in all hominin mtDNA fragments. The libraries from layers 2, 3, 4, and 7 have between 15.6 and 50% C \rightarrow T terminal substitutions, indicating the presence of ancient hominin DNA (table S19). We thus prepared additional DNA extracts from layers 2, 3, 4, and 7 (table S18). To determine which hominin groups may have contributed mtDNA to these samples, we examined sequences for substitutions found to be specific to modern humans, Neanderthals, Denisovans, and a $\sim 430,000$ -year-old hominin individual from Sima de los Huesos (Sima) (18) in phylogenetic analyses of mtDNAs as described (14, 19).

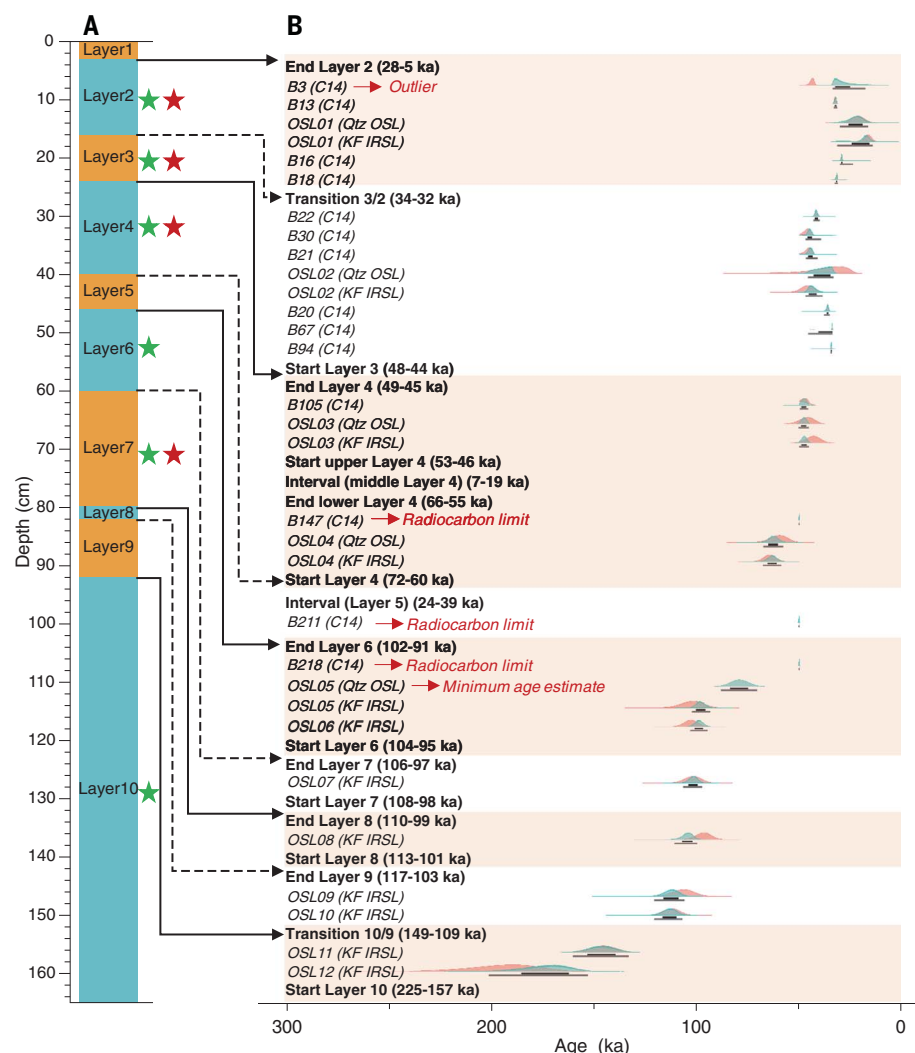


Fig. 2. Stratigraphy and dating results of T2. (A) Composite schematic stratigraphy of excavation area T2. The alternating colors are for illustration purposes only. The positions of sedimentary DNA samples from which Denisovan and animal DNA were found are shown as red stars and green stars, respectively. (B) Bayesian modeling results for all radiocarbon and optical ages. Red probability distributions represent the unmodeled ages (likelihoods), and green distributions represent the modeled ages (posterior probabilities). The narrow and wide bars beneath each distribution represent the 68.2 and 95.4% probability ranges of the modeled ages. Modeled ages for each layer boundary and phase and duration for each interval are given in parentheses (95.4% probability, random only errors).

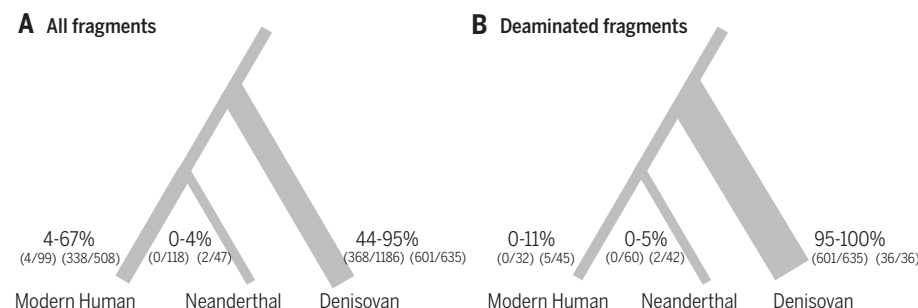


Fig. 3. Lineage inferences for layer 4. Using modern human, Neanderthal, and Denisovan branch-specific substitutions for all fragments (A) and deaminated fragments (B) from layer 4 for lineage inferences. Ranges for the percentage of lineage-matching sites for all libraries from layer 4 are given. The fractions give the absolute number of sequenced fragments carrying derived lineage-specific alleles over the total number of fragments covering positions where such alleles occur. The lineage inferences for layers 2, 3, and 7 are in fig. S28.

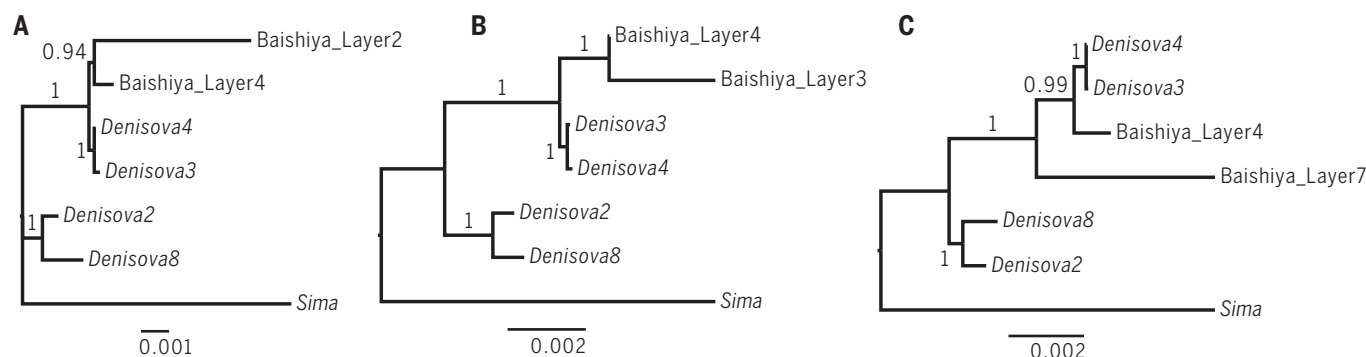


Fig. 4. mtDNA phylogenetic trees. mtDNA phylogenetic trees for sediment samples from layers 4 and 2 (**A**), layers 4 and 3 (**B**), or layers 4 and 7 (**C**) of the BKC as well as mtDNA from four Denisovans from Denisova Cave and a ~430,000-year-old hominin from Sima de los Huesos in Spain. Consensus sequences with deaminated fragments were used for BKC samples, except for layer 2 mtDNA, which is from all fragments from low contamination libraries and deaminated fragments from potentially contaminated libraries (see “decision” column in table S11). The phylogeny was estimated with a Bayesian approach under a general time reversible substitution model with a gamma distributed correction of sequence evolution.

From the 24 libraries, between 31% (368/1186) and 95% (601/635) of the mtDNA fragments that covered informative positions matched the Denisovan state (Fig. 3 and fig. S28), whereas 0 to 14% (1/7) matched the Neanderthal state (Fig. 3 and fig. S28), 0 to 3.7% (5/135) the Sima state (table S19), and 0 to 67% (338/508) the modern human state (Fig. 3 and fig. S28). Restricting the analysis to DNA fragments with first and last three C→T substitutions indicating cytosine deamination (Fig. 3 and fig. S28) decreased the proportion of fragments matching the modern human state to 0 to 43% (3/7) and increased the proportion matching the Denisovan state to 71 to 100%. To reduce the influence of modern human contamination, we restricted subsequent analyses to deaminated mtDNA fragments and excluded two DNA libraries in which modern human mtDNA fragments were slightly deaminated, albeit much less so than Denisovan mtDNA fragments (14).

By merging deaminated hominin mtDNA fragments from libraries for each layer, we arrive at an average mtDNA coverage for layers 2, 3, 4, and 7 of 0.37-fold, 1.5-fold, 40-fold, and 1.3-fold, respectively. DNA recovered from sediments may be derived from multiple different individuals, and this is the case at least in layer 4, where we have sufficient information to estimate the number of mtDNA fragments present (14). However, to gauge the average relationships of mtDNA in each layer, we called a consensus mtDNA sequence for each layer using positions covered by at least two different DNA fragments, excluding positions covered by only two fragments and where they differ. We also required that more than two-thirds of the fragments covering each position must carry an identical base at positions covered by more than two fragments (18). These sequences covered 7, 36, 99, and 26% of the mtDNA, respectively (table S15).

We then estimated phylogenetic trees using previously published mtDNA sequences from four Denisovans from Denisova Cave (Denisova 2, Denisova 3, Denisova 4, and Denisova 8) and the individual from Sima de los Huesos. The composite consensus mtDNA from layer 4 that is of comparatively high quality falls within the mtDNA variation of Denisovans, forming a clade with Denisova 3 and 4 to the exclusion of Denisova 2 and 8 (Fig. 4). When the consensus mtDNA sequences that are of lower quality are analyzed separately (Fig. 4), the mtDNA sequences from layers 2 and 3 form a clade with the layer 4 mtDNA, whereas the consensus layer 7 mtDNA diverges earlier from the lineage leading to Denisova 3 and 4. Thus, the mtDNA sequences from BKC form a clade (100% posterior support) with the mtDNA sequences for Denisova 3 and 4 (20, 21). The depositional age for the lower part of layer 4 (~60 ka) (Fig. 2) is comparable to the date of Denisova 3 (76 to 52 ka) and Denisova 4 (84 to 55 ka) (20, 21). Besides, the depositional age for layer 7 (108 to 97 ka) (Fig. 2) is older than those for Denisova 3 and 4 but younger than the ages for Denisova 2 (194 to 122 ka) and Denisova 8 (136 to 105 ka) (20, 21). Although Denisovan mtDNA is present in layers 3 and 2, it is tenuous to associate them to their corresponding depositional ages (~30 to 50 ka), given the reworked nature of the layers. Therefore, whether the BKC Denisovans had survived until the arrival of modern humans on the Tibetan Plateau by 30 to 40 ka (22) remains an open question.

In conclusion, the stratigraphic, chronological, and sedimentary DNA results presented show that Denisovans occupied BKC at ~100 and ~60 ka. This confirms that Denisovans were widely distributed in Asia during the Late Pleistocene. Together with the older Xiahe mandible, this evidence suggests a long occupation of the Tibetan Plateau by individuals who may have become adapted to the high-

altitude environment. The genetic adaptations to high altitudes seen in modern Tibetans could be associated with a haplotype introgressed from Denisovans (23) that perhaps evolved during the extended occupation of this high-altitude environment by Denisovans. Deeper investigations at BKC and other Paleolithic sites in a broader region surrounding the Tibetan Plateau may help to understand the relationship and evolution of Denisovans, modern humans, and possible other archaic humans in East Asia.

REFERENCES AND NOTES

1. J. Krause et al., *Nature* **464**, 894–897 (2010).
2. D. Reich et al., *Nature* **468**, 1053–1060 (2010).
3. M. Meyer et al., *Science* **338**, 222–226 (2012).
4. K. Prüfer et al., *Nature* **505**, 43–49 (2014).
5. D. Reich et al., *Am. J. Hum. Genet.* **89**, 516–528 (2011).
6. P. Qin, M. Stoneking, *Mol. Biol. Evol.* **32**, 2665–2674 (2015).
7. S. Sankararaman, S. Mallick, N. Patterson, D. Reich, *Curr. Biol.* **26**, 1241–1247 (2016).
8. B. Vernot et al., *Science* **352**, 235–239 (2016).
9. S. R. Browning, B. L. Browning, Y. Zhou, S. Tucci, J. M. Akey, *Cell* **173**, 53–61.e9 (2018).
10. S. Sawyer et al., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15696–15700 (2015).
11. V. Slon et al., *Sci. Adv.* **3**, e1700186 (2017).
12. B. Viola et al., *Am. J. Phys. Anthropol.* **168**, 258 (2019).
13. F. Chen et al., *Nature* **569**, 409–412 (2019).
14. Materials and methods are available as supplementary materials.
15. V. Slon et al., *Science* **356**, 605–608 (2017).
16. Q. Fu et al., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2223–2227 (2013).
17. G. Zong, W. Chen, X. Huang, Q. Xu, *Cenozoic Mammals and Environment of Hengduan Mountains Region* (China Ocean Press, 1996).
18. M. Meyer et al., *Nature* **505**, 403–406 (2014).
19. M. Meyer et al., *Nature* **531**, 504–507 (2016).
20. K. Douka et al., *Nature* **565**, 640–644 (2019).
21. Z. Jacobs et al., *Nature* **565**, 594–599 (2019).
22. X. L. Zhang et al., *Science* **362**, 1049–1051 (2018).
23. E. Huerta-Sánchez et al., *Nature* **512**, 194–197 (2014).
24. X. Xu et al., *Nucleic Acids Res.* **46**, D14–D20 (2018).

ACKNOWLEDGMENTS

We thank G. Dong, Z. Wang, J. Brantingham, and D. Rhode for participating in early investigation of BKC; Z. Li and D. Lin for help in the interpretation of sedimentary data; M. Qiu for taking photos of the stone artifacts and bones; and S. Pang, Z. Jiang, and Z. Jia for

measuring the cave. We are grateful to the support of National Cultural Heritage Administration of China, Provincial Cultural Heritage Administration of Gansu, Gansu Provincial Archaeological Institute, and local government for our archaeological excavation in BKC. We are also grateful to the support from the Baishiya Temple and the local residents in Ganjia town. **Funding:** This study was funded by the Second Tibetan Plateau Scientific Expedition and Research Program (STEP) (2019QZKK0601) and the Strategic Priority Research Program (XDB26000000, XDA20040000) of CAS; NSFC (41771225) to D.Z.; NSFC (91731303, 41925009, 41672021, 41630102), Tencent Foundation through the EXPLORER PRIZE, and "Research on the roots of Chinese civilization" of Zhengzhou University (XKZDJC202006) to Q.F.; The Strategic Innovation Fund of the Max Planck Society to S.P.; and Australian Research Council Future Fellowships to B.L. (FT140100384) and Z.J. (FT150100138). **Author contributions:** D.Z. and F.C. designed the study. D.Z.,

F.C., H.X., T.C., X.S., J.W., J. Yao, J. Yang, D.B.M., C.P., Y.H., and X.C. carried out field investigation, excavated the site, and carried out sampling processes. D.Z., F.C., H.X., B.L., D.B.M., and C.P. conducted stratigraphic and taphonomy analysis. B.L., Z.J., H.X., and T.C. performed the OSL and radiocarbon dating. Q.F., R.Y., Q.D., X.F., P.C., W.P., and F.L. performed the ancient DNA experiments and analysis. S.P., M.M., J.K., V.S., and D.M. helped for discussing and interpreting the genetic data. D.Z., Q.F., S.P., B.L., M.A.Y., Z.J., and F.C. wrote the paper with contributions from all authors. **Competing interests:** The authors declare no competing interests; **Data and materials availability:** All relevant data are available in the main text or the accompanying supplementary materials. The new mitochondrial consensus files reported in this paper have been deposited in the Genome Warehouse in National Genomics Data Center (24), Beijing Institute of Genomics (China National Center for Bioinformation), Chinese Academy of Sciences, under

accession number PRJCA002765, which is publicly accessible at <https://bigd.big.ac.cn/gwh>. Artifacts and animal fossils referred to in this study are curated in Lanzhou University.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/584/suppl/DC1
Materials and Methods
Figs. S1 to S29
Tables S1 to S19
References (25–127)
MDAR Reproducibility Checklist

[View/request a protocol for this paper from Bio-protocol.](#)

27 April 2020; accepted 10 September 2020
10.1126/science.abb6320

MAGNETISM

Metal-organic magnets with large coercivity and ordering temperatures up to 242°C

Panagiota Perlepe^{1,2}, Itziar Oyarzabal^{1,3*}, Aaron Mailman⁴, Morgane Yquel^{1,2}, Mikhail Platonov^{5,†}, Iurii Dovgaliuk^{6,‡}, Mathieu Rouzières¹, Philippe Négrier⁷, Denise Mondieig⁷, Elizaveta A. Suturina⁸, Marie-Anne Dourges⁹, Sébastien Bonhommeau⁹, Rebecca A. Musgrave¹, Kasper S. Pedersen^{1,10}, Dmitry Chernyshov⁶, Fabrice Wilhelm⁵, Andrei Rogalev⁵, Corine Mathonière², Rodolphe Clérac^{1*}

Magnets derived from inorganic materials (e.g., oxides, rare-earth-based, and intermetallic compounds) are key components of modern technological applications. Despite considerable success in a broad range of applications, these inorganic magnets suffer several drawbacks, including energetically expensive fabrication, limited availability of certain constituent elements, high density, and poor scope for chemical tunability. A promising design strategy for next-generation magnets relies on the versatile coordination chemistry of abundant metal ions and inexpensive organic ligands. Following this approach, we report the general, simple, and efficient synthesis of lightweight, molecule-based magnets by postsynthetic reduction of preassembled coordination networks that incorporate chromium metal ions and pyrazine building blocks. The resulting metal-organic ferrimagnets feature critical temperatures up to 242°C and a 7500-oersted room-temperature coercivity.

Magnets that operate at room temperature are usually pure metals, metal oxides, or intermetallic compounds, and they have applications in numerous aspects of our daily lives. For example, magnets are key components in data processing and storage devices, are commonly used in electrical motors that power most household appliances, and are essential in renewable energy technologies (1). Despite their extensive use and tremendous success in technological applications, conventional magnets present several drawbacks, such as energetically expensive fabrication (e.g., for SmCo and AlNiCo) and limited availability of key component elements (e.g., in the rare-earth-based magnets NdFeB and SmCo). Over the last three decades, various approaches have been developed to address these limitations and to target next-generation magnets. One particularly appealing strategy relies on the

rational assembly of molecular building blocks, such as organic ligands and paramagnetic metal ions. These molecule-based materials exhibit behavior similar to that of traditional magnets; however, unlike the exclusively inorganic examples, they benefit from the synthetic and postsynthetic versatility that results from the molecular and coordination chemistries, which allow precise tailoring and optimization of their properties (2–4). This synthetic approach has already led to a vast number of systems with peculiar magnetic behaviors, several of which have no counterpart in inorganic materials. Among these molecule-based magnets are discrete high-spin molecules known as single-molecule magnets (SMMs) (5, 6), one-dimensional (1D) magnets (single-chain magnets) (7), and 2D and 3D networks exhibiting magnetically ordered phases (8). By separating magnetic metal ions with organic ligands, these molecule-based materials feature remarkably low densities (~1 g cm⁻³) compared with those of exclusively inorganic materials (generally >5 g cm⁻³). Although state-of-the-art inorganic magnets are indispensable because of their high maximum energy product [i.e., high magnetic density (9)], complementary molecule-based magnetic materials will be of great relevance to emergent magnetoelectronic, magnetic sensing, and recording technologies as a result of their low density. However, most of these molecule-based materials suffer from low operating temperatures, which has precluded technological application.

To raise the operating temperature of molecule-based magnets, closed-shell ligands have been replaced by radicals to link paramagnetic metal ions in 2D or 3D coordination networks (8). The presence of a radical's spin leads to particularly strong magnetic interactions

with the metal centers, which can be controlled by the chemical identity of the organic radical and metal ion, and the overlap of their magnetic orbitals containing an unpaired electron (8, 10, 11). This methodology is exemplified by the pioneering work of J. S. Miller on a family of magnets incorporating paramagnetic metal ions and organic radical species such as the tetracyanoethylene radical ([TCNE]^{•-}) (12). In these systems, the strong magnetic coupling between spins localized in the metal 3d orbitals and those of the radicals result in magnetically ordered phases with critical temperatures (T_C) as high as 400 K ($V[TCNE]_x$ where $x \sim 2$) (13). Apart from displaying the current record T_C value measured for a molecule-based magnet, $V[TCNE]_x$ has also shed light on the applicability of molecule-based, lightweight magnets in spintronic devices, quantum information, and microwave electronics (14–17). More recently, a synthetic strategy has been developed that subjects preassembled metal-organic coordination networks to postsynthetic oxidation or reduction (acting on the ligands to form radicals or acting on the metal ions to induce mixed-valency) to obtain magnetically ordered materials at a higher temperature (maximum of 105 K until this work) than their precursors (18–21). These combined efforts have resulted in magnets with hysteresis effects on the field ($\mu_0 H$) dependence of the magnetization (M) at ambient temperature in a small number of systems, such as TCNE-based compounds and derivatives as well as several Prussian blue analogs (2, 3, 12, 22, 23) and covalently linked organic radical frameworks (24, 25). However, all of these molecule-based magnets have failed so far to exhibit substantial room-temperature coercivity, which in the best case is on the order of hundreds of oersteds.

Herein, we report on the postsynthetic chemical reduction of two 2D coordination networks— $CrCl_2(py)_2$ and $Cr(OSO_2CH_3)_2(py)_2$ (py = pyrazine) (26, 27)—to enhance magnetic interactions and thus increase the critical temperature of any resulting ferrimagnetic order. Although structurally similar, these two materials [$CrX_2(py)_2$, where X is either $[CH_3SO_3]^-$ or Cl^-] exhibit contrasting physical properties. In $Cr(OSO_2CH_3)_2(py)_2$, the octahedral Cr^{II} metal ions are bridged by neutral, closed-shell pyrazine ligands (py)⁰, which transmit only weak magnetic interactions between $S = 2$ Cr^{II} spins. As a result, this material is an antiferromagnet below 10 K and an insulator (27). By contrast, $CrCl_2(py)_2$ features octahedral Cr^{III} metal ions and a mixed-valence pair of pyrazine ligands (i.e., $Cr^{III}Cl_2[(py)_2]^{•+}$). This electronic configuration generates strong magnetic interactions between the $S = 3/2$ Cr^{III} and delocalized $S = 1/2$ pyrazine spins, which leads to ferrimagnetic ordering below 55 K and a substantial room-temperature electrical conductivity (26). In this work, we describe the postsynthetic reduction of these coordination

¹Université de Bordeaux, CNRS, Centre de Recherche Paul Pascal, UMR 5031, F-33600 Pessac, France. ²Université de Bordeaux, CNRS, Bordeaux INP, ICMCB, UMR 5026, F-33600 Pessac, France. ³Chemistry Faculty, University of the Basque Country, UPV/EHU, 20018 Donostia-San Sebastián, Spain. ⁴Department of Chemistry, University of Jyväskylä, FI-40014 Jyväskylä, Finland. ⁵ESRF-The European Synchrotron, CS 40220, F-38043 Grenoble Cedex 9, France. ⁶Swiss-Norwegian Beamlines at the European Synchrotron Radiation Facility, F-38000 Grenoble, France. ⁷Université de Bordeaux, CNRS, Laboratoire Ondes et Matière d'Aquitaine, UMR 5798, F-33400 Talence, France. ⁸Department of Chemistry, University of Bath, Claverton Down, Bath BA2 7AY, UK. ⁹Université de Bordeaux, CNRS, Bordeaux INP, ISM, UMR 5255, F-33400 Talence, France. ¹⁰Department of Chemistry, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark. *Corresponding author. Email: itziar.oyarzabal@ehu.eus (I.O.); clerac@crpp-bordeaux.cnrs.fr (R.C.)

[†]Present address: Kirensky Institute of Physics, Federal Research Center KSC SB RAS, 660036 Krasnoyarsk, Russia.

[‡]Present address: Institut des Matériaux Poreux de Paris, UMR 8004 CNRS, Ecole Normale Supérieure, Ecole Supérieure de Physique et de Chimie Industrielles de Paris, PSL Université, 75005 Paris, France.

networks, resulting in lightweight ferrimagnets with T_C up to 515 K and 7500-Oe room-temperature coercivity.

The chemical reduction of the methanesulfonate-paired 2D material was carried out through the addition of two molar equivalents of lithium 1,2-dihydroacenaphthylene ($\text{Li}^+[\text{C}_{12}\text{H}_{10}^-]$; $E_{1/2} = -3.23$ V versus $[(\text{C}_5\text{H}_5)_2\text{Fe}]^{+/0}$) (19) to a suspension of $\text{Cr}^{\text{II}}(\text{OSO}_2\text{CH}_3)_2(\text{pyz})_2$ in tetrahydrofuran (THF) [Fig. 1A; see (28) for the detailed synthetic procedure]. A notable color change was evident upon reduction of the light brown $\text{Cr}(\text{OSO}_2\text{CH}_3)_2(\text{pyz})_2$ to afford product **1** as an air-sensitive, dark gray, and microcrystalline powder. The powder x-ray diffraction (PXRD)

pattern of **1** (fig. S1) (28, 29) revealed a highly crystalline phase, with Bragg diffraction peaks corresponding exclusively to $\text{Li}[\text{SO}_3\text{CH}_3]$, which indicates that the reduced metal-organic product was either poorly crystalline, nanocrystalline, or amorphous. However, the presence of THF-insoluble $\text{Li}[\text{SO}_3\text{CH}_3]$ implied that the methanesulfonate anions were partially or fully extracted from the 2D precursor upon reduction to yield a Cr-based network possibly featuring two reduced pyrazines ($[\text{pyz}]^{\cdot-}$). X-ray absorption spectroscopy (XAS) measurements at the Cr K-edge were performed on product **1**, the $\text{Cr}(\text{OSO}_2\text{CH}_3)_2(\text{pyz})_2$ precursor containing Cr^{II} in an octahedral coordination

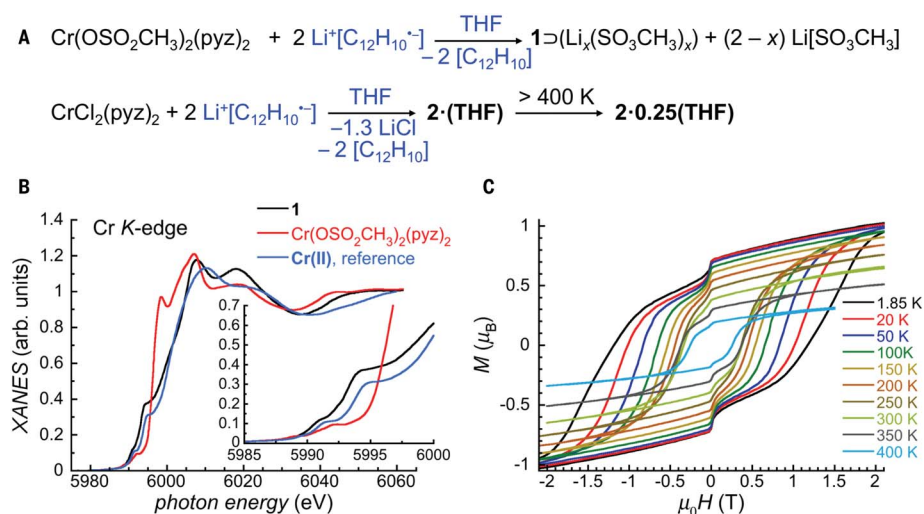


Fig. 1. Chemical reduction of $\text{Cr}(\text{OSO}_2\text{CH}_3)_2(\text{pyz})_2$ and $\text{CrCl}_2(\text{pyz})_2$. (A) Reaction schemes [solvents and soluble species in blue; solid materials in black; the \rightarrow symbol indicates the presence of $\text{Li}_x(\text{SO}_3\text{CH}_3)_x$ within **1**]. (B) Normalized XANES spectra at the Cr K-edge region of **1** (black trace), $\text{Cr}(\text{OSO}_2\text{CH}_3)_2(\text{pyz})_2$ (red trace), and $\text{Cr}(\text{II})$ reference (blue trace) at 295 K. Inset shows a magnified view of the near-edge region. (C) Magnetization versus applied dc magnetic field data (at 7 Oe s^{-1}) in the -2.1 to 2.1 T field range for **1** between 1.85 and 400 K. As the exact composition of **1** is unknown, the magnetization data were normalized using the molecular weight of the $\text{Cr}(\text{OSO}_2\text{CH}_3)_2(\text{pyz})_2$ parent compound.

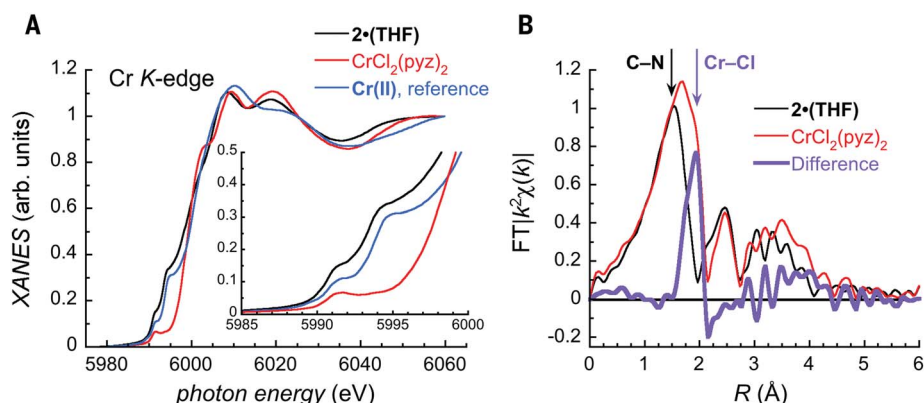


Fig. 2. XANES and EXAFS spectra at the Cr K-edge at 295 K. (A) Normalized XANES spectra of **2**·(THF) (black trace), $\text{CrCl}_2(\text{pyz})_2$ (red trace), and $\text{Cr}(\text{II})$ reference (blue trace). Inset shows a magnified view of the near-edge region. (B) Fourier-transform (FT) EXAFS spectra for **2**·(THF) (black trace) and $\text{CrCl}_2(\text{pyz})_2$ (red trace) (see figs. S9 and S10 for additional EXAFS data) (28). The difference between the two datasets is shown in purple.

sphere, and a square-planar Cr^{II} reference complex $\text{Cr}(\text{N}(\text{TMS})_2)_2(\text{py})_2$ [noted **Cr(II)**; $\text{TMS} = \text{Si}(\text{CH}_3)_3$, $\text{py} = \text{pyridine}$] (30). The x-ray absorption near-edge structure (XANES) spectrum of **1** is markedly different from that of its precursor, as well as that of Cr oxides and Cr metal (31), but its features at low energy (near-edge region) and at the rising edge are notably similar to those of the square-planar **Cr(II)** reference (Fig. 1B). The near-edge structures for both product **1** and **Cr(II)**, which are the fingerprint of the Cr oxidation state in a given ligand field, show two shoulders at the same energies (~ 5991 and ~ 5994 eV). These XAS results unequivocally support that the Cr electronic structure and coordination geometry in **1** are essentially the same as in **Cr(II)**; i.e., a high-spin $S = 2$ Cr^{II} metal ion in a square-planar $\{\text{Cr}^{\text{II}}\text{N}_4\}$ environment (30). The structure of **1** is thus compatible with a 2D square coordination network, $\text{Cr}^{\text{II}}(\text{pyz}^{\cdot-})_2$, reminiscent of the precursor $\text{Cr}^{\text{II}}(\text{pyz}^0)_2$ layer (27). The field dependency of the magnetization was recorded for powder **1** at various temperatures (Fig. 1C and figs. S2 to S4) (28), revealing broad M versus $\mu_0 H$ hysteresis loops ($\mu_0 H_{\text{coer}} = 3400$ Oe at 300 K) up to at least 400 K [the temperature limit of the magnetic properties measurement system (MPMS)]. This notable magnetic behavior is in sharp contrast to that reported for any known chromium-based materials (e.g., metal, nanoparticles, or oxides; table S1) (32–38).

The above experimental evidence unequivocally confirms that the postsynthetic chemical reduction of $\text{Cr}^{\text{II}}(\text{OSO}_2\text{CH}_3)_2(\text{pyz})_2$ resulted in a mixture of insoluble crystalline $\text{Li}[\text{SO}_3\text{CH}_3]$ and an amorphous $\text{Cr}^{\text{II}}(\text{pyz}^{\cdot-})_2$ phase, which displayed hard magnet properties ($\mu_0 H_{\text{coer}} = 3400$ Oe at 300 K) and a critical temperature above 400 K. To target a pure analog of this room-temperature magnet, our focus turned to the related $\text{CrCl}_2(\text{pyz})_2$ system, which is expected to yield the same magnetic material upon reduction of both the pyrazine scaffold $[(\text{pyz})_2]^{2+} \rightarrow (\text{pyz}^{\cdot-})_2$ and the Cr^{III} metal ion ($\text{Cr}^{\text{III}} \rightarrow \text{Cr}^{\text{II}}$). Notably, the anticipated by-product of this reaction, LiCl (Fig. 1A), should be far easier to remove from the Cr-based product than $\text{Li}[\text{SO}_3\text{CH}_3]$ because of its increased solubility in organic media, particularly THF. Using identical experimental conditions to those for the synthesis of **1**, the $\text{CrCl}_2(\text{pyz})_2$ precursor was exposed to $\text{Li}^+[\text{C}_{12}\text{H}_{10}^-]$ (Fig. 1A) (28). Synchrotron PXRD experiments on the resulting dark gray solid (90% isolated yield) revealed several prominent diffraction peaks and, notably, the absence of an independent crystalline LiCl phase (fig. S5) (28). The diffractogram was refined in the orthorhombic $Pmmm$ space group with the following cell parameters: $a = 6.9239(9)$, $b = 6.9524(2)$, and $c = 8.478(2)$ Å [$V = 408.1(1)$ Å³; the number between parentheses is the estimated standard deviation]. It is worth noting that the a and b

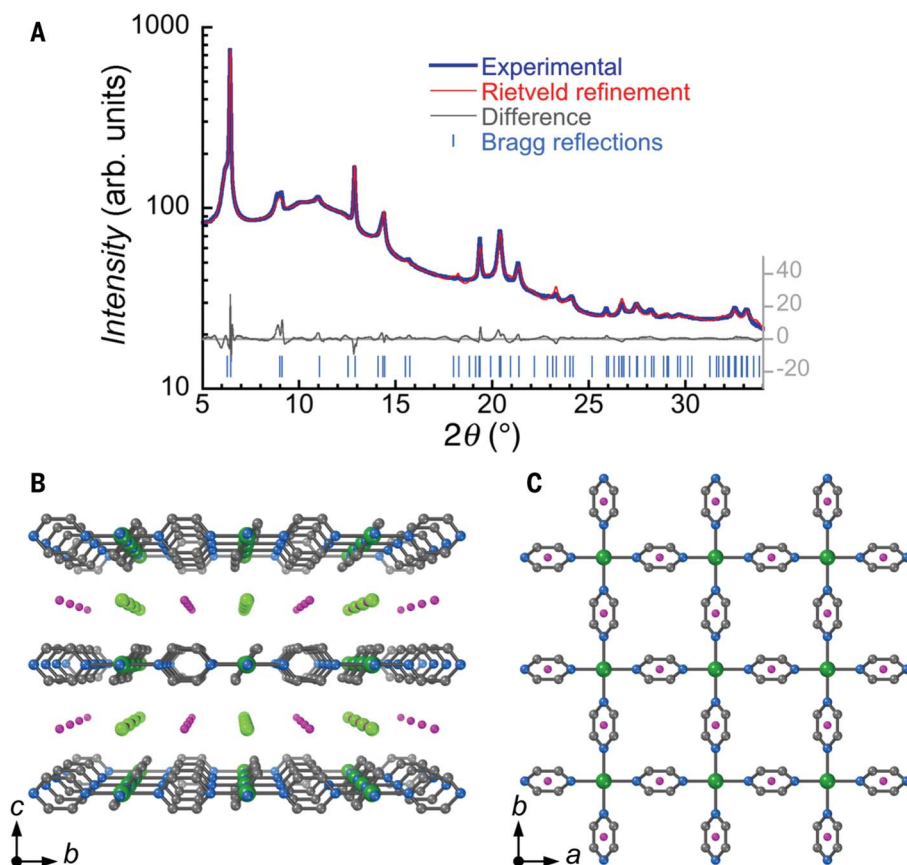


Fig. 3. Structural model of $\text{Li}_{0.7}[\text{Cr}(\text{pyz})_2]\text{Cl}_{0.7}\cdot 0.25(\text{THF})$ [2·0.25(THF)**].** (A) The best Rietveld refinement (red trace; $R_1 = 4.23\%$) of the synchrotron PXRD pattern of **2·0.25(THF)** at 290 K (after cooling from 500 K in a sealed capillary; blue trace) is shown with the experimental-model difference (gray trace) and calculated Bragg reflections (blue bars). (B) Perspective view (along the a direction) of **2·0.25(THF)** showing the alternation of $\text{Li}_{0.7}\text{Cl}_{0.7}$ and neutral $\text{Cr}^{\text{II}}(\text{pyz}^-)_2$ layers stacking along the c direction. (C) Eclipsed layered structure viewed along the c direction. Cr is shown in dark green, N in blue, and C in dark gray. Cl (light green) and Li (pink) are shown at a fixed occupancy of 70% according to the elemental composition. Hydrogen atoms have been omitted for clarity.

lattice parameters obtained for this material [referred to as **2·(THF)** herein] are close to those found in the $\text{CrCl}_2(\text{pyz})_2$ precursor [$Immm$ space group: $a = 6.90351(4)$, $b = 6.97713(4)$, and $c = 10.82548(6)$ Å; $V = 521.427(6)$ Å³; $\rho = 1.803$ g cm⁻³]. This suggests the presence of a similar 2D $\text{Cr}(\text{pyz})_2$ -type network in both $\text{CrCl}_2(\text{pyz})_2$ and **2·(THF)**.

The Cr K-edge XANES spectra of **2·(THF)** is remarkably similar to that of both the model square-planar complex **Cr(II)** and **1** (Fig. 1B), whereas it is clearly different from that of the precursor $\text{Cr}^{\text{III}}\text{Cl}_2(\text{pyz})_2$ (Fig. 2A). Therefore, we conclude that the reduction of $\text{CrCl}_2(\text{pyz})_2$ results in the reduction of octahedral high-spin Cr^{III} metal ions into square-planar high-spin Cr^{II} sites, which implies the loss of both axial chloride ligands in **2·(THF)**. Raman spectroscopy further supports this conclusion, as evidenced by the absence of the characteristic Cr–Cl symmetric stretching band (~ 260 cm⁻¹), whereas the position and narrow nature of Raman bands in the pyrazine

fingerprint region of the spectrum (600 to 1700 cm⁻¹) suggest the presence of only reduced $[\text{pyz}]^-$ ligands (figs. S6 and S7 and tables S2 to S4) (28). To understand the reduction mechanism and the concomitant chloride decoordination, quantum chemical geometry optimizations were performed on $[\text{CrCl}_x(\text{pyz})_4]^{2-x-q}$ molecular fragments (where the number of chloride ligands, $x = 2, 1$, or 0; and successive reductions are represented by $q = 0, 1, 2, 3$, or 4; table S5) (28). Upon the addition of four electrons to the initial $[\text{CrCl}_2(\text{pyz})_4]^0$ fragment, the calculations show a minimum energy at infinite Cr···Cl distance, which indicates the instability of the fully reduced bis-chloride-bound fragment, $[\text{CrCl}_2(\text{pyz})_4]^{4-}$. Upon the removal of chloride anions, the ability of the $[\text{CrCl}(\text{pyz})_4]^+$ and $[\text{Cr}(\text{pyz})_4]^{2+}$ fragments to stabilize reduced pyrazine ligands increases around the Cr^{II} center (fig. S8) (28). The pyrazine reduction is thus facilitated by the chloride dissociation. Moreover, the optimized structure of the $[\text{Cr}(\text{pyz})_4]^{2-}$

fragment ($q = 4$, pyrazines are all reduced around the Cr^{II} center) adopts a perfect square-planar geometry (table S5) (28), which makes this moiety ideal for forming extended $\text{Cr}(\text{pyz})_2$ -type sheets.

A comparison of the extended x-ray absorption fine structure (EXAFS) at the Cr K-edge of **2·(THF)** and its precursor provides further experimental evidence for chloride decoordination. As shown in Fig. 2B, the Fourier-transform EXAFS spectra of $\text{CrCl}_2(\text{pyz})_2$ and **2·(THF)** are similar except at $R \sim 1.9$ Å, where a marked difference is evident (see figs. S9 and S10 for EXAFS and k^2 -weighted EXAFS spectra) (28). This corresponds to a major modification in the local environment of the Cr^{II} site. In the precursor compound, $\text{CrCl}_2(\text{pyz})_2$, the broad feature at $R = 1.67$ Å and the shoulder at 1.91 Å can be attributed to the Cr–N (2.003 to 2.059 Å) and Cr–Cl (2.337 Å) bonds, respectively. However, although the Cr–N bond in **2·(THF)** is found at $R = 1.53$ Å, the signature of the Cr–Cl bond around $R \sim 1.9$ Å is pronouncedly attenuated (see difference spectrum in Fig. 2B), which corroborates the loss of axial Cl^- ions and the resulting square planar geometry at the Cr center.

Combustion elemental analysis (EA) and inductively coupled plasma optical emission spectroscopy (ICP-OES) measurements support the above assumptions and reveal the presence of 0.7(1) Li ion and 0.99(6) THF molecule per $\text{Cr}(\text{pyz})_2$ moiety, each assumed to reside between the $\text{Cr}^{\text{II}}(\text{pyz}^-)_2$ layers in **2·(THF)** (table S6) (28). Additionally, XANES measurements at the Cl K-edge for $\text{CrCl}_2(\text{pyz})_2$ and **2·(THF)** provide evidence for remaining chlorine anions (~ 0.7 per Cr) in the reduced material (fig. S11) (28). The near-edge feature, which corresponds to an electron excitation from the Cl 1s orbital to molecular orbitals of hybridized Cl 3p and Cr 3d orbitals (26), is much lower in intensity in the spectrum of **2·(THF)** versus its precursor. This feature reflects a considerable weakening of the chromium-chloride interaction and subsequent elongation of the Cr···Cl distance (i.e., decoordination from the Cr metal ion) (39) in **2·(THF)**, which corroborates the theoretical XANES calculations (fig. S12) (28) and experimental EXAFS data (Fig. 2B and figs. S9 and S10) (28). However, the presence of a detectable near-edge signal indicates that chloride ions are still close enough to the Cr^{II} ions to interact electronically (see a comparison with LiCl in fig. S13 in which Li^+ and Cl^- ions are ionically independent) (28). To summarize, **2·(THF)** is a material that has neutral 2D $\text{Cr}^{\text{II}}(\text{pyz}^-)_2$ layers of square planar Cr^{II} metal ions and two singly reduced pyrazines, separated by one THF molecule and 0.7(1) equivalents of Li^+ and Cl^- ions. The chemical formula of **2·(THF)** can therefore be defined as $\text{Li}_{0.7}[\text{Cr}(\text{pyz})_2]\text{Cl}_{0.7}(\text{THF})$, with a calculated density of 1.278 g cm⁻³.

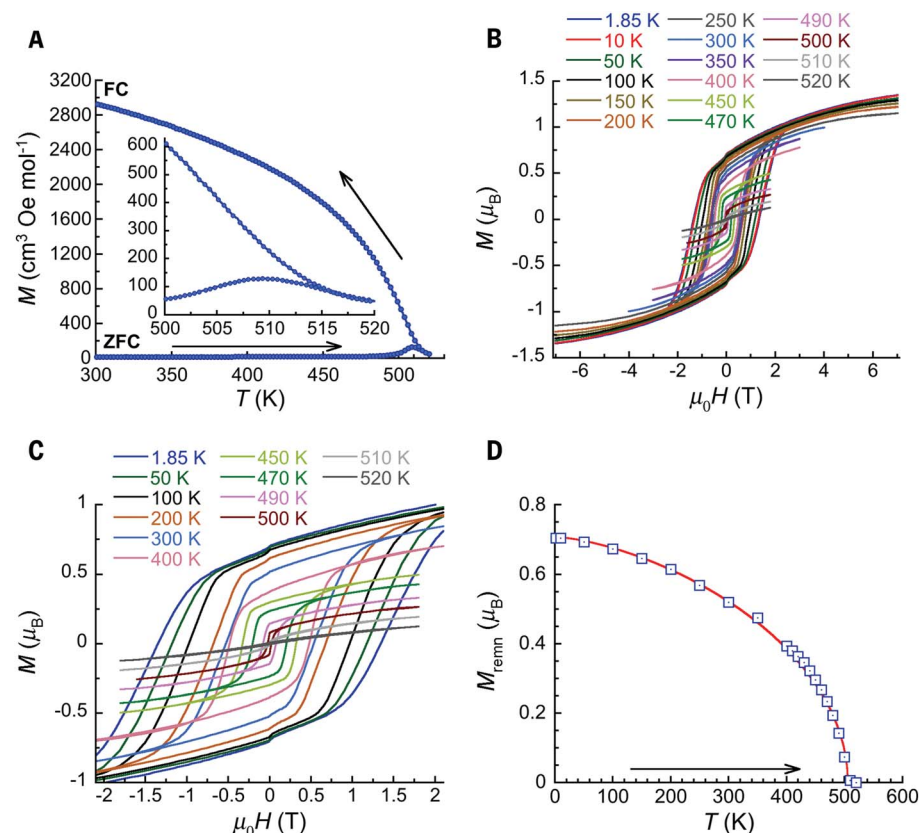


Fig. 4. Magnetic properties of 2-(THF) and 2-0.25(THF) (after partial desolvation above 400 K).

(A) Zero field-cooled (ZFC) and field-cooled (FC) magnetization data obtained under an applied dc magnetic field of 50 Oe at 5 K min^{-1} . Inset shows a magnified view of the main plot in the 500 to 520 K temperature range. The solid lines are a guide for the eye. (B) Magnetization versus applied dc magnetic field data (at 5 to 12 Oe s^{-1}) in the -7 to 7 T field range, from 1.85 to 520 K (42). (C) Magnified view of selected data from (B) in the -2.1 to 2.1 T field range (42). (D) Temperature dependence of the remnant magnetization, M_{remn} , determined from the M versus $\mu_0 H$ data between 1.85 and 520 K [(B) and (C)]. The solid red line is the best fit to the mean-field (MF) Bloch law: $M_{\text{remn}} \propto [1 - (T/T_{\text{C-MF}})^{3/2}]^{1/2}$ with $T_{\text{C-MF}} = 506 \text{ K}$ (considering data up to 490 K, $\chi^2(\text{GoF}) = 0.00057$) (43).

Variable-temperature PXRD measurements on **2-(THF)** revealed an irreversible phase transformation upon heating above 380 K (fig. S14) (28). This phase is stable up to 500 K and remains the sole phase upon cooling to room temperature. As evidenced by thermogravimetric analysis-mass spectrometry (TGA-MS) measurements (figs. S15 to S20) (28), this irreversible structural change is caused by a loss of the interlayer THF molecules. However, Fourier transform infrared (FTIR) spectroscopy (figs. S21 to S24) (28) and EA measurements on samples of **2-(THF)** after prolonged heating (18 hours at 400 K or 10 hours at 500 K) show a residual presence of ~ 0.25 THF per formula unit, which is in good agreement with TGA-MS measurements (fig. S17 and table S6) (28). XANES measurements at the Cr K-edge of this heat-treatment product, **2-0.25(THF)**, indicated no appreciable change in the Cr^{II} coordination environment upon partial THF loss (fig. S25) (28). On the other hand, the Cr

K-edge spectrum showed an attenuation in the near-edge region versus **2-(THF)** (fig. S26) (28), which indicates reduced mixing between Cr 3p and Cr 3d orbitals—i.e., greater localization of Cr p orbitals. The chlorine anions are thus slightly further from the Cr ions after the partial loss of the interlayer THF molecules. The diffractogram of **2-0.25(THF)** is less affected than that of **2-(THF)** by the anisotropic broadening of the diffraction peaks systematically associated with planes that have a nonzero l Miller index (Fig. 3A and figs. S14 and S27) (28). For both compounds, this observation implies less crystallographic order along the c direction (and thus in the spacing between the sheets) than within the ab plane of the 2D network, as is expected for layered materials (40, 41). The diffractogram of **2-0.25(THF)** at 290 K was fully indexed in the tetragonal $P4/mmm$ space group with $a = b = 6.9893(1)$ and $c = 7.195(3) \text{ \AA}$ [$V = 351.5(1) \text{ \AA}^3$; $\rho = 1.228 \text{ g cm}^{-3}$; Fig. 3A and table S7] (28),

showing that, upon removing the THF molecules, the interlayer distance diminishes, as evidenced by the decrease of the c parameter from $\sim 8.5 \text{ \AA}$ in **2-(THF)** to $\sim 7.2 \text{ \AA}$ in **2-0.25(THF)** [for comparison the interlayer distance in the $\text{CrCl}_2(\text{pyz})_2$ precursor is 5.4 \AA].

Considering all of the above information, an appropriate structural model was used to refine the experimental PXRD pattern obtained for **2-0.25(THF)**, which yielded reasonable agreement factors (e.g., $R_1 = 4.23\%$, $R_p = 1.63\%$, and $R_{\text{wp}} = 2.32\%$ at 290 K; Fig. 3 and table S7) (28). As expected, the Cr sites in **2-0.25(THF)** are bridged by pyrazine ligands forming a 2D square $\text{Cr}(\text{pyz})_2$ network in the crystallographic ab plane. In contrast to the precursors $\text{CrCl}_2(\text{pyz})_2$ and $\text{Cr}(\text{OSO}_2\text{CH}_3)_2(\text{pyz})_2$ (26, 27), these $\text{Cr}(\text{pyz})_2$ layers in **2-0.25(THF)** are eclipsed along the c direction (Fig. 3C). The square-planar coordination sphere of Cr is occupied by four nitrogen atoms from four pyz ligands, with a Cr–N distance of $2.0440(6) \text{ \AA}$. As evidenced by other techniques on **2-THF** and **2-0.25(THF)** (vide supra), the Cl^- anions are not coordinated to the Cr metal ions, but are located in between Cr sites of two adjacent layers with a $\text{Cr}\cdots\text{Cl}$ distance of $3.598(2) \text{ \AA}$. In this structural model, it was not possible to localize THF molecules, but the Li cations are most likely positioned in between pyrazines for electrostatic reasons and slightly better refinement of the PXRD data.

The magnetic properties of **2-(THF)** were studied and compared with those collected for **1** (figs. S2 to S4) (28). Zero-field cooled (ZFC) and field-cooled (FC) magnetization data were collected under a dc field of 50 Oe for a sample of **2-(THF)** [and thus for **2-0.25(THF)** after partial desolvation from $\sim 400 \text{ K}$ onward] on a MicroSense vibrating sample magnetometer (VSM) capable of reaching temperatures up to 600 K. The magnetization bifurcation point of the ZFC and FC data, which corresponds to the temperature at which the coercive field vanishes, is found at $\sim 510 \text{ K}$ (Fig. 4A). This temperature agrees well with the M versus $\mu_0 H$ curves, which show a crossover between S-shape and linear (typical of a paramagnetic state) forms between 510 and 520 K (Fig. 4, B and C). As also confirmed by the vanishing of the remnant magnetization at 510 K (Fig. 4D), the critical temperature of **2-0.25(THF)** is thus $\sim 510 \text{ K}$, which exceeds the ordering temperature measured for $[\text{V}(\text{TCNE})_x]$ by $\sim 110 \text{ K}$ (13). It is worth emphasizing that the linear dependence of the magnetization at 520 K after the disappearance of the M versus $\mu_0 H$ hysteresis loops rules out the presence of any superparamagnetic nanoparticles. On cooling from 520 K to room temperature, the M versus $\mu_0 H$ hysteresis loops are recovered as expected for a phase transition between paramagnetic and ferri- or ferromagnetic states. Nevertheless, the absolute value of the magnetization is systematically lower than that during

the heating process (figs. S28 and S29) (28), which suggests that the critical temperature of **2-0.25(THF)** is in close proximity to its decomposition temperature under the experimental conditions of the magnetic measurements (i.e., in a sealed quartz tube under inert atmosphere). Similar to the powder **1** (Fig. 1C and figs. S2 to S4) (28), **2-(THF)** is a hard magnet with a remarkably large coercive field of 5300 Oe at room temperature (13500 Oe at 1.85 K; Fig. 4, B and C, and fig. S30) (28). This value compares well with those of widely used inorganic magnets and is larger than any of those observed for molecule-based magnets (which are on the order of hundreds of oersteds in the best cases; table S8) (28). The coercive field reproducibly displays an anomalous increase at ~350 K on the initial heating of **2-(THF)** (fig. S30) (28) and **1** (fig. S4) (28) samples, whereas no such feature is observed in the remnant magnetization, shown in Fig. 4D and figs. S29 and S3 (28). This anomaly of the coercive field, which is absent in measurements on **2-0.25(THF)** [prepared by annealing **2-(THF)**; figs. S30 to S35] (28), is thus likely linked to the partial loss of interlayer THF between ~300 and 400 K (vide supra) and the associated irreversible structural rearrangement (fig. S14) (28). The magnetic properties (magnetization and x-ray spectroscopy measurements; figs. S36 and S37) of **2-(THF)** and **2-0.25(THF)** samples are similar, with only slight variation in T_C (510 versus 515 K, respectively), but there is a noticeable 50% difference in the coercivity at 300 K (5300 versus 7500 Oe, respectively; Fig. 4C and figs. S32 and S37) (28). These results show that the annealing of **2-(THF)** to remove most of the THF molecules improves the T_C and $\mu_0 H_{\text{coer}}$ characteristics of the resulting magnet, **2-0.25(THF)**.

At 1.85 K, the magnetization of both **2-(THF)** and **2-0.25(THF)** does not saturate at 7 T, but reproducibly reaches a maximum value of 1.34 μ_B (Fig. 4B and fig. S31) (28). This value is systematically lower than that of 2 μ_B predicted for an ordered ferrimagnetic state with antiferromagnetically coupled spins (one $S = 2 \text{ Cr}^{\text{II}}$ center and two $S = \frac{1}{2}$ pyrazine radicals; assuming g factors of 2). In an analogous manner to $\text{CrCl}_2(\text{pyz})_2$ (26), the low magnetization may originate from the high degree of conjugation between the Cr d orbitals and pyrazine π orbitals, which leads to a partial delocalization of the Cr spin density over the organic scaffold and strong Cr-radical antiferromagnetic interactions. As evidenced by single-point broken-symmetry calculations performed on a $[\text{Cr}(\text{pyz})_4]^{2-}$ fragment taken from the **2-0.25(THF)** structural model (Fig. 3, B and C), the Cr-radical exchange coupling is strongly antiferromagnetic and ranges from -336 to -427 cm^{-1} (-483 to -614 K with the $-2J$ convention) depending on the tilting of the radical pyrazines (tables S9 and S10 and

figs. S38 and S39) (28). These strong antiferromagnetic interactions are in agreement with the ferrimagnetic order experimentally observed at high temperature for these materials.

This work reports molecule-based metal-organic magnets with high critical temperatures up to 515 K and large, room-temperature coercivity, which compete well with the characteristics of the traditional inorganic magnets and surpass the properties of previously known molecule-based magnets (table S8) (28). We demonstrate here that the postsynthetic chemical reduction of coordination networks is a general, simple, and efficient synthetic approach that offers broad perspectives for the preparation of a new generation of high-temperature, lightweight magnets, with yet unrealized application in emergent technologies. Finally, it should be highlighted that the reduced materials reported in this study are electrically insulating (as expected because of the absence of mixed-valency), whereas the mixed-valence $[\text{Cr}^{\text{III}}\text{Cl}_2(\text{pyz})_2]^{+}$ precursor shows a substantial room-temperature conductivity of 32 mS cm^{-1} (26). Therefore, notable potential exists in fine-tuning the postsynthetic reduction of these metal-organic materials, which will lead to the further development of new high- T_C conducting molecule-based magnets.

REFERENCES AND NOTES

- O. Gutfleisch et al., *Adv. Mater.* **23**, 821–842 (2011).
- J. S. Miller, *Mater. Today* **17**, 224–235 (2014).
- J. S. Miller, *Chem. Soc. Rev.* **40**, 3266–3296 (2011).
- O. Kahn, *Molecular Magnetism* (VCH-Verlag, 1993).
- D. Gatteschi, R. Sessoli, J. Villain, *Molecular Nanomagnets* (Oxford Univ. Press, 2006).
- S. Demir, I.-R. Jeon, J. R. Long, T. D. Harris, *Coord. Chem. Rev.* **289–290**, 149–176 (2015).
- C. Coulon, V. Pianet, M. Urdampilleta, R. Clérac, in *Molecular Nanomagnets and Related Phenomena*, S. Gao Ed. (Springer, 2014), pp. 143–184.
- A. E. Thorarinnsson, T. D. Harris, *Chem. Rev.* **120**, 8716–8789 (2020).
- J. M. D. Coey, *Magnetism and Magnetic Materials* (Cambridge Univ. Press, 2010).
- N. Motokawa, H. Miyasaka, M. Yamashita, K. R. Dunbar, *Angew. Chem. Int. Ed.* **47**, 7760–7763 (2008).
- X. Ma et al., *Angew. Chem. Int. Ed.* **57**, 7841–7845 (2018).
- J. S. Miller, S.-I. Ohkoshi, in *Molecular Magnetic Materials: Concepts and Applications*, B. Sieklucka, D. Pinkowicz, Eds. (Wiley-VCH, 2017), pp. 161–186.
- J. M. Manriquez, G. T. Yee, R. S. McLean, A. J. Epstein, J. S. Miller, *Science* **252**, 1415–1417 (1991).
- H. Liu et al., *Nat. Mater.* **17**, 308–312 (2018).
- N. Zhu et al., *Appl. Phys. Lett.* **109**, 082402 (2016).
- M. Chilcote et al., *APL Mater.* **7**, 111108 (2019).
- A. Franson et al., *APL Mater.* **7**, 121113 (2019).
- K. Taniguchi, J. Chen, Y. Sekine, H. Miyasaka, *Chem. Mater.* **29**, 10053–10059 (2017).
- L. Liu, J. A. DeGayer, L. Sun, D. Z. Zee, T. D. Harris, *Chem. Sci.* **10**, 4652–4661 (2019).
- J. A. DeGayer, I.-R. Jeon, L. Sun, M. Dincă, T. D. Harris, *J. Am. Chem. Soc.* **139**, 4175–4184 (2017).
- L. E. Darago, M. L. Aubrey, C. J. Yu, M. I. Gonzalez, J. R. Long, *J. Am. Chem. Soc.* **137**, 15703–15711 (2015).
- J. S. Miller, *Polyhedron* **28**, 1596–1605 (2009).
- S. Ferlay, T. Mallah, R. Ouahès, P. Veillet, M. Verdaguer, *Nature* **378**, 701–703 (1995).
- J. Mahmood et al., *Chem* **4**, 2357–2369 (2018).
- H. Phan et al., *Chem* **5**, 1223–1234 (2019).
- K. S. Pedersen et al., *Nat. Chem.* **10**, 1056–1061 (2018).
- P. Perlepe et al., *Polyhedron* **153**, 248–253 (2018).

28. See the supplementary materials.

- T. Trella, W. Frank, poster presented at the 17th Conference of the GDCh Division of Solid State Chemistry and Materials Research, Dresden, Germany, 15 to 17 September 2014.
- Y.-F. Deng et al., *Chem. Commun.* **51**, 17688–17691 (2015).
- S. Cuervo et al., *J. Anal. At. Spectrom.* **31**, 1818–1829 (2016).
- E. Fawcett, *Rev. Mod. Phys.* **60**, 209–283 (1988).
- C. G. Shull, M. K. Wilkinson, *Rev. Mod. Phys.* **25**, 100–107 (1953).
- G. E. Bacon, N. Cowlam, *J. Phys. C: Solid State Phys.* **2**, 238–251 (1969).
- T. Furubayashi, I. Nakatani, *J. Appl. Phys.* **73**, 6412–6413 (1993).
- W. Abdul-Razzaq, M. S. Seehra, *Phys. Status Solidi* **193**, 94–102 (2002).
- S. Foner, *Phys. Rev.* **130**, 183–197 (1963).
- J. M. D. Coey, M. Venkatesan, *J. Appl. Phys.* **91**, 8345–8350 (2002).
- S. D. George, P. Brant, E. I. Solomon, *J. Am. Chem. Soc.* **127**, 667–674 (2005).
- A. B. Cairns, A. L. Goodwin, *Chem. Soc. Rev.* **42**, 4881–4893 (2013).
- T. N. Ramesh, R. S. Jayashree, P. V. Kamath, *Clays Clay Miner.* **51**, 570–576 (2003).
- P. Perlepe et al., Magnetization versus applied dc magnetic field data for $\text{Li}_0.2[\text{Cr}(\text{pyz})_2]\text{Cl}_{0.7}(\text{THF})$, version 1, Zenodo (2020).
- R. Skomski, in *Novel Functional Magnetic Materials*, A. Zhukov Ed. (Springer Series in Materials Science, Springer, 2016), pp. 359–395.

ACKNOWLEDGMENTS

The authors thank the GdR MCM-2 (Magnétisme et Commutation Moléculaires) and the MOLSPIN COST action CA15128. A. Väisänen (JYU, Jyväskylä, Finland), E. Hautakangas (JYU, Jyväskylä, Finland), P. Voisin (ESRF, Grenoble, France), J. M. Seco (UPV/EHU, Donostia-San Sebastián, Spain), G. Le Bourdon (ISM, Talence, France), L. Vellutini (ISM, Talence, France), L. Voigt (DTU, Lyngby, Denmark), Y.-G. Li (Northeast Normal University, Changchun, China; who we thank for the hydrothermal reactors), P. Dechambenoit (CRPP, Pessac, France), X. Ma (CRPP, Pessac, France), and S. De (CRPP, Pessac, France) are thanked for helpful discussions, comments, and experimental assistance. We would like to dedicate this paper to the loving memory of our colleague Dr. Ángela Valentín Pérez. **Funding:** This work was supported by the University of Bordeaux, the Région Nouvelle Aquitaine, Quantum Matter Bordeaux, and the Centre National de la Recherche Scientifique (CNRS). I.O. and R.C. are grateful to the Basque Government for I.O.'s postdoctoral grant. K.S.P. thanks the VILLUM FONDEN for a Villum Young Investigator grant (15374). A.M. thanks JYU and the Academy of Finland (project 289172) for support. **Author contributions:** R.C. conceived the original idea and formulated the research aims. P.P. executed the syntheses. A.R., F.W., M.P., P.P., I.O., M.Y., R.C., and C.M. performed the x-ray absorption spectroscopy experiments and analyzed the results. Crystallographic work was carried out by I.D., D.C., P.N., and D.M. Raman and infrared spectra were collected by S.B., A.M., I.O., and M.Y. Elemental analyses and inductively coupled plasma measurements were carried out by A.M., K.S.P., and P.P. Thermogravimetric analysis–mass spectrometry data were collected and analyzed by M.-A.D., P.P., and I.O. Magnetic measurements were conducted by P.P., M.R., R.C., and C.M. E.A.S. performed the density functional theory studies. After a first complete draft of the manuscript was written by I.O., P.P., R.A.M., C.M., and R.C., all authors were involved in finalizing the manuscript and gave their consent to its publication. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Crystallographic data of $\text{Li}_{0.2}[\text{Cr}(\text{pyrazine})_2]\text{Cl}_{0.7} \cdot 0.25(\text{THF})$ [**2-0.25(THF)**] at 290 and 500 K are available free of charge from the Cambridge Crystallographic Data Centre under reference numbers 1983877 and 2007863, respectively. The M versus $\mu_0 H$ data shown in Fig. 4, B and C, are publicly available on Zenodo (42). All other data are in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/587/suppl/DC1
Materials and Methods
Figs. S1 to S39
Tables S1 to S10
References (44–76)

26 June 2020; accepted 11 September 2020
10.1126/science.abb3861

QUANTUM SYSTEMS

Parallel single-shot measurement and coherent control of solid-state spins below the diffraction limit

Songtao Chen*, Mouktik Raha*, Christopher M. Phenicie, Salim Ourari, Jeff D. Thompson†

Solid-state spin defects are a promising platform for quantum science and technology. The realization of larger-scale quantum systems with solid-state defects will require high-fidelity control over multiple defects with nanoscale separations, with strong spin-spin interactions for multi-qubit logic operations and the creation of entangled states. We demonstrate an optical frequency-domain multiplexing technique, allowing high-fidelity initialization and single-shot spin measurement of six rare-earth (Er^{3+}) ions, within the subwavelength volume of a single, silicon photonic crystal cavity. We also demonstrate subwavelength control over coherent spin rotations by using an optical AC Stark shift. Our approach may be scaled to large numbers of ions with arbitrarily small separation and is a step toward realizing strongly interacting atomic defect ensembles with applications to quantum information processing and fundamental studies of many-body dynamics.

A central appeal of solid-state atomic defects for quantum technology is the possibility of realizing strong dipolar interactions between closely spaced spins (1). This coupling enables multi-qubit logic operations [to realize, for example, error correction (2, 3) or deterministic teleportation over a quantum network (4)] as well as fundamental studies of many-body quantum phenomena (5). Typically, these interactions are appreciable for defect separations less than several tens of nanometers. However, for optically addressed spins, it is an open challenge to achieve simultaneous, high-fidelity initialization, control, and readout of spins separated by less than the diffraction limit of the addressing light, typically several hundred nanometers. Several techniques have been demonstrated to simultaneously address pairs of closely spaced nitrogen vacancy centers, such as super-resolution microscopy (6), and variations in the Larmor frequency arising from different defect orientations (7, 8) or magnetic field gradients (9); however, these approaches have not been extended to high-fidelity operations such as single-shot spin readout, or to larger numbers of defects. Alternatively, an array of nuclear spins surrounding a single atomic defect can be distinguished by their positions in the gradient of the hyperfine coupling (10–12); although this approach has been used to generate entanglement between as many as 10 spins (13), it suffers from the bottleneck that all operations are performed through a single, central electron spin.

Rare-earth ions (REIs) in solid-state hosts are a promising platform for many applications because of their demonstrated long coherence times [for example, exceeding 6 hours for Eu^{3+} (14)] as well as operation in the tele-

com band and compatibility with silicon photonics [in the case of Er^{3+} (15)]. Furthermore, their distinctive spectral characteristics enable frequency-domain addressing of many defects within the same spatial volume. REIs experience random, static shifts of their optical transition frequencies that give rise to an inhomogeneous (ensemble) linewidth Γ_{inh} (typically 1 to 10 GHz in crystalline hosts) that is much broader than the homogeneous linewidth of an individual ion, Γ_h (typically <1 MHz). In a given sample volume, this allows a large number of distinct subsets of ions to be separately addressed, on the order of $N_{\text{ad}} \approx \Gamma_{\text{inh}}/\Gamma_h > 10^3$. This approach can be applied to any solid-state emitter, in principle, but the especially small magnitude of Γ_{inh} and Γ_h in REIs allows the entire inhomogeneous distribution to be addressed with electro-optic sidebands on a single laser. Using spectral hole burning, this effect has been exploited in rare-earth ensembles to realize multimode atomic memories for quantum networks (16, 17). Quantum gate architectures based on ensemble spectral-hole qubits have also been proposed (18–21) and demonstrated (22–24).

Frequency-domain addressing can also be used to address individual REIs within a diffraction-limited volume, if the total number of ions N is less than N_{ad} . Although detecting individual REIs is challenging (because of their low photon emission rates), this problem can be overcome by using Purcell enhancement in nanophotonic optical cavities (15, 25), as exemplified by recent demonstrations of single-shot spin readout of single REIs (26, 27). We combine frequency-domain addressing and high-fidelity optical control to realize initialization and single-shot spin readout of six Er^{3+} spins with submicrometer separations, coupled to a single photonic crystal cavity. Our experimental approach consists of an Er^{3+} -doped Y_2SiO_5 (YSO) crystal coupled to a silicon photonic crystal cavity (Fig. 1A). Among REIs, Er^{3+} is

attractive for its telecom-wavelength optical transition and potential application to quantum networks. The cavity enhances the emission rate of the ions (15) and modifies the selection rules to make the optical transitions highly cyclic, enabling single-shot spin readout (26). The zero-field photoluminescence excitation (PLE) spectrum (Fig. 1B) shows several hundred ions within the $0.05\text{-}\mu\text{m}^3$ mode volume of the optical cavity, with an inhomogeneous linewidth of several gigahertz. We first focus on a pair of ions located in the blue tail of the inhomogeneous distribution, labeled ion 1 and ion 2, which couple to the cavity with Purcell factors of 330 and 200, respectively, when resonant with the cavity. Because the ions are addressed through a single-mode cavity, the optical signal provides no spatial information about the ions: They are, by definition, within a single, diffraction-limited volume. Instead, the ions are addressed in the frequency domain, relying on a separation between their transitions of ~ 250 MHz, which is considerably larger than their linewidths (24 and 10 MHz) but smaller than the cavity linewidth of 4.2 GHz. In a magnetic field, each ion's optical transition splits into four lines that can be used to interface with its spin (Fig. 1, C and D).

First, we demonstrate simultaneous initialization and single-shot spin measurement of ion 1 and ion 2. The measurement relies on cavity-enhanced cyclicity of the optical transitions, which is controlled by the alignment of the magnetic field to the local cavity polarization (26). A magnetic field orientation of $[(\theta, \varphi) = (90, 150)^\circ]$ allows high cyclicity for both ions (Fig. 2A), indicating similar cavity polarization at their respective positions. We initialize the spins by optical pumping, driving the excited-state spin transition with microwaves to mix the spin levels (Fig. 2, B and C) (27, 28). Then, we perform a simultaneous single-shot spin measurement by alternately exciting the spin-conserving optical transitions (A, B) on each ion. For both initialization and measurement, the laser frequency is rapidly switched between transitions by using a sideband from a fiber-coupled electro-optic modulator. We infer an initialization fidelity of $\geq 95\%$, 97% (29) and an average readout fidelity of 76%, 88% (Fig. 2D) for ion 1 and ion 2, respectively. The ions' spins can be coherently manipulated by using microwave pulses that address both ions equally (Fig. 2E), because the disorder in the Larmor frequency is much smaller than that of the optical transition. Details about the spin lifetime and coherence times can be found in the supplementary materials (29).

Next, we turn to demonstrating individually addressed spin manipulations. To achieve this, we make use of the AC Stark shift from a detuned optical pulse to induce a net phase shift ϕ between $|\uparrow\rangle$ and $|\downarrow\rangle$ (30). The optical pulses

Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA

*These authors contributed equally to this work.

†Corresponding author. Email: jdthompson@princeton.edu

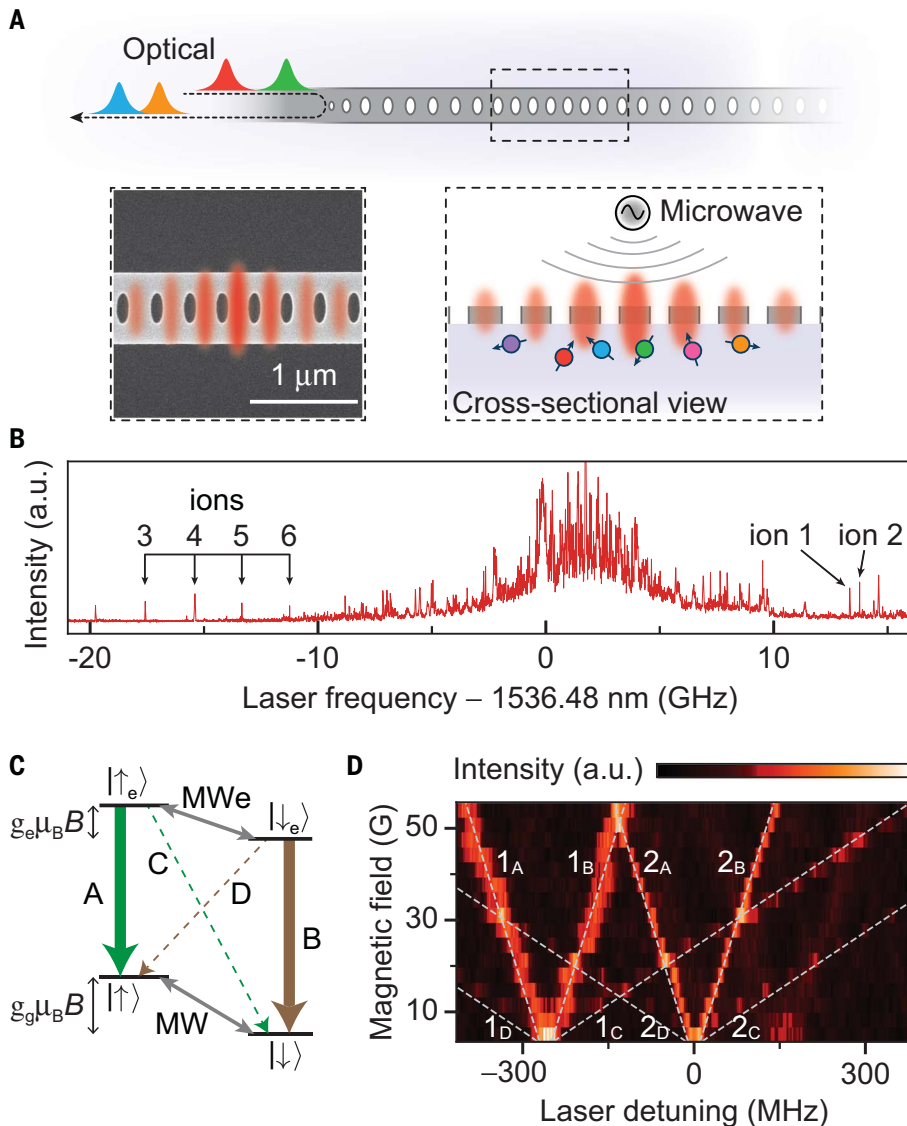


Fig. 1. Spectrally addressing multiple ions in a diffraction-limited volume. (A) Schematic of the device, showing multiple ions with different transition frequencies (colors) coupled to the cavity. (Inset) Scanning electron microscope image of a representative cavity, showing the extent of the optical mode. (B) PLE spectrum of Er^{3+} ions in a single device with magnetic field $B = 0$. Arrows indicate the six ions used in this work. (C) Level structure of Er^{3+} :YSO in a magnetic field, with optical (A to D) and microwave (MW, MWe) transitions indicated. g_g (g_e) denotes the ground- (excited-) state magnetic g -factor. (D) PLE spectrum of ion 1 and ion 2 in the presence of a magnetic field (oriented along the D_2 axis of the YSO crystal). ($1_A, 1_B$) and ($2_A, 2_B$) correspond to the spin-conserving optical transitions (A, B) of the two ions, respectively. Zero detuning in this panel and subsequent figures refers to the ion 2 resonance when $B = 0$. a.u., arbitrary units.

are inserted into an XY8 sequence that mitigates low-frequency magnetic field noise during the phase accumulation time. For each ion, the accumulated phase shift is $\phi = T\Omega^2(\Delta_B^{-1} - \Delta_A^{-1})/4$, where T is the pulse duration, Ω is the optical Rabi frequency, and Δ_A, Δ_B are the detunings of the laser from the spin-conserving transitions A, B (Fig. 3A, inset). For a given laser frequency and intensity, the detuning and Rabi frequency are different for each ion, enabling local control of the phase shift. To control N ions, $N - 1$ laser frequencies are

needed, because microwave rotations provide an additional control axis (29). Here, with $N = 2$, we can control both ions independently using a single laser frequency. In addition to the phase shift, there is also a loss of coherence from photon scattering and fluctuations in the optical transition frequency (e.g., from spectral diffusion), which happens at a rate $\Gamma' \propto \Gamma\Omega^2(\Delta_A^{-2} + \Delta_B^{-2})$, where Γ is the effective transition linewidth (29).

We measure the optically induced phase shift and decoherence using Ramsey spec-

troscopy, where the decoherence manifests as a change in visibility (Fig. 3A). To select the optimum operational point, we characterize the phase shift and decoherence as a function of laser frequency (Fig. 3B). The results are in good agreement with a theoretical model. The ratio of phase shift to decoherence is optimized for large detunings, and we predict a 2.55% loss of visibility per radian of differential phase shift $\Delta\phi = \phi_2 - \phi_1$ at 350 MHz detuning. In this regime, the experimental loss of visibility over a $\pi/2$ phase shift is below our measurement resolution.

In combination with a global microwave rotation $R_z(-\phi_1)$, the differential phase $\Delta\phi$ gives rise to a net rotation on ion 2 alone: $R_z^{(2)}(\Delta\phi) = \mathbb{I} \otimes R_z(\Delta\phi)$ (Fig. 3C). Similarly, a global microwave rotation $R_z(-\phi_2)$ generates a rotation on ion 1 alone. Here, $R_{\hat{n}}(\alpha)$ denotes a rotation by an angle α about axis \hat{n} . Universal control of a single qubit requires arbitrary angle rotations around two orthogonal axes. However, global microwave rotations can transform ion-selective optical z rotations into rotations around an arbitrary axis (29). As an example, we demonstrate rotations about the x axis, $R_x^{(i)}(\Delta\phi)$, where $i = 1, 2$ denotes the target ion (Fig. 3, C and D), realizing more than 2π rotation as the optical pulse duration is varied.

Lastly, we extend our approach to demonstrate simultaneous spin initialization and readout with four additional ions, labeled ion 3 to ion 6 (Fig. 1B). To access them, we shift the cavity resonance to -14.8 GHz (with respect to Fig. 1B), resulting in Purcell factors of 130, 260, 360, and 50. After choosing a magnetic field orientation that allows high cyclicity for all ions (29), we perform single-shot readout measurements. Because of the larger spread of these ions' frequencies (6.4 GHz) with respect to the cavity linewidth, it is advantageous to perform the readout using only one of the A or B transitions for each ion, whichever has larger Purcell enhancement (Fig. 4A). The average readout fidelities for each ion are 80, 74, 87, and 71%, respectively (Fig. 4B). The mean readout fidelity of the four-ion set is 78%, which is slightly lower than that of the two-ion set of 82%, because of the low Purcell factor of ion 6. Although the ions are measured sequentially, the total measurement duration (300 ms) is much shorter than the ground-state spin T_1 [typically > 10 s (29)], such that the measurements are effectively simultaneous.

In Fig. 4C, we show simultaneous microwave-driven Rabi oscillations on all four ions after initializing into $|\uparrow\uparrow\downarrow\downarrow\rangle$. Because ion 4 is situated in a crystallographic site rotated from that of the other ions, it has a different coupling to the microwave waveguide and correspondingly different Rabi frequency. In this measurement, the static field B lies in the $D_1 - D_2$ plane such that all ions have the same Larmor frequency, but we note that rotating B out of this plane

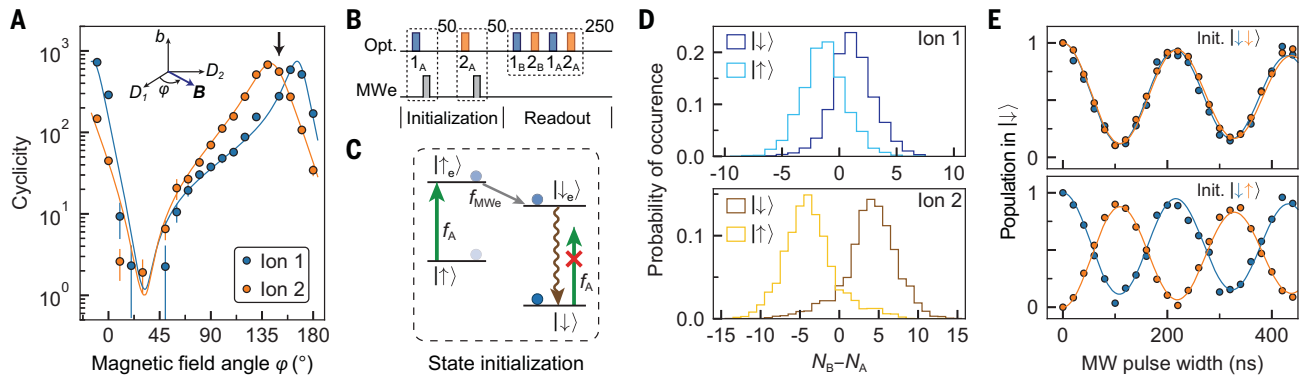


Fig. 2. Simultaneous initialization and readout of ions 1 and 2. (A) Cyclicity of the optical transitions (defined as $1 + \Gamma_{A,B}/\Gamma_{C,D}$, where Γ_i is the Purcell-enhanced decay rate on transition i) as a function of magnetic field angle (see inset). The solid lines are fits to a theoretical model from (26), and the black arrow indicates the orientation used in subsequent experiments ($\varphi = 150^\circ$). (B) Pulse scheme used for spin initialization (in this case, to $|\downarrow\rangle$) and readout. The exponents 50, 250 denote the number of repetitions of the sequence. All optical and microwave pulses are π pulses. (C) Diagram of the initialization sequence.

A dark state emerges in $|\downarrow\rangle$ from the combination of optical excitation of the spin-conserving A transition and microwave driving of the excited-state spin. Exciting B instead will initialize the spin to $|\uparrow\rangle$. (D) Photon histograms showing simultaneous single-shot readout for both ions. The ions are probed alternately on their A and B transitions, and $N_B - N_A$ denotes the difference in detected counts. $N_B > N_A$ indicates that the spin state is $|\downarrow\rangle$. (E) Simultaneous Rabi oscillations are observed while driving the ground-state spin of both ions with a microwave pulse after initialization into the indicated states. The vertical axes are corrected for initialization and readout fidelity.

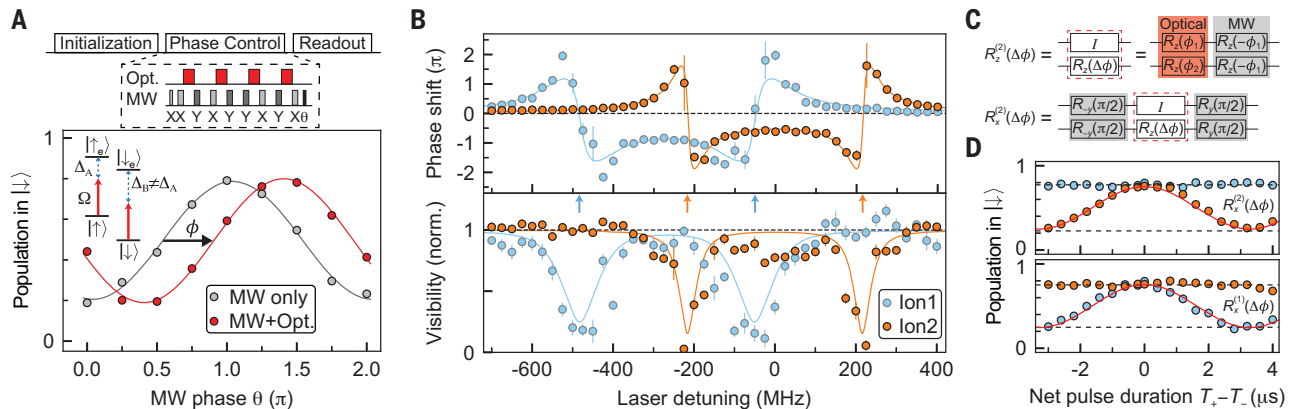


Fig. 3. Coherent optical spin rotation using the AC Stark shift. (A) (Upper) Optical phase shifts are generated by using a sequence of detuned optical pulses interleaved with a microwave-driven XY8 decoupling sequence. (Lower) The phase shift is detected by varying the phase of the final microwave $\pi/2$ pulse before measuring the spin population. The phase shift (ϕ) and visibility are extracted from a sinusoidal fit. (B) Frequency dependence of the phase shift and change in visibility for each ion after a 2- μ s optical pulse. The solid lines are numerical simulations.

including spectral diffusion. The arrows indicate the frequencies of the A, B transitions for each ion. (C) Circuit diagram for implementing $R_{z,x}^{(i)}(\Delta\phi)$ rotations. (D) Ion-selective Rabi oscillations $R_x^{(i)}(\Delta\phi)$. Positive (negative) phase shifts are generated by placing optical pulses after the odd- (even-) numbered π pulses in the XY8 sequence, with total duration T_+ (T_-) (the laser detuning is 275 MHz). The dashed lines show the loss of contrast from dephasing measured in the absence of any optical pulses; excess dephasing from the optical pulses is not observable.

would make the ion 4 Larmor frequency different, enabling spectral addressing of its microwave transition. In Fig. 4D, we show the initialization and single-shot measurement outcomes for all 16 four-spin states.

We have demonstrated simultaneous frequency-domain addressing of multiple Er^{3+} spins within a diffraction-limited volume, realizing a complete set of operations: initialization, coherent control, and single-shot spin measurement. This is a step toward two applications. First, combined with realistic improvements in optical and spin coherence enabling indistinguishable single-photon emission and long-term quantum memories [using host crystals with

higher site symmetry (27) and lower nuclear spin content (31)], this work may lead to frequency-multiplexed quantum repeaters with substantially faster entanglement generation rates. Second, this approach enables probing and manipulating single defects within dense, strongly interacting ensembles. The creation of these ensembles is within reach by using ion implantation: $N \approx 20$ ions implanted into a $(30 \text{ nm})^3$ volume would have a typical separation of 11 nm, corresponding to a magnetic dipolar interaction strength of 1 MHz (given the large magnetic moment in Er^{3+}), faster than the decoherence rate of the ground-state spin with dynamical decoupling (29).

REFERENCES AND NOTES

1. D. D. Awschalom, L. C. Bassett, A. S. Dzurak, E. L. Hu, J. R. Petta, *Science* **339**, 1174–1179 (2013).
2. T. H. Taminiau, J. Cramer, T. van der Sar, V. V. Dobrovitski, R. Hanson, *Nat. Nanotechnol.* **9**, 171–176 (2014).
3. G. Waldherr et al., *Nature* **506**, 204–207 (2014).
4. W. Pfaff et al., *Science* **345**, 532–535 (2014).
5. S. Choi et al., *Nature* **543**, 221–225 (2017).
6. P. C. Maurer et al., *Nat. Phys.* **6**, 912–918 (2010).
7. P. Neumann et al., *Nat. Phys.* **6**, 249–253 (2010).
8. F. Dolde et al., *Nat. Phys.* **9**, 139–143 (2013).
9. M. Grinolds et al., *Nat. Phys.* **7**, 687–692 (2011).
10. S. Kolkowitz, Q. P. Unterreithmeier, S. D. Bennett, M. D. Lukin, *Phys. Rev. Lett.* **109**, 137601 (2012).
11. T. H. Taminiau et al., *Phys. Rev. Lett.* **109**, 137602 (2012).
12. N. Zhao et al., *Nat. Nanotechnol.* **7**, 657–662 (2012).
13. C. Bradley et al., *Phys. Rev. X* **9**, 031045 (2019).
14. M. Zhong et al., *Nature* **517**, 177–180 (2015).

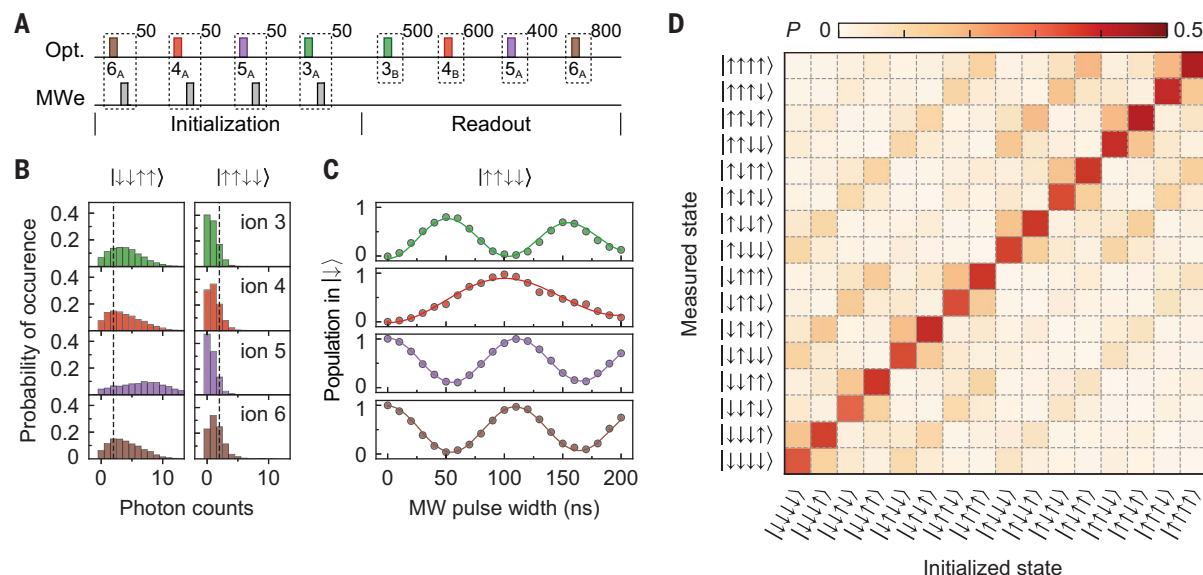


Fig. 4. Simultaneous initialization and readout of ions 3 to 6. (A) Pulse sequence for initialization and readout. (B) Histograms of detected photon counts for different initial states. The dashed lines indicate the threshold used for state discrimination. The spin state giving rise to higher counts

depends on which transition is used for readout, which differs between ions. (C) Simultaneous microwave Rabi oscillations of all ions after being initialized to $|\uparrow\uparrow\uparrow\downarrow\rangle$. (D) Single-shot readout results for all 16 four-ion initial states. The average probability of obtaining the correct state for all four ions is $P = 0.37$.

15. A. M. Dibos, M. Raha, C. M. Phenicie, J. D. Thompson, *Phys. Rev. Lett.* **120**, 243601 (2018).
16. H. de Riedmatten, M. Afzelius, M. U. Staudt, C. Simon, N. Gisin, *Nature* **456**, 773–777 (2008).
17. E. Saglamyurek *et al.*, *Nature* **469**, 512–515 (2011).
18. K. Ichimura, *Opt. Commun.* **196**, 119–125 (2001).
19. N. Ohlsson, R. Krishna Mohan, S. Kröll, *Opt. Commun.* **201**, 71–77 (2002).
20. J. H. Wesenberg, K. Mølmer, L. Rippe, S. Kröll, *Phys. Rev. A* **75**, 012304 (2007).
21. R. Ahlefeldt, M. Pearce, M. Hush, M. Sellars, *Phys. Rev. A* **101**, 012309 (2020).
22. G. J. Pryde, M. J. Sellars, N. B. Manson, *Phys. Rev. Lett.* **84**, 1152–1155 (2000).
23. J. Longdell, M. Sellars, *Phys. Rev. A* **69**, 032307 (2004).
24. J. J. Longdell, M. J. Sellars, N. B. Manson, *Phys. Rev. Lett.* **93**, 130503 (2004).
25. T. Zhong *et al.*, *Phys. Rev. Lett.* **121**, 183603 (2018).
26. M. Raha *et al.*, *Nat. Commun.* **11**, 1605 (2020).
27. J. M. Kindem *et al.*, *Nature* **580**, 201–204 (2020).
28. M. Afzelius *et al.*, *J. Lumin.* **130**, 1566–1571 (2010).

29. See the supplementary materials.
30. B. B. Buckley, G. D. Fuchs, L. C. Bassett, D. D. Awschalom, *Science* **330**, 1212–1215 (2010).
31. C. M. Phenicie *et al.*, *Nano Lett.* **19**, 8928–8933 (2019).
32. S. Chen, M. Raha, C. M. Phenicie, S. Ourari, J. D. Thompson, Replication Data for: Parallel single-shot measurement and coherent control of solid-state spins below the diffraction limit, Harvard Dataverse, V1 (2020); <https://doi.org/10.7910/DVN/5XMCMU>.

ACKNOWLEDGMENTS

We gratefully acknowledge conversations with N. de Leon and M. Tuna Uysal. **Funding:** Support for this research was provided by the National Science Foundation (NSF, EFRI ACQUIRE program grant 1640959), the Princeton Center for Complex Materials (PCCM), an NSF MRSEC (DMR-1420541), the Air Force Office of Scientific Research (grant FA9550-18-1-0081), and the DARPA DRINQS program (grant D18AC00015). We acknowledge the use of Princeton's Imaging and Analysis Center, which is partially supported by the PCCM, as well as the Princeton

Micro-Nano Fabrication Lab and Quantum Device Nanofabrication Lab facilities. C.M.P. was supported by the Department of Defense through the National Defense Science and Engineering Graduate Fellowship Program. **Author contributions:** S.C., M.R., C.M.P., S.O., and J.D.T. contributed to the design and execution of the experiment. S.C., M.R., and J.D.T. analyzed the data and wrote the manuscript. **Completing interests:** The authors declare no competing interests. **Data and materials availability:** All (other) data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials and are publicly available at Harvard Dataverse (32).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/592/suppl/DC1
Supplementary Text
Figs. S1 to S9
References (33, 34)

13 May 2020; accepted 11 September 2020
10.1126/science.abc7821

PROTON MEMBRANES

CdPS₃ nanosheets-based membrane with high proton conductivity enabled by Cd vacanciesXitang Qian^{1,2}, Long Chen¹, Lichang Yin¹, Zhibo Liu¹, Songfeng Pei¹, Fan Li^{3,4}, Guangjin Hou³, Shuangming Chen⁵, Li Song⁵, Khalid Hussain Thebo¹, Hui-Ming Cheng^{1,2,6}, Wencai Ren^{1,2*}

Proton transport in nanochannels under humid conditions is crucial for the application in energy storage and conversion. However, existing materials, including Nafion, suffer from limited conductivity of up to 0.2 siemens per centimeter. We report a class of membranes assembled with two-dimensional transition-metal phosphorus trichalcogenide nanosheets, in which the transition-metal vacancies enable exceptionally high ion conductivity. A Cd_{0.85}PS₃Li_{0.15}H_{0.15} membrane exhibits a proton conduction dominant conductivity of ~0.95 siemens per centimeter at 90° Celsius and 98% relative humidity. This performance mainly originates from the abundant proton donor centers, easy proton desorption, and excellent hydration of the membranes induced by cadmium vacancies. We also observed superhigh lithium ion conductivity in Cd_{0.85}PS₃Li_{0.3} and Mn_{0.77}PS₃Li_{0.46} membranes.

Nanofluidics, which consist of fluid flow in channels of 1- to 100-nm scale, can cause unusual flow effects because of the strong confinement of molecules and ions (1). The ability to construct nanofluidic channels has led to the discovery of unexpected water molecule- and ion-transport behaviors (2–5). Proton transport through nanochannels under humid conditions—for example, the proton-exchange membrane (PEM)—is crucial for the applications in energy storage and conversion (6, 7). Various materials—including polymers (8, 9), such as Nafion (DuPont Company) (8), metal organic frameworks (MOFs) (10, 11), bio-derived materials (12, 13), and two-dimensional (2D) materials (14, 15)—have been developed for PEMs. Among them, Nafion, the benchmark for PEMs, has a proton conductivity of up to 0.2 S cm^{−1} at high relative humidity (RH). However, Nafion is easily dehydrated under high temperature (> 80°C) and/or at low RH, resulting in a severe decline in conductivity and efficiency. For 2D materials assembled membranes, the large capillary-like force (>50 bar) (4, 5) generated inside the nanochannels can substantially help the ion transport. Nevertheless, the reported membranes show low ion conductivity: less than 0.02 S cm^{−1} (15).

Transition-metal phosphorus trichalcogenides (TMPTCs) are a large family of layered materials with a formula of MPX₃ (M = Cd,

Mn, Fe, Co, Ni, Zn, or Cr; X = S or Se) (16). CdPS₃ nanosheets containing Cd vacancies were synthesized by means of an alkaline ion intercalation exchange method with bulk CdPS₃ crystals as raw material (Fig. 1, A to C, and supplementary materials, materials and methods) (17). We used inductively coupled plasma atomic emission spectroscopy (ICP-AES) to identify the chemical composition of the products (table S1). Cd_{0.85}PS₃K_{0.3} containing Cd vacancies was first obtained by means of K ion intercalation exchange (figs. S1 and S2) and was then subjected to Li ion intercalation exchange (fig. S1), followed by exfoliation, to obtain Cd_{0.85}PS₃Li_{0.3} nanosheets (Fig. 1B). Cd_{0.85}PS₃Li_{0.15}H_{0.15} nanosheets were synthesized by fully soaking Cd_{0.85}PS₃Li_{0.3} nanosheets in HCl solution (Fig. 1C). Nuclear magnetic resonance (NMR) measurements confirmed that ~50% Li ions in Cd_{0.85}PS₃Li_{0.3} were exchanged with protons during the above process (Fig. 1D), which is attributed to the low energy requirement (fig. S3).

The Fourier transform infrared (FTIR) spectra (Fig. 1E) show that the PS₃ asymmetric stretching band at 563 cm^{−1} in CdPS₃ was split into two peaks at 598 and 560 cm^{−1} in Cd_{0.85}PS₃K_{0.3}, Cd_{0.85}PS₃Li_{0.3}, and Cd_{0.85}PS₃Li_{0.15}H_{0.15}, which verified the appearance of Cd vacancies in the layers (18). Compared with Cd_{0.85}PS₃K_{0.3}, the relative intensity of ~598 cm^{−1} peak to ~560 cm^{−1} peak was decreased for Cd_{0.85}PS₃Li_{0.3} and Cd_{0.85}PS₃Li_{0.15}H_{0.15}. Considering the much smaller radius of protons (8 × 10^{−6} Å) (19) and Li ions (0.69 Å) than that of K ions (1.38 Å) and Cd ions (0.95 Å) (20), this result suggests that the Li ions and protons occupy the negatively charged divalent Cd vacancies through electrostatic attraction (21), which is consistent with the NMR results (fig. S4 and materials and methods).

The Cd K-edge x-ray absorption fine structure (XAFS) spectra were also measured to explore the chemical environment of Cd spe-

cies (fig. S5 and materials and methods). The reduced coordination number of Cd and increased disorder degree in Cd_{0.85}PS₃Li_{0.3} and Cd_{0.85}PS₃Li_{0.15}H_{0.15} support the appearance of Cd vacancies (table S2). According to the Cd K-edge extended XAFS (EXAFS) spectra (fig. S5), S L-edge and P L-edge x-ray absorption spectra (fig. S6), and x-ray photoelectron spectroscopy (XPS) spectra (fig. S7), the valence states of Cd, P, and S (+2, +4, and −2, respectively) remain the same in CdPS₃, Cd_{0.85}PS₃Li_{0.3}, and Cd_{0.85}PS₃Li_{0.15}H_{0.15} (materials and methods), suggesting that all the samples keep charge balance.

Atomic force microscope (AFM) measurements show that both Cd_{0.85}PS₃Li_{0.3} and Cd_{0.85}PS₃Li_{0.15}H_{0.15} nanosheets are predominantly monolayers, with a thickness of ~1 nm and average sizes of ~300 and 500 nm, respectively (Fig. 1F and fig. S8). Similar to graphene oxide (GO) (22), these nanosheets can form well-dispersed stable aqueous colloids, as confirmed with the Tyndall effect (Fig. 1, B and C). No sediment was observed, even after being stored for 1 month. The formation of well-dispersed stable colloids is attributed to the electrostatic repulsion between the nanosheets, which are highly negatively charged in water (fig. S9).

We fabricated free-standing Cd_{0.85}PS₃Li_{0.3} and Cd_{0.85}PS₃Li_{0.15}H_{0.15} membranes from their dispersions by means of vacuum filtration (Fig. 2, A and C). The cross-sectional scanning electron microscope (SEM) images reveal that both membranes have well-ordered lamellar structure (Fig. 2, B and D). Moreover, the Cd_{0.85}PS₃Li_{0.3}/Cd_{0.85}PS₃Li_{0.15}H_{0.15} membrane shows a strong x-ray diffraction (XRD) peak at 2θ of ~9.1/9.2° and a weak peak at ~13.8/13.6° (Fig. 2E), corresponding to interlayer distance (*d*) of 0.97/0.96 and 0.64/0.65 nm, respectively. The Cd vacancies enables these two membranes to be highly hydrophilic, with a contact angle of ~20° (Fig. 2, F and G), which is much smaller than those of the pristine CdPS₃ nanosheet-based membrane (~60°) (fig. S10) and GO membrane (~45° to 60°) (23). As a result, water molecules are readily adsorbed in both membranes (fig. S11). The strong XRD peak gives evidence of the intercalated water molecules in the nanochannels, whereas the weak XRD peak is similar to that of CdPS₃ crystals (fig. S1), which is related to the restacking of nanosheets in some regions.

Two-probe alternating current impedance was measured at different temperatures and 98% RH (fig. S12 and materials and methods). Before measurements, the membranes were stored at 98% RH for more than 24 hours, which guaranteed that all the measurements were carried out under equilibrium conditions (fig. S13). At 98% RH, the uptake of water in Cd_{0.85}PS₃Li_{0.15}H_{0.15} and Cd_{0.85}PS₃Li_{0.3}

¹Shenyang National Laboratory for Materials Science, Institute of Metal Research, Chinese Academy of Sciences, 72 Wenhua Road, Shenyang 110016, China. ²School of Materials Science and Engineering, University of Science and Technology of China, 72 Wenhua Road, Shenyang 110016, China. ³State Key Laboratory of Catalysis, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China. ⁴University of Chinese Academy of Sciences, Beijing 100049, China. ⁵National Synchrotron Radiation Laboratory, University of Science and Technology of China, Hefei 230029, China. ⁶Tsinghua-Berkeley Shenzhen Institute (TBSI), Tsinghua University, 1001 Xueyuan Road, Shenzhen 518055, China.

*Corresponding author. Email: wrcen@imr.ac.cn

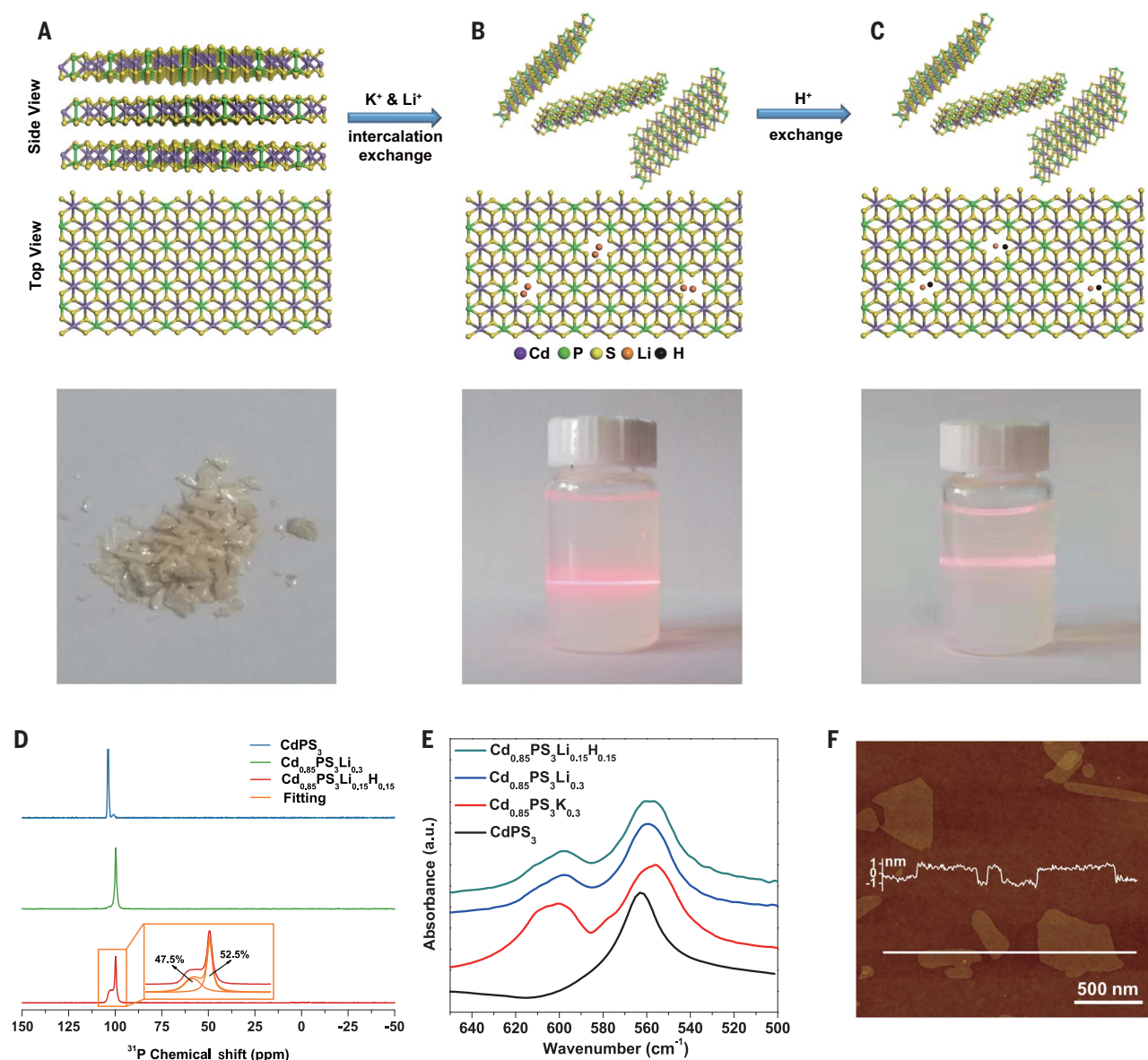


Fig. 1. Synthesis and characterizations of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ and $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ nanosheets. (A to C) Schematic illustration of the synthesis process with CdPS_3 crystals as raw material. (Bottom) (A) CdPS_3 crystals, (B) $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$, and (C) $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ nanosheet water dispersions. (D) ^{31}P single-pulse magic angle spinning NMR spectra. (E) FTIR spectra. (F) A typical AFM image of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ nanosheets, showing thickness of ~ 1 nm.

membranes was 2.97 and 3.9 mol mol^{-1} , respectively (fig. S14). The conductivities were calculated by using the equation

$$\sigma = \frac{L}{R \cdot S} \quad (1)$$

where σ (siemens per centimeter) is conductivity, L (centimeters) is the length of membrane, R (ohms) is the resistance calculated from Nyquist plots (figs. S15 and S16), and S (square centimeters) is the cross-sectional area of membrane. The $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ and $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ membranes exhibit conductivity of 0.952 ± 0.091 and $0.805 \pm 0.085 \text{ S cm}^{-1}$, respectively, at 90°C (Fig. 3A). More-

over, the conductivities are not influenced by the geometry of the membranes (fig. S17).

The ionic transference number of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membranes was determined to be ~ 0.998 at 90°C (fig. S18) (24), proving that they are ion conductors. The proton conduction is the dominant contribution to the conductivity. The measured proton transference number is in the range of 0.55 to 0.72 (fig. S19), which is higher than that of a Nafion membrane (~ 0.5) (fig. S20). The motions of Cd^{2+} ions have negligible contributions to the conductivity (fig. S21). In $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ membranes, the concentration of Li ions is $\sim 10^6$ times higher than that of dissociated protons from adsorbed water, even at $98\% \text{ RH}$.

The huge concentration difference of mobile species suggests that the Li ions contribute predominantly to the conductivity.

In the temperature range from 30° to 90°C , the proton conductivity of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membranes is several times higher than that of the polymeric materials (Fig. 3B and table S3) (8, 9), including Nafion (fig. S22A) (8), and one to four orders of magnitude higher than those of the GO-based materials (fig. S22B) (14, 15), MOFs (10, 11), and bio-derived materials (12, 13). The high-temperature stability is another advantage of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membrane over the Nafion membrane, which shows decreased proton conductivity when the temperature is higher than 80°C (8). Similarly,

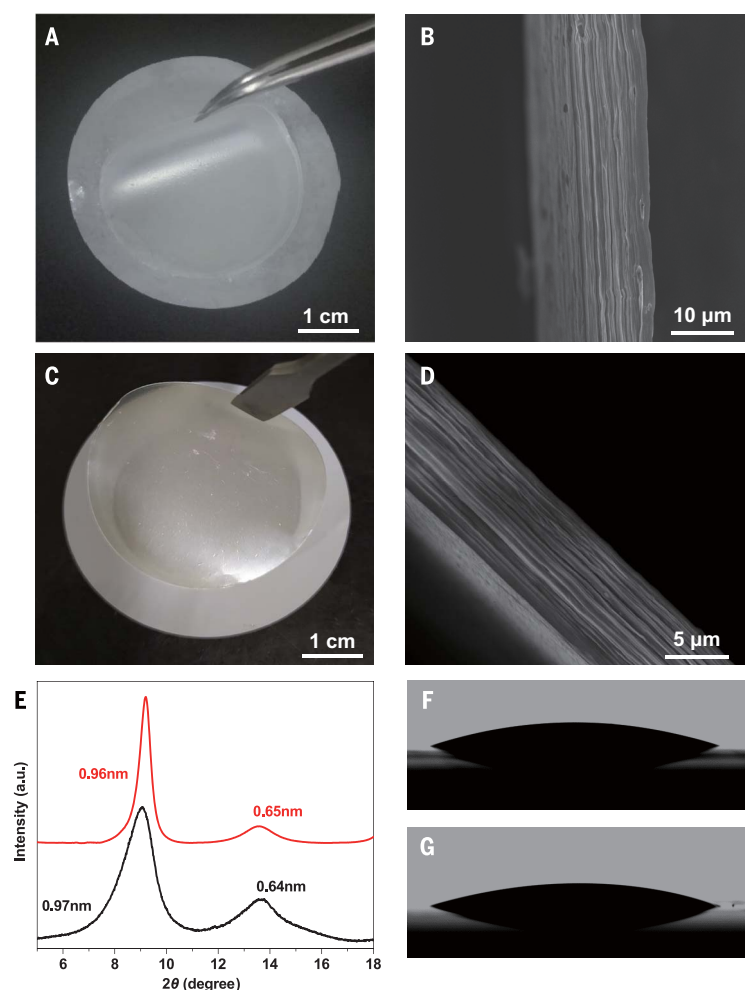


Fig. 2. Structure characterizations of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ and $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membranes. (A and C) Photos of the free-standing (A) $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ and (C) $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membranes. (B and D) Cross-sectional SEM images of the (B) $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ and (D) $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membranes. (E) XRD patterns of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ (black) and $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ (red) membranes measured at ambient conditions with ~30 to 50% RH. (F and G) Wettability of (F) $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ and (G) $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membranes toward water.

$\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ membranes show superior Li ion conductivity that is three to five orders of magnitude higher than those of MOF composite membranes (25) and ionic crystals (26) in the presence of water.

The ion conductivities of both membranes exhibit Arrhenius-type behavior following the equation

$$\ln(\sigma T) = \ln\sigma_0 - \frac{E_a}{RT} \quad (2)$$

where E_a is the activation energy for ion transport, σ_0 is a preexponential factor, R is the gas constant, and T is temperature (Fig. 3C). The fitted E_a for proton and Li ion transport at 98% RH is 0.177 ± 0.016 and 0.21 ± 0.022 eV, respectively. These results suggest that the proton transport is governed by the Grotthuss mechanism (≤ 0.4 eV) (27), in which the protons hop among the hydrogen-bonded water

network in the 2D nanochannels. The E_a value for proton transport is the lowest among the reported PEMs (table S3) and approaches the theoretical upper limit (0.13 eV) (27), indicating that the obstacles to proton transport in $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membranes are very small. Moreover, the E_a for Li ion transport in $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ membranes is lower than most of those for proton transport in the reported PEMs (table S3).

We further studied the kinetic isotope effect (KIE) of both membranes (materials and methods). For the Grotthuss mechanism, the KIE is usually no less than $(m_D/m_H)^{1/2} \sim 1.4$ because the mass of D (m_D) is twice that of H atom (m_H) (27, 28). By contrast, the KIE for vehicle mechanism is equal to the viscosity ratio of D_2O to H_2O , ~ 1.2 (28, 29). At 98% RH, the conductivity of H_2O -humidified $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membrane is 1.87 times higher than that of D_2O -humidified membrane (Fig. 3D), confirm-

ing that the proton transport is governed by the Grotthuss mechanism. The much smaller KIE value (1.25) obtained for $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ membranes further demonstrates its difference in ion conduction from $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membranes (fig. S23).

We also investigated the proton conductivity of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membranes under different RHs (30 to 98%) at 60° and 30°C (fig. S24), respectively. The proton conductivity increases as RH increases, with a sharp increase at high RH over 80% (Fig. 4, A and B). Our membranes also show superior proton conductivity at low RHs. The conductivity (0.260 S cm^{-1}) at 60% RH and 60°C is about three times higher than that of Nafion at 100% RH and the same temperature (8). At similar low temperature (30°C) and low RHs ($\leq 75\%$), the conductivity is one to six orders of magnitude higher than those of reported proton-conducting materials (table S4). At 33% RH, E_a is ~ 0.2 eV, and KIE is ~ 1.4 (figs. S25 and S26), suggesting that the proton transport is also governed by the Grotthuss mechanism at low RHs.

We performed XRD measurements to identify the structure evolution of the $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membranes at different RHs (fig. S27). The XRD peak related to the restacked nanosheets gradually disappears until 53% RH, along with a d increase of ~ 0.3 nm (Fig. 4, C and D), suggesting that the water clusters gradually expanded and joined each other to form a monolayer water network (30). Further increasing humidity leads to a sharp increase in d by another 0.3 nm until 100% RH, suggesting the formation of bilayer water networks. Such ordered hydrogen-bonded water networks (fig. S28) (31) may help proton transport governed by the Grotthuss mechanism. We observed similar trends for the structure evolution and Li ion conductivity of the $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ membrane as a function of RH (fig. S29).

The Cd vacancy plays a key role in boosting ion conductivity in our membranes. The proton conductivity of $\text{Cd}_{0.9}\text{PS}_3\text{Li}_{0.1}\text{H}_{0.1}$ membranes is around 0.442 S cm^{-1} at 60°C and 98% RH, which is lower than that of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membranes (0.641 S cm^{-1}) at the same atmosphere (fig. S30A). Similarly, $\text{Cd}_{0.9}\text{PS}_3\text{Li}_{0.2}$ membranes show much lower Li ion conductivity (0.346 S cm^{-1}) than that of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ membranes (0.583 S cm^{-1}) at 60°C and 98% RH (fig. S30B). The pristine CdPS_3 nanosheet-assembled membrane shows a temperature-independent ion conductivity of $\sim 2.5 \times 10^{-4} \text{ S cm}^{-1}$ at 98% RH (fig. S31), which is about four orders of magnitude lower than those of the Cd vacancy-containing membranes.

We performed density functional theory calculations to understand the role of Cd vacancy in proton transport. The Cd vacancy-free monolayer CdPS_3 shows a strong repulsion against protons in terms of the adsorption

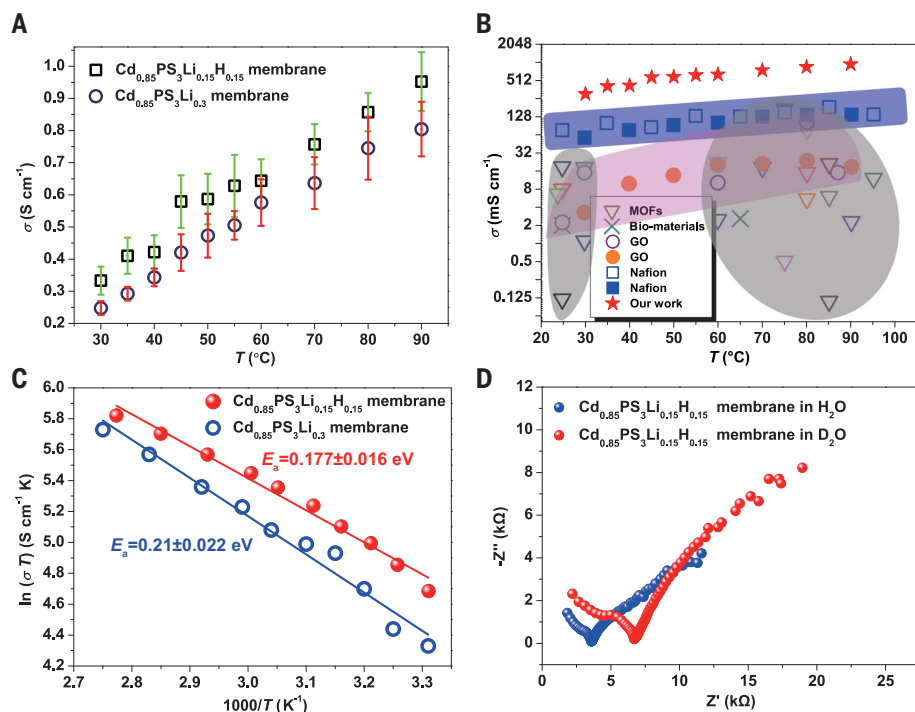


Fig. 3. Proton and Li ion transport properties of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ and $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ membranes. (A) Temperature-dependent conductivity of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membrane and $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ membrane at 98% RH. (B) A comparison of conductivity of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membrane (stars) with other proton-conducting materials at different temperatures and high RH ($\geq 90\%$). Open squares and circles indicate the reported conductivities of Nafion and GO membranes, respectively, and the solid shapes indicate those measured with our platform. (C) Arrhenius plots of conductivities of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ and $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.3}$ membranes at 98% RH. (D) Typical Nyquist plots for the $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membrane in the presence of H_2O (blue) and D_2O (red) at 98% RH.

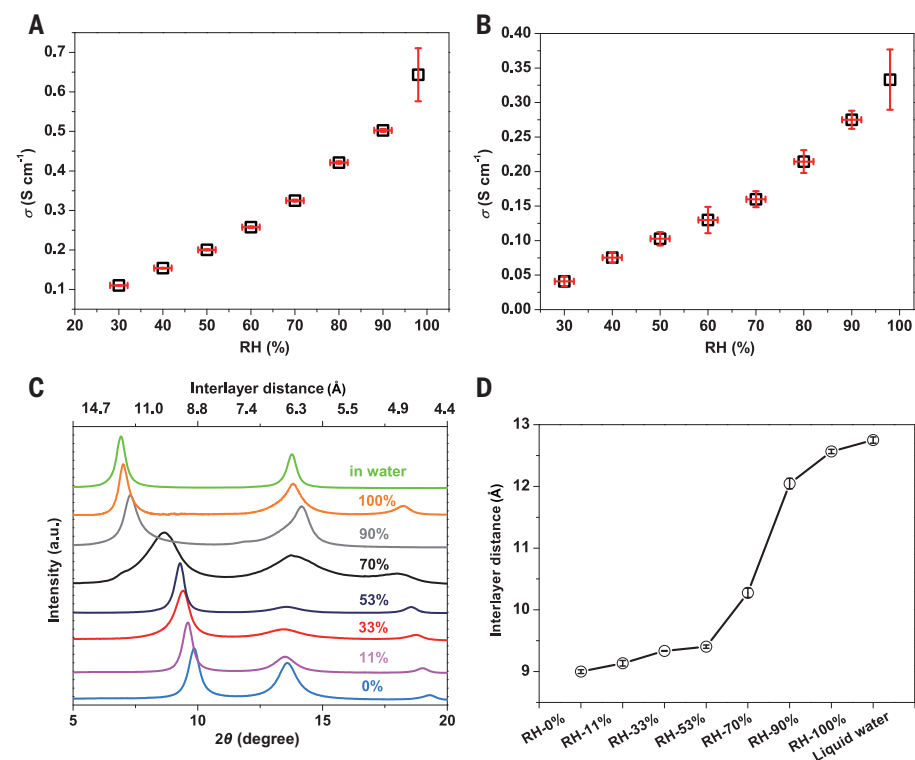


Fig. 4. Proton conductivity and interlayer distance evolution of $\text{Cd}_{0.85}\text{PS}_3\text{Li}_{0.15}\text{H}_{0.15}$ membranes at different RHs. (A and B) RH-dependent conductivity at (A) 60°C and (B) 30°C. (C) XRD patterns at different RHs and in water. (D) RH-dependent interlayer distance extracted from the XRD peak in (C).

energy 1.416 eV per proton (fig. S32, A and B). By contrast, protons can easily adsorb on the Cd vacancy, with an adsorption energy of -0.908 eV per proton (fig. S32, C and D). The Li ions that adsorbed on the Cd vacancy

can also be exchanged with protons (Fig. 1D and fig. S3). These results indicate that the Cd vacancies in CdPS_3 play a key role in providing a proton source for proton transport. Moreover, the adsorbed protons on Cd vacancies

can easily detach from the nanosheet in the presence of water and then attach to water clusters, with an energy change (ΔE) of -1.199 eV (fig. S33), forming hydronium ions.

On the basis of the above studies, the ultra-fast ion transport in our membranes at a broad range of RHs and temperatures can be understood as follows. The negatively charged divalent Cd vacancies in the nanosheets serve as a donor center to provide ions, which resemble the function of SO_3^- groups in Nafion (8). The high hydrophilicity allows the formation of an ordered water network even at a low RH. Under such humidified conditions, the interaction between ions and Cd vacancies is eliminated by the water molecules, and consequently, ions readily get rid of the control from Cd vacancies for transport in the nanochannels. At high RHs, the larger d and continuous bilayer water networks further reduce the ion transport barrier, leading to rapid increase in the ion conductivity (Fig. 4, A and B, and fig. S29). In addition, the regular 2D nanofluidic channels and the resulting nanocapillary force (4, 5) may also play a role in improving the water diffusion and consequently the proton and Li ion transport.

To show the promise of our membranes for practical applications, we constructed a simulated PEM fuel cell (materials and methods). The proton conductivity shows a small variation in the range of 0.6 to 0.7 S cm^{-1} at 60°C without degradation after 36 hours of testing (fig. S34), confirming the good stability of our membranes. Moreover, the membranes show good flexibility (movie S1) and higher tensile strength than that of Nafion membranes (fig.

S35) (32, 33). We also fabricated Mn vacancy-containing $\text{Mn}_{0.7}\text{PS}_3\text{Li}_{0.46}$ membranes (figs. S36 to S41 and materials and methods), which show ultrahigh Li ion conductivity of $\sim 0.75 \text{ S cm}^{-1}$ at 90°C and 98% RH with $E_a \sim 0.24 \pm 0.02 \text{ eV}$, demonstrating the versatility of the role of transition-metal vacancies for boosting ion transport. These transition-metal vacancies-containing 2D TMPTC membranes add to the large family of ion transport membranes.

REFERENCES AND NOTES

1. L. Bocquet, E. Charlaix, *Chem. Soc. Rev.* **39**, 1073–1095 (2010).
2. J. C. T. Eijkel, A. V. D. Berg, *Microfluid. Nanofluid.* **1**, 249–267 (2005).
3. R. B. Schoch, J. Han, P. Renaud, *Rev. Mod. Phys.* **80**, 839–883 (2008).
4. R. K. Joshi *et al.*, *Science* **343**, 752–754 (2014).
5. R. R. Nair, H. A. Wu, P. N. Jayaram, I. V. Grigorieva, A. K. Geim, *Science* **335**, 442–444 (2012).
6. B. E. Logan, M. Elimelech, *Nature* **488**, 313–319 (2012).
7. K. D. Kreuer, *J. Membr. Sci.* **185**, 29–39 (2001).
8. L. Liu, W. Chen, Y. Li, *J. Membr. Sci.* **504**, 1–9 (2016).
9. R. Wang, X. Wu, X. Yan, G. He, Z. Hu, *J. Membr. Sci.* **479**, 46–54 (2015).
10. N. T. Nguyen *et al.*, *J. Am. Chem. Soc.* **137**, 15394–15397 (2015).
11. M. Liu *et al.*, *Nat. Commun.* **7**, 12750 (2016).
12. D. D. Ordinario *et al.*, *Nat. Chem.* **6**, 596–602 (2014).
13. E. E. Josberger *et al.*, *Sci. Adv.* **2**, e1600112 (2016).
14. G. He *et al.*, *Adv. Funct. Mater.* **25**, 7502–7511 (2016).
15. M. R. Karim *et al.*, *J. Am. Chem. Soc.* **135**, 8097–8100 (2013).
16. F. Wang *et al.*, *Adv. Funct. Mater.* **28**, 1802151 (2018).
17. R. Clement, O. Garnier, J. Jegoudez, *Inorg. Chem.* **25**, 1404–1409 (1986).
18. P. Jeevanandam, S. Vasudevan, *Solid State Ion.* **104**, 45–55 (1997).
19. R. Pohl *et al.*, *Nature* **466**, 213–216 (2010).
20. Y. Marcus, *Biophys. Chem.* **51**, 111–127 (1994).
21. I. Lagadic, P. G. Lacroix, R. Clément, *Chem. Mater.* **9**, 2004–2012 (1997).
22. D. Li, M. B. Müller, S. Gilje, R. B. Kaner, G. G. Wallace, *Nat. Nanotechnol.* **3**, 101–105 (2008).
23. K. H. Thebo *et al.*, *Nat. Commun.* **9**, 1486 (2018).
24. J. B. Wagner, C. Wagner, *J. Chem. Phys.* **26**, 1597–1601 (1957).
25. Y. Guo, Y. Ying, Y. Mao, X. Peng, B. Chen, *Angew. Chem. Int. Ed.* **55**, 15120–15124 (2016).
26. N. Yoshinari, S. Yamashita, Y. Fukuda, Y. Nakazawa, T. Konno, *Chem. Sci.* **10**, 587–593 (2018).
27. N. Agmon, *Chem. Phys. Lett.* **244**, 456–462 (1995).
28. T. E. DeCoursey, V. V. Cherny, *J. Gen. Physiol.* **109**, 415–434 (1997).
29. N. K. Roberts, H. L. Northey, *J. Chem. Soc., Faraday Trans.* **70**, 253–262 (1974).
30. B. Rezanian, N. Severin, A. V. Talyzin, J. P. Rabe, *Nano Lett.* **14**, 3993–3998 (2014).
31. K. Gopinadhan *et al.*, *Science* **363**, 145–148 (2019).
32. L. Yu, F. Lin, L. Xu, J. Xi, *RSC Adv.* **6**, 3756–3763 (2016).
33. J. Li *et al.*, *J. Mater. Chem. A* **8**, 196–206 (2020).

ACKNOWLEDGMENTS

We thank X. Yu, T. Zhou, and W. Zhang for the help on AFM measurements and Z. Jiang and L. Cao from Tianjin University for kind teaching on ion conductivity measurements. **Funding:** This work was financially supported by the National Natural Science Foundation of China (51325205, 51290273, 51521091, and 51972312), the Key Research Program of Frontier Sciences of the Chinese Academy of Sciences (ZDBS-LY-JSC027), the Strategic

Priority Research Program of the Chinese Academy of Sciences (XDB30000000), the National Key R&D Program of the Ministry of Science and Technology of China (2016YFA0200101), LiaoNing Revitalization Talents Program (XLYC1808013), the Program for Guangdong Introducing Innovative and Entrepreneurial Teams, and the Development and Reform Commission of Shenzhen Municipality for the development of the “Low-Dimensional Materials and Devices” discipline. The theoretical calculations were performed on TianHe-1(A) at the National Supercomputer Center in Tianjin and TianHe-2 at the National Supercomputer Center in Guangzhou. **Author contributions:** W.R. conceived and supervised the project. W.R. and X.Q. designed the experiments. X.Q. performed the experiments. L.C. helped with the sample preparation. L.Y. performed calculations. Z.L. carried out TEM characterizations. S.P. helped with the property measurements. F.L. and G.H. carried out NMR measurements and analyses. S.C. and L.S. carried out XAFS measurements and analyses. X.Q. and W.R. analyzed the data and wrote the manuscript with input from all authors. K.H.T. commented on the manuscript. H.-M.C. advised the project and commented on the manuscript. **Competing interests:** W.R., X.Q., L.C., and H.-M.C. are listed as coinventors on a pending Chinese patent application related to this work filed by Institute of Metal Research, Chinese Academy of Sciences (application no. 202010687242.8). All other authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/596/suppl/DC1
Materials and Methods
Figs. S1 to S41
Tables S1 to S4
References (34–73)
Movie S1

30 March 2020; accepted 10 September 2020
10.1126/science.abb9704

TOPOLOGICAL SYSTEMS

Generation of helical topological exciton-polaritons

Wenjing Liu¹, Zhurun Ji¹, Yuhui Wang¹, Gaurav Modi¹, Minsoo Hwang¹, Biyuan Zheng², Volker J. Sorger³, Anlian Pan², Ritesh Agarwal^{1*}

Topological photonics in strongly coupled light-matter systems offer the possibility for fabricating tunable optical devices that are robust against disorder and defects. Topological polaritons, i.e., hybrid exciton-photon quasiparticles, have been proposed to demonstrate scatter-free chiral propagation, but their experimental realization to date has been at deep cryogenic temperatures and under strong magnetic fields. We demonstrate helical topological polaritons up to 200 kelvin without external magnetic field in monolayer WS₂ excitons coupled to a nontrivial photonic crystal protected by pseudo time-reversal symmetry. The helical nature of the topological polaritons, where polaritons with opposite helicities are transported to opposite directions, is verified. Topological helical polaritons provide a platform for developing robust and tunable polaritonic spintronic devices for classical and quantum information-processing applications.

The recognition of the role of topology in condensed matter systems provided a route for understanding and creating new phases of quantum matter and exploring new phenomena and their applications (1). Topologically inequivalent insulators cannot be adiabatically transformed into one another without a phase transition that involves closing the bandgap. As such, the transport through interface states formed at the boundary of topologically inequivalent insulators is topologically protected and free from back-scattering by various defects. Depending on the presence of time-reversal symmetry, two-dimensional (2D) topological insulators can be broadly divided into two classes: quantum Hall (QH) systems, which support chiral charge currents, and quantum spin Hall (QSH) systems, which support helical spin currents. With the extension of the topological formalism to photonic systems (2), topological photonics is now an active research field (3). Photonic analogs of QH and QSH effects have been demonstrated (4–7), which gave rise to topological devices including robust waveguides, routers, and lasers (8–10).

Polaritons have emerged as an important platform for exploring new topological matter that bridges electronic and photonic systems (11–19). Exciton-polaritons are half-light, half-matter quasiparticles resulting from the strong hybridization between excitons and photons when the coherent energy transfer rate between them exceeds their decay rates. Proposals on realizing topological polaritonic states use excitons in a topologically nontrivial system (17), the spin-orbit coupling of

photons (16), or the photon-exciton hybridization process (11), allowing manipulation of topological properties by independent control of each component (20). Compared with bare photonic systems, the excitonic component in polaritons introduces strong interparticle interactions and enhanced sensitivity to external stimuli that can be useful for assembling nonlinear and actively controllable topological devices.

Recently, a 2D polaritonic QH system was experimentally achieved in quantum-well lattices with nontrivial topology originating from the TE-TM splitting of the photonic mode and the topological bandgap opened by the Zeeman splitting of excitons. Although this is a promising demonstration of a new phase of topological matter, the small bandgap (~0.1 meV) opened under strong magnetic fields and the 4 K operation temperature may restrict its widespread utilization. Fortunately, it is possible to avoid these challenges in an analogous QSH polaritonic system. QSH insulators do not require a magnetic field or magnetic order,

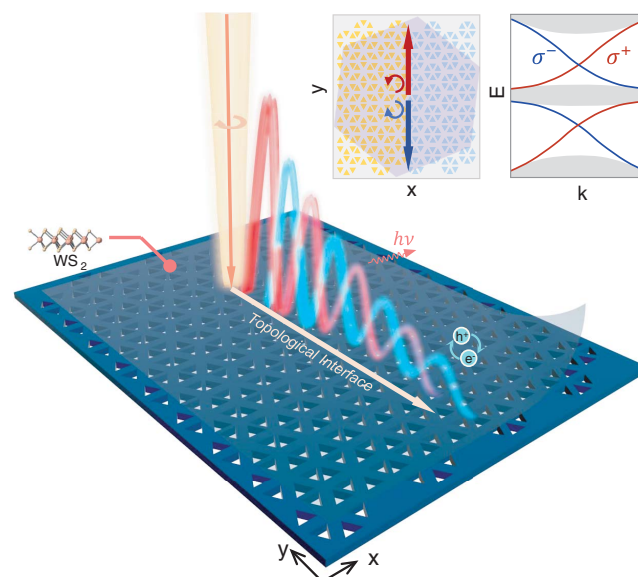
and their optical analogs can be realized solely by exploiting the system symmetry and interactions with large topological bandgaps (>10 meV), thus providing opportunities for high-temperature operation. In addition, the additional spin-degree of freedom and helical nature of the analog QSH polaritons can be further exploited for polariton-based spintronic devices.

Here, we demonstrate helical topological polaritons in an analogous QSH system by strongly coupling monolayer WS₂ excitons with a topologically nontrivial hexagonal photonic crystal (Fig. 1). The polariton topology is protected by a pseudo time-reversal symmetry originating from the C_{6v} symmetry of the coupled system. Analogous to the helical edge states in a QSH insulator, topological exciton-polaritons in this system are distinct from their trivial counterparts. Polaritons of different helicity propagate in opposite directions and are topologically protected from back-scattering. Monolayer transition metal dichalcogenides are excellent candidates for exploring 2D polaritonic physics at higher temperatures because of their large exciton oscillator strength (21–25), and their ultrathin geometry ensures minimum perturbation to the band structure of the underlying photonic crystal.

Our photonic crystals are formed of hexagonal lattices, with each unit cell composed of six particles located on the perpendicular bisector of its six sides (Fig. 2A) (6, 7, 26–28). While keeping the sixfold rotational symmetry intact, the six particles can be simultaneously arranged toward or away from the unit cell center, which changes the intercell and intracell nearest-neighbor couplings. The topology of the lattice is controlled by swapping the intercell and intracell coupling strengths

Fig. 1. Schematic of the helical topological polariton system.

The topological polaritonic system consists of monolayer WS₂ strongly coupled to a nontrivial photonic crystal. Spin momentum-locked helical topological polaritons are confined to the domain wall connecting the two bulk lattice regions with different topology. Inset is a top view of the topological polariton device and a schematic of the helical polariton dispersion.



¹Department of Materials Science and Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA. ²Key Laboratory for Micro-Nano Physics and Technology of Hunan Province, State Key Laboratory of Chemo/Biosensing and Chemometrics, College of Materials Science and Engineering, Hunan University, Changsha, Hunan 410082, P. R. China.

³Department of Electrical and Computer Engineering, George Washington University, Washington, DC 20052, USA.

*Corresponding author. Email: riteshag@seas.upenn.edu

(29). We need to consider the coupling between such a photonic crystal and the A-excitons with isotropic polarization in monolayer WS₂ (24). The excitonic material was divided into six sublattices corresponding to each photonic unit cell. The tight-binding approximation was then applied by treating each of these sublattices as an isolated excitonic particle located at its mass center (Fig. 2A). The coupled system thus contains two particles within each sublattice corresponding to the excitonic and photonic structures, respectively. The Hamiltonian of the coupled system can be written as (see the supplementary materials, section 1A) follows:

$$H = \hbar \sum_i (\omega_{ph} \hat{a}_i^\dagger \hat{a}_i + \omega_{ex} \hat{b}_i^\dagger \hat{b}_i) + \sum_{NN} (f_{ij} \hat{a}_i^\dagger \hat{a}_j + H.c.) + g \sum_i (\hat{a}_i^\dagger \hat{b}_i + H.c.) \quad (1)$$

where \hat{a} (\hat{a}^\dagger) and \hat{b} (\hat{b}^\dagger) are the annihilation (creation) operators for photons and excitons, respectively; $i, j = 1, 2, \dots, 6$ are the six sublattices; $\hbar\omega_{ph}$ and $\hbar\omega_{ex}$ are the energies of the uncoupled photons and excitons, respectively; g_{ij} is the exciton-photon coupling strength; f_{ij} is the nearest-neighbor coupling factor of the photonic lattice, which is t_1 for the intracell coupling and $t_2 \exp(i\mathbf{k} \cdot \mathbf{a}_i)$ for the intercell coupling, with \mathbf{a}_i being the lattice vector.

Based on the C_{6v} symmetry of the exciton-photon coupled system, the Hamiltonian can be projected onto the basis $(|p_{\pm}^{ph}\rangle, |d_{\pm}^{ph}\rangle, |p_{\pm}^{ex}\rangle,$

$|d_{\pm}^{ex}\rangle)^T$, where p and d denote modes with odd and even parities, respectively, and $+$ or $-$ signs represent pseudospin-up or pseudospin-down states associated with the positive or negative angular momenta, respectively. Near the Γ point of the Brillouin zone, the pseudospin-up and pseudospin-down states are decoupled and degenerate, each leading to four polariton bands (see the supplementary materials, section 1A):

$$\begin{aligned} \lambda_{upper} &= \hbar\omega_0 + \frac{1}{2} \left(\pm\alpha + \sqrt{4g^2 + (\mp\alpha + \delta_{ep})^2} \right) \\ \lambda_{lower} &= \hbar\omega_0 + \frac{1}{2} \left(\pm\alpha - \sqrt{4g^2 + (\mp\alpha + \delta_{ep})^2} \right) \end{aligned} \quad (2)$$

where upper and lower denote the bands above and below the exciton energy ω_{ex} respectively, $\omega_0 = \frac{1}{2}(\omega_{ex} + \omega_{ph})$, $\delta_{ep} = \hbar(\omega_{ex} - \omega_{ph})$, and α corresponds to the dispersion of the uncoupled photonic crystal with $\alpha = (t_2 - t_1)^2 + \frac{3}{4}t_2^2k^2$. When $t_1 = t_2$, the two upper and lower polariton bands are degenerate at the Γ point with linear dispersion for small k values. These represent two Dirac cones with energy splitting of $\sqrt{4g^2 + \delta_{ep}^2}$, and a typical polariton dispersion can be characterized by a bulk band Rabi splitting, $\Omega_{bulk} = 2g$. This degenerate condition is the topological phase transition point of the designed exciton-photon coupled system. If $t_1 \neq t_2$, then a finite “mass” term is induced by α , which opens a bandgap in both

the upper and lower Dirac cones (Fig. 2, B and C), with two different topological phases identified for $t_1 > t_2$ and $t_1 < t_2$ conditions, respectively (see the supplementary materials, section 1A). The topological phase transition can also be clearly observed by the inversion of band parity near the Γ point (Fig. 2, B and C) at both the upper and lower bandgaps: For $t_1 > t_2$, the band below each bandgap has an odd parity ($|p\rangle$ band), whereas the band above the bandgap has an even parity ($|d\rangle$ band), which reverses for the $t_1 < t_2$ condition.

Finite-difference time-domain (FDTD) simulations were performed on a WS₂-photonic crystal coupled structure (Figs. 1 and 2) that confirmed the predictions of the tight-binding model (see the supplementary materials, sections 2A to 2C). The key for obtaining a topologically nontrivial phase in the coupled system is the preservation of the C_{6v} symmetry, which is fulfilled by isotropic WS₂ excitons (24). The helical topological polariton interface states were extracted analytically by solving the real space wavefunction of the interface polaritons. For a polaritonic system with $t_2 > t_1$ at $x < 0$, and $t_2 < t_1$ at $x > 0$, the topological interface mode can be written as follows [see the supplementary materials, section 1B (7)]:

$$|\Psi\rangle = \begin{cases} \Sigma_i A_i |\phi_i^A\rangle \exp(ik_{y,i}^A y) \exp(|\delta_x^A| x) & \text{for } x < 0 \\ \Sigma_i B_i |\phi_i^B\rangle \exp(ik_{y,i}^B y) \exp(-|\delta_x^B| x) & \text{for } x > 0 \end{cases} \quad (3)$$

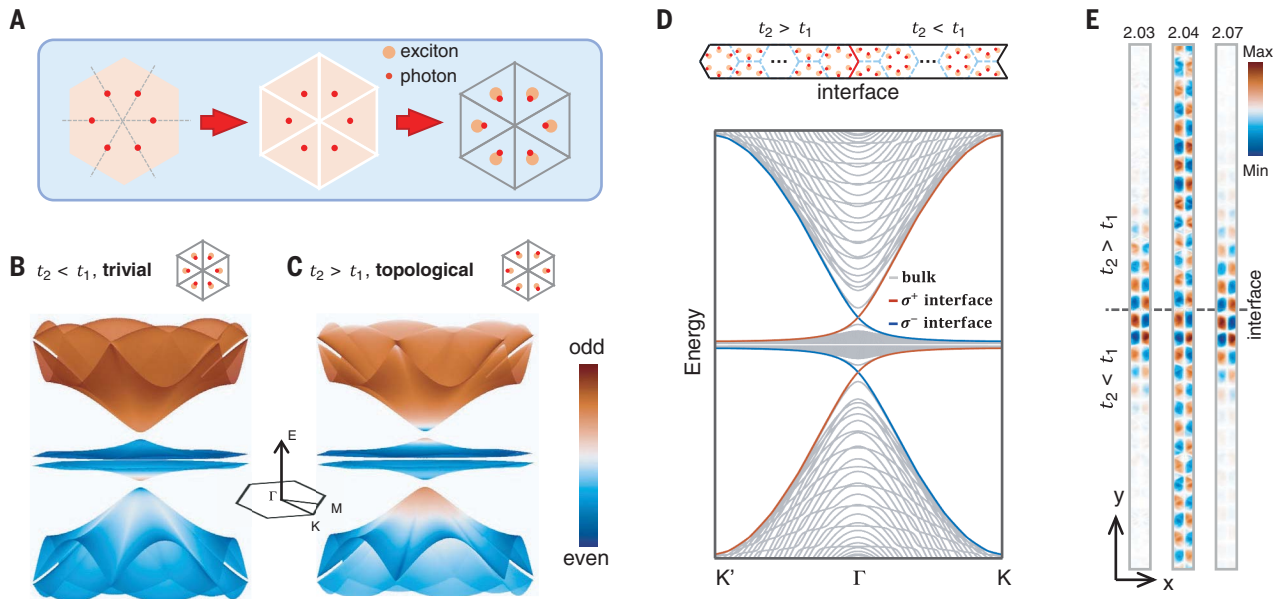
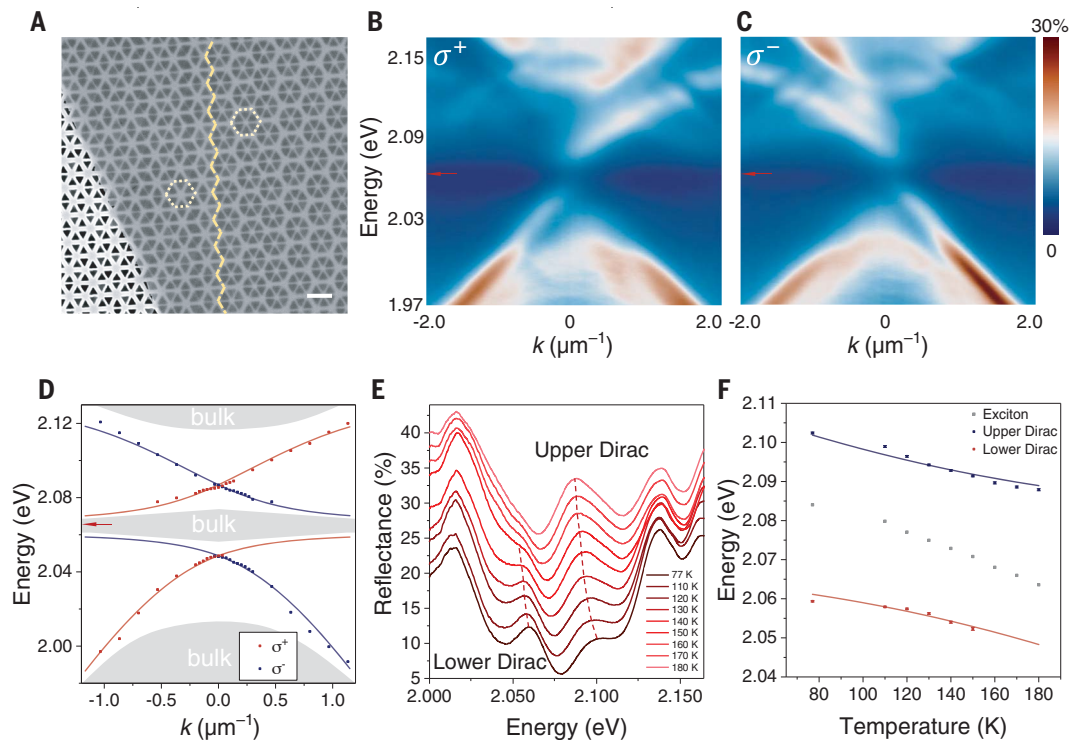


Fig. 2. Theoretical and numerical studies of the helical topological exciton-polaritons. (A) Illustration of the tight-binding model. The coupled structure contains six photonic resonators (red dots) in each hexagonal unit cell. The 2D excitonic layer (orange) is divided into six sublattices according to the photonic lattice symmetry, and excitonic particles are assumed to be located at the mass center of each sublattice. (B and C) Calculated band structures of the trivial and topological polaritonic lattices. The color plots the spatial parity of the bands. (D) Helical

topological interface polariton states calculated from two polaritonic crystals with different topology connected by a zigzag domain wall. The pseudospin-up (red) and pseudospin-down (blue) interface polariton bands are associated with the positive and negative group velocities, respectively. (E) Spatial H_2 field distribution of the WS₂-photonic crystal system at the domain wall calculated at 2.03 and 2.07 eV (two polaritonic Dirac points with field confinement) and at 2.04 eV, where the field extends into the bulk region between the Dirac points (fig. S2E).

Fig. 3. Identification of the helical topological polaritons through momentum-space reflectance spectra. (A) Scanning electron microscope image of the WS₂-photonic crystal coupled structure at the topological domain wall (yellow dashed line). Scale bar, 500 nm. (B and C) Angle-resolved reflectance spectra of the helical interface polaritons excited by σ^+ and σ^- polarized light. The red arrows indicate exciton energy.



which is propagating along the y direction while evanescently decaying in the x direction. Here, $|\phi_i^A\rangle$ and $|\phi_i^B\rangle$ are the eigenfunctions for $x < 0$ and $x > 0$ and k_y and $|\delta_x|$ are the wave vector and decay constant along and perpendicular to the interface, respectively. The interface polariton dispersion can subsequently be obtained as follows (for pseudospin-up states):

$$E_{\pm} = \hbar\omega_0 + \frac{1}{2}\hbar vk_y \pm \frac{1}{2}\sqrt{4g^2 + (-\hbar vk_y + \delta_{ep})^2} \quad (4)$$

where $v = \frac{\sqrt{3}}{2}a_0t_2$, and the + and - signs represent the upper and lower branches of the topological interface polariton modes, respectively. The dispersion curves always have a positive group velocity, whereas a set of interface polariton dispersions with opposite group velocities can be found for pseudospin-down states, therefore demonstrating the helical nature of the topological interface polaritons. The dispersion described by Eq. 4 exhibits an anticrossing behavior at the exciton energy that is characteristic of exciton-polariton dispersion, in which the Rabi frequency, $\Omega_{\text{topo}} = 2g$, is the same as the Rabi frequency of the bulk polariton, i.e., Ω_{bulk} . Although this result was obtained from a simplified tight-binding model, such consistency demonstrates that the anticrossing in the topological interface polariton bands is a direct consequence of the eigenmode splitting in the bulk polariton lattice, which can be viewed as the expression of the bulk-edge correspondence principle in the strong coupling regime. In other words,

there are two equivalent descriptions of the anticrossing feature in the topological interface polaritonic dispersion: strong coupling between the exciton and the photonic interface state and, equivalently, the consequence of non-trivial topology of the split bulk polariton bands.

Numerical calculations were also performed on the tight-binding model to obtain the interface polariton dispersion (Fig. 2D and supplementary materials, section 1B). Topological interface states with Dirac type dispersion emerged spanning both the upper and lower bandgaps, with positive (negative) group velocities corresponding to the pseudospin-up (pseudospin-down) states, confirming their helical nature. FDTD simulations were performed on a WS₂-photonic crystal system with a topological domain wall (Fig. 2E and supplementary materials, section 2E). At the two Dirac points at 2.03 and 2.07 eV, the electromagnetic field is spatially confined to the topological interface, whereas in between, the field is found to spatially extend into the bulk lattice, corresponding to the central bulk polariton bands. The gap-spanning interface mode dispersion and the observation of the central bulk bands in the simulation further emphasize that the splitting in the topological interface polariton dispersion does not result from the gapping out of the interface state as in a topologically trivial system; instead, it is a natural phenomenon in a strongly coupled topological structure.

Experimentally, photonic crystals were patterned on suspended SiN_x membranes followed by transfer of monolayer WS₂ flakes (see the supplementary materials and methods) (Fig.

3A). The SiN_x photonic crystals were composed of six triangular holes arranged in a hexagonal lattice, in which the intercell and intracell interaction strengths were controlled by moving the holes toward or away from the unit cell center. The structure supports a TE electromagnetic mode with the pseudospin corresponding to the angular momentum of the out-of-plane \mathbf{H} -field, and hence the helicity of the in-plane \mathbf{E} -field at the center of each unit cell, which can be directly probed by a far-field circularly polarized light (7).

The energy-wavevector (E - k) dispersion of the topological polaritons was studied by angle-resolved reflectance measurements at 160 K. In this particular sample, the WS₂ A-exciton is located at 2.067 eV, in close resonance with ω_{ph} (detuning, $\delta_{\text{ep}} = -7$ meV). Only at the topological interface is a set of unidirectional dispersion curves found between the upper and lower bulk polariton bands, with negative (positive) group velocities associated with the σ^+ (σ^-) polarizations, respectively, corresponding to the helical topological states (Fig. 3, B and C). Away from the interface, the unidirectional dispersion disappeared, leaving only the gapped bulk polariton bands (see the supplementary materials, section 3D). An anticrossing in the interface polariton dispersion curve was observed at the exciton energy, whereas the bulk central polariton bands were not observable because of low far-field coupling efficiency (see the supplementary materials, section 2D). All four interface bands with both helicities fit very well to the tight-binding model (Fig. 3D), resulting in a

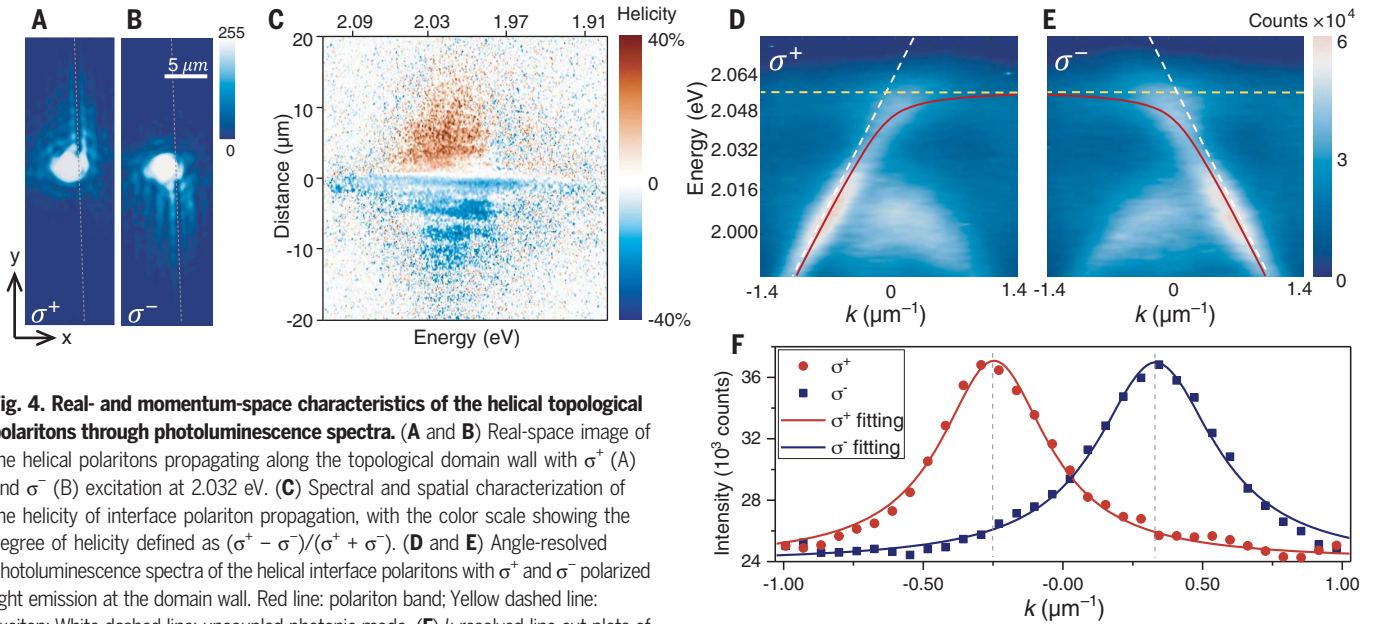


Fig. 4. Real- and momentum-space characteristics of the helical topological polaritons through photoluminescence spectra. (A and B) Real-space image of the helical polaritons propagating along the topological domain wall with σ^+ (A) and σ^- (B) excitation at 2.032 eV. (C) Spectral and spatial characterization of the helicity of interface polariton propagation, with the color scale showing the degree of helicity defined as $(\sigma^+ - \sigma^-)/(\sigma^+ + \sigma^-)$. (D and E) Angle-resolved photoluminescence spectra of the helical interface polaritons with σ^+ and σ^- polarized light emission at the domain wall. Red line: polariton band; Yellow dashed line: exciton; White dashed line: uncoupled photonic mode. (F) k -resolved line cut plots of (D) and (E) at $\hbar\omega = 2.032$ eV and the corresponding single Lorentz peak fitting.

Rabi splitting $\Omega_{\text{topo}} = 38 \pm 1$ meV, suggesting that the system is in the strong coupling limit, given that $\gamma_{\text{ex}} = 15$ meV and $\gamma_{\text{ph}} = 6.5$ meV, satisfying the strong coupling criterion, $\Omega > \sqrt{2(\gamma_{\text{ex}}^2 + \gamma_{\text{ph}}^2)}$. It should be noted that in the bare photonic crystals, a small bandgap also exists in the dispersion of the interface states caused by the C_{6v} symmetry breaking at the interface (26). In our structure, this small gap (< 3 meV) is much smaller than the polariton Rabi splitting, so it does not affect our analysis (see the supplementary materials, section 3H). In a different sample, interface-confined topological polariton formation was also confirmed with $\Omega_{\text{topo}} = 32 \pm 1$ meV (see the supplementary materials, section 3C) at 200 K. Therefore, these results demonstrate the realization of topological polaritons at relatively high temperatures because of the large exciton oscillator strength of monolayer WS_2 and the appropriately designed photonic structure. These topological polaritons have their in-plane momentum locked to the helicity of light, which are confined spatially at the interface and spectrally within the bulk bandgap, exhibiting striking differences from conventional exciton-photon polaritons and from the dispersion of the interface photonic modes. A Rabi splitting of $\Omega_{\text{bulk}} = 40 \pm 1$ meV was extracted from the same flake, which is comparable to Ω_{topo} , thus verifying our theoretical prediction that the Rabi splitting of the topological interface state is in accordance with that of the bulk (see the supplementary materials, section 3E).

The temperature-dependent polariton response was investigated by tuning the A-exciton energy from 77 to 180 K (Fig. 3, E and F). Both the upper and lower Dirac points red-shift with

increasing temperature, which can be fitted to Eq. 4 at $k_y = 0$, resulting in a Rabi splitting $\Omega_{\text{topo}} = 38 \pm 2$ meV (Fig. 3F), consistent with the values extracted in Fig. 3D and those from the FDTD simulations (Fig. 2E). The mode shifting with temperature also demonstrates the ability to actively engineer the dispersion of the topological modes through the excitonic component. Precisely tuning the dispersion is critical in controlling the information transport speed, light-matter interaction strength, and nonlinear processes in devices, which is challenging in photonic topological systems. In polaritonic systems, the dispersion can be actively controlled spanning a broad parameter space by tuning the relative weights of excitons and photons. In the device demonstrated in Fig. 3, by changing the device temperature from 77 to 180 K, the group velocity can be tuned from 30 $\mu\text{m}/\text{ps}$ to nearly 0, and effective polariton mass can be varied by two orders of magnitude within a 10 K range at a single operating wavelength and without breaking the topological protection of these modes (see the supplementary materials, section 3G).

Helical topological polariton transport was then studied in both the real and momentum space (Fig. 4) through real-space imaging and photoluminescence spectroscopy (see the supplementary materials and methods). The experiments were first performed at 200 K in a device with $\hbar\omega_{\text{ex}} = 2.055$ eV and $\hbar\omega_{\text{ph}} = 2.060$ eV with a circularly polarized laser ($\hbar\omega = 2.032$ eV) incident at the topological domain wall. A real-space image was taken to directly visualize the waveguiding of the helical interface states, with the unidirectional, spin momentum-locked waveguiding along the interface readily observed (Fig. 4, A and B).

By contrast, excitation in the bulk lattice region leads to no observable waveguiding at energies within the polaritonic bulk bandgap (see the supplementary materials, section 3J). Topological protection was also confirmed in a device with a domain wall with a 60° turn, and polaritons were successfully guided through the sharp turn without back scattering (see the supplementary materials, section 3I).

To investigate the spectral characteristics of the helical polarization of the polariton emission, the degree of circular polarization of the polariton emission was extracted as a function of energy and propagation distance along the interface (Fig. 4C; see the supplementary materials and methods). A large degree of helicity (33%) defined as $(\sigma^+ - \sigma^-)/(\sigma^+ + \sigma^-)$ was observed ~ 7.5 μm away from the excitation spot between ~ 1.98 and 2.05 eV, corresponding to the lower-branch helical interface polaritons. The degree of helicity is below unity because of the free exciton and trion emission excited directly by the pump laser, which leaks to the far field without coupling to the interface mode (see the detailed analysis in the supplementary materials, section 3K). The propagation length of helical polaritons at 2.023 eV is 8.0 ± 0.1 μm , which is mainly limited by the intrinsic far-field radiation loss of the photonic mode (see the supplementary materials, section 3K), purposely designed to enable the optical characterization of the system. Nonetheless, much longer propagation length can be achieved in devices with smaller photonic bandgap to reduce the radiation loss or by fabricating a valley Hall topological photonic crystal (30).

The polariton transport was further investigated in the momentum space through angle-resolved photoluminescence (Fig. 4, D and E)

measurements. At 200 K, photoluminescence from neutral A-excitons and negatively charged trions are observed at 2.055 and 2.015 eV, respectively. The strong coupling between the interface state and the neutral excitons can be identified by the anticrossing dispersion near the exciton energy and a redshift of the Dirac point from 2.060 eV in the uncoupled photonic mode, with the Rabi splitting Ω_{topo} of 32 ± 2 meV, confirming the strong coupling condition (see the supplementary materials, section 3F). The interface state and trions couple only weakly because of their smaller oscillator strength. To quantitatively investigate the degree of helicity of the topological interface exciton-polariton transport, the k -resolved line cuts at 2.032 eV corresponding to Fig. 4, D and E, are plotted in Fig. 4F. Both the interface modes corresponding to σ^+ and σ^- polarization fit to a single peak exclusively from the branch with positive and negative group velocities, respectively, thus demonstrating the helical nature of the topological polariton states.

Our approach of realizing strongly coupled topological polaritons opens the path to designing and eventually actively controlling the topological properties in a polaritonic system. Although the photonic topological devices are suitable for waveguides and transmission lines with low loss, by incorporating excitons, topological polaritons may find applications in low-threshold coherent light sources, nonlinear switchable and computational devices with precisely tailored dispersion and robust properties

immune to defects. The design methodology presented here can be readily generalized to other topological structures such as 1D and 2D Su-Schrieffer-Heeger systems and higher-order topological states, to explore the rich topological physics in the strong coupling regime.

REFERENCES AND NOTES

1. M. Z. Hasan, C. L. Kane, *Rev. Mod. Phys.* **82**, 3045–3067 (2010).
2. F. D. M. Haldane, S. Raghu, *Phys. Rev. Lett.* **100**, 013904 (2008).
3. T. Ozawa *et al.*, *Rev. Mod. Phys.* **91**, 015006 (2019).
4. Z. Wang, Y. Chong, J. D. Joannopoulos, M. Soljacic, *Nature* **461**, 772–775 (2009).
5. M. Hafezi, S. Mittal, J. Fan, A. Migdall, J. M. Taylor, *Nat. Photonics* **7**, 1001–1005 (2013).
6. S. Barik *et al.*, *Science* **359**, 666–668 (2018).
7. N. Parappurath, F. Alpeggiani, L. Kuipers, E. Verhagen, *Sci. Adv.* **6**, eaaw4137 (2020).
8. T. Ma, A. B. Khanikaev, S. H. Mousavi, G. Shvets, *Phys. Rev. Lett.* **114**, 127401 (2015).
9. M. A. Bandres *et al.*, *Science* **359**, eaar4005 (2018).
10. Y. Zeng *et al.*, *Nature* **578**, 246–250 (2020).
11. T. Karzig, C.-E. Bardyn, N. H. Lindner, G. Refael, *Phys. Rev. X* **5**, 031001 (2015).
12. A. V. Nalitov, D. D. Solnyshkov, G. Malpuech, *Phys. Rev. Lett.* **114**, 116401 (2015).
13. C.-E. Bardyn, T. Karzig, G. Refael, T. C. H. Liew, *Phys. Rev. B* **91**, 161413 (2015).
14. Y. Zhang, Y. V. Kartashov, A. Ferrando, *Phys. Rev. A* **99**, 053836 (2019).
15. Y. V. Kartashov, D. V. Skryabin, *Phys. Rev. Lett.* **119**, 253904 (2017).
16. S. Klemmt *et al.*, *Nature* **562**, 552–556 (2018).
17. A. Janot, B. Rosenow, G. Refael, *Phys. Rev. B* **93**, 161111 (2016).
18. C. Li *et al.*, *Phys. Rev. B* **97**, 081103 (2018).
19. G. Hu *et al.*, *Nature* **582**, 209–213 (2020).
20. M. Jung, Z. Fan, G. Shvets, *Phys. Rev. Lett.* **121**, 086807 (2018).
21. X. Liu *et al.*, *Nat. Photonics* **9**, 30–34 (2015).
22. L. Zhang, R. Gogna, W. Burg, E. Tutuc, H. Deng, *Nat. Commun.* **9**, 713 (2018).
23. S. Dufferwiel *et al.*, *Nat. Commun.* **6**, 8579 (2015).
24. Y.-J. Chen, J. D. Cain, T. K. Stanev, V. P. Dravid, N. P. Stern, *Nat. Photonics* **11**, 431–435 (2017).
25. W. Liu *et al.*, *Nano Lett.* **16**, 1262–1269 (2016).
26. L.-H. Wu, X. Hu, *Phys. Rev. Lett.* **114**, 223901 (2015).
27. S. Barik, H. Miyake, W. DeGottardi, E. Waks, M. Hafezi, *New J. Phys.* **18**, 113013 (2016).
28. W. Liu *et al.*, *Nano Lett.* **20**, 1329–1335 (2020).
29. Materials and methods are available as supplementary materials.
30. T. Ma, G. Shvets, *New J. Phys.* **18**, 025012 (2016).

ACKNOWLEDGMENTS

Funding: R.A. is supported by the US-ARO (grant no. W911NF-16-2-0194), ONR-MURI (grant no. N00014-17-1-2661), the National Science Foundation (NSF-QII-TAQS grant no. 1936276 and NSF 2-DARE grant no. EFMA-1542879), and by a seed grant from MRSEC/DMR-1720530. Device fabrication work was performed at the Singh Center for Nanotechnology at the University of Pennsylvania, which is supported by the NSF National Nanotechnology Coordinated Infrastructure Program under grant no. NNCI-1542153. A.P. is supported by the National Natural Science Foundation of China (grant nos. U19A2090 and 51525202).

Author contributions: R.A. supervised the project; W.L. and R.A. conceived the idea and designed the experiments; W.L. and Z.J. fabricated and characterized devices with help from Y.W. and G.M.; W.L., Y.W., and Z.J. performed optical measurements; W.L. developed the theoretical model and performed numerical simulations with help from Z.J., Y.W., and M.H.; B.Z. and A.P. grew the monolayer WS_2 ; V.J.S. provided input on waveguide characterization; and W.L. and R.A. analyzed the data and wrote the manuscript. All authors discussed the results and contributed to the final manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/600/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S23
Tables S1 and S2
References (31–36)

28 April 2020; resubmitted 17 July 2020
Accepted 9 September 2020
Published online 8 October 2020
10.1126/science.abc4975

GEOPHYSICS

Surface deformation associated with fractures near the 2019 Ridgecrest earthquake sequence

Xiaohua Xu^{1*}, David T. Sandwell¹, Lauren A. Ward², Chris W. D. Milliner³, Bridget R. Smith-Konter², Peng Fang¹, Yehuda Bock¹

Contemporary earthquake hazard models hinge on an understanding of how strain is distributed in the crust and the ability to precisely detect millimeter-scale deformation over broad regions of active faulting. Satellite radar observations revealed hundreds of previously unmapped linear strain concentrations (or fractures) surrounding the 2019 Ridgecrest earthquake sequence. We documented and analyzed displacements and widths of 169 of these fractures. Although most fractures are displaced in the direction of the prevailing tectonic stress (prograde), a large number of them are displaced in the opposite (retrograde) direction. We developed a model to explain the existence and behavior of these displacements. A major implication is that much of the prograde tectonic strain is accommodated by frictional slip on many preexisting faults.

On 4 to 5 July 2019, the moment magnitude (M_w) 6.4 and M_w 7.1 Ridgecrest earthquake sequence ruptured a geometrically complex system of conjugate faults within the Walker Lane belt to the north of the eastern California shear zone (ECSZ), 17 km northeast of the city of Ridgecrest, California (1). Field scientists reported more than 5 m of dextral slip near the epicenter of the NW-SE M_w 7.1 rupture and 1 to 2 m of sinistral slip along the NE-SW M_w 6.4 rupture, whereas observations of distributed slip went largely undetected in rapid-response field surveys (2). Crustal deformation from the Ridgecrest earthquake sequence, however, was extremely well imaged by the twin Sentinel-1 radar satellites, providing an opportunity to examine high-resolution slip behavior of hundreds of very small crustal fractures in near real time. Not only did these distinctive remote-sensing observations aid early ground survey efforts, they also provide insight into candidate mechanisms of strain concentrations nearby the main rupture. Here, we investigate compelling evidence of discrete and distributed, yet contrasting, stress-induced surface fractures from two physically distinct deformation processes surrounding the Ridgecrest earthquake sequence: frictional sliding on preexisting faults and elastic deformation of compliant fault zones. These fractures are largely invisible to ground surveys and are only associated with infrequent major earthquakes.

Previous studies have documented small amounts of surface displacement on mapped faults nearby major large earthquakes (3–8).

Most of the triggered faults slip in the direction of the prevailing tectonic stress (prograde). This slip could be driven by dynamic stresses from elastic wave propagation or by permanent Coulomb stress changes. However, in a small number of cases associated with the 1992 Landers and 1999 Hector Mine earthquakes, the polarity (sense of offset across the fault) of the displacement was opposite to the prevailing stress (retrograde) (9–11). If these features represented true retrograde fault slip, the observations may imply that the relatively small stress release from the earthquake exceeded the tectonic stress, representative of a frictional slip model. An alternate explanation is that the retrograde displacement is due to strain localization on a compliant fault zone (10). If faults exhibit a relatively wide damage zone of low shear modulus (12), then the elastic strain will be concentrated in the damage zone and have a polarity that matches the Coulomb stress change (10, 11). This compliant fault zone model requires a preexisting fault that was damaged by historic ruptures.

The Sentinel-1 (14) and Advanced Land Observing Satellite-2 (ALOS-2) (15) satellites acquired high coherence interferograms that span the two Ridgecrest events. We generated phase gradient maps (Fig. 1A) (13, 16) [(17) and table S1] from stacks of the C-band Sentinel-1 data to identify 291 surface fractures. We then constructed east-west and up-down/south-north deformation maps from the difference and sum of stacked ascending and descending line-of-sight (LOS) displacement maps (17). We used these to measure the polarity and width of the lineaments' deformations. The polarities of 109 of the deformations are consistent with the prevailing tectonic stress (18), although 60 have well-resolved retrograde displacements (Fig. 1B). We could not classify 122 of the fractures.

To interpret these displacements, we developed a Coulomb stress forward model from a

detailed earthquake rupture model. We constructed the rupture model from an inversion of a compilation of geodetic data (19) [(17) and figs. S2 and S3]. Measurements of surface motion are provided by LOS displacements from two look directions of both the Sentinel-1A and -1B and ALOS-2 satellites and along-track interferometry from the ALOS-2 satellite. Global Navigation Satellite System (GNSS) offsets from 92 stations were used to constrain the large-scale deformation (20), where the GNSS product, which captured surface motion between the times of the two events, enabled a partial decomposition of the two ruptures, suggesting that there was substantial slip on both faults from each event. We constrained near-fault deformation in areas of possible interferometric synthetic aperture radar (InSAR) decorrelation with offsets estimated from Sentinel-2 imagery (21). Our slip model reveals a shallow slip deficit (SSD) (22) of 29%, which is larger than any recent major strike-slip earthquake in the ECSZ and Baja California (19) (i.e., 18% for Landers, 3% for Hector Mine, and 11% for El-Mayor Cucapah). This larger SSD might indicate the existence of a much thicker sedimentary layer or distributed shallow faulting over this region.

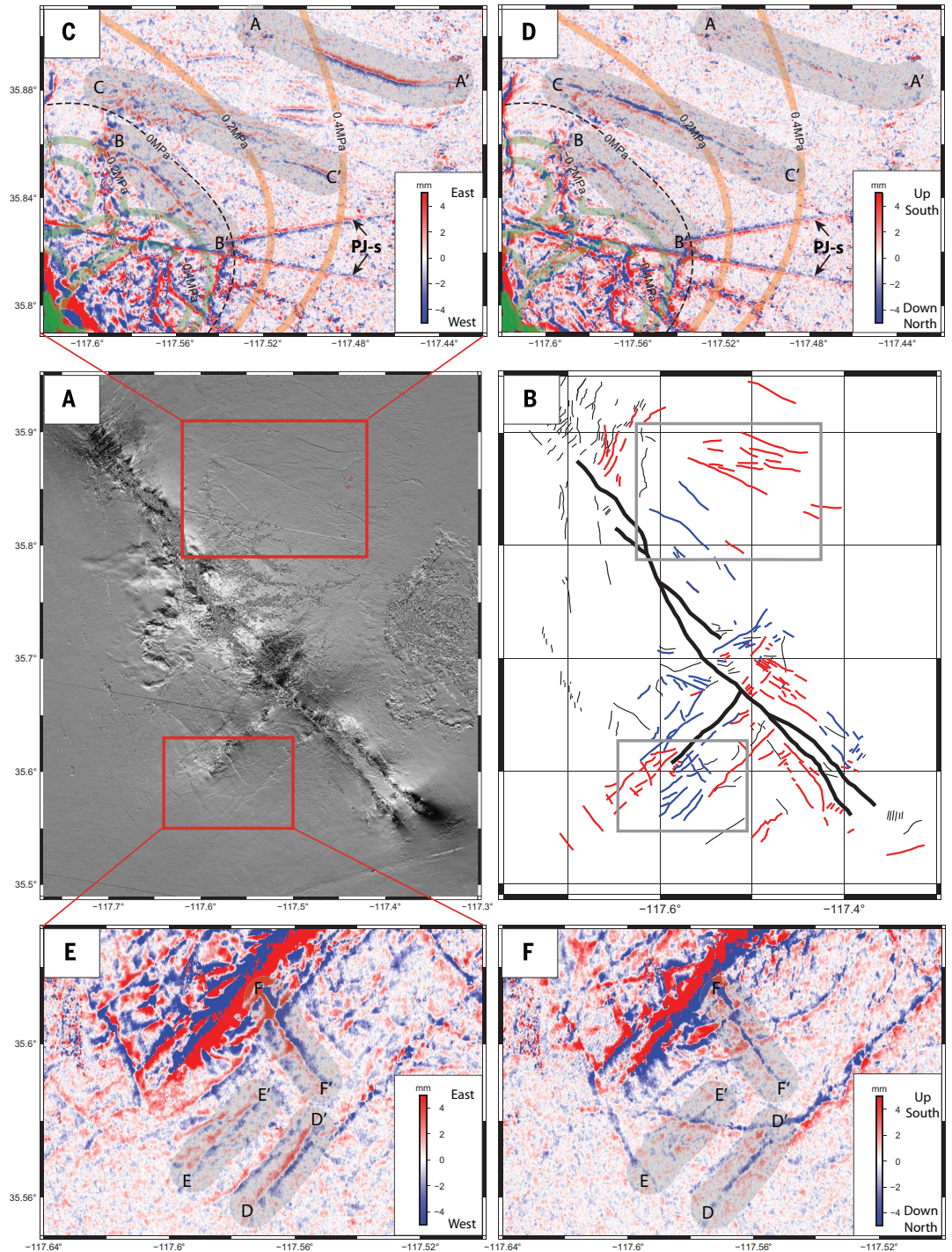
Using our slip model, we calculated tensor stress change at the surface of the Earth (23) [shear modulus (G) of 10 GPa] and projected it onto vertical planes parallel to each of the fracture segments to compute Coulomb stress change (fig. S4). We did this for a range of frictional coefficients (0.01 to 0.99) and background stress amplitudes (0.01 to 1 MPa). We assumed the regional tectonic deviatoric stress to be north-south compression and equal east-west extension (18). The polarity of the fracture displacements we observed (Fig. 1) was compared with the combined earthquake and regional deviatoric tectonic stress to achieve the best agreement (number matching to total). We found that the best agreement was a relatively small additional deviatoric tectonic stress of 0.4 MPa (fig. S4, bottom) and a low frictional coefficient (0.1). The transition in displacement polarity from prograde to retrograde is quite sensitive to the amplitude of this regional deviatoric tectonic stress. We speculate that this small background stress reflects only a portion of real surface stress and that the stress deeper in the crust is much higher.

Results from our analysis imply that the static stress change from the earthquake controls the polarity of the surface fractures. We then sought to determine if the observed deformation is from true fault slip due to the fault frictional strength being exceeded or an elastic response of a wide compliant fault zone due to the Coulomb stress change (Fig. 2A). The simplest compliant fault zone model has a damage zone with certain width (w) and lower

¹Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA. ²Department of Earth Sciences, University of Hawaii at Manoa, Honolulu, HI, USA. ³Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA.

*Corresponding author. Email: xix016@ucsd.edu

Fig. 1. Observations of fractures nearby the Ridgecrest earthquake sequence revealed by Sentinel-1 radar interferometry. (A) Phase gradient map from stacked Sentinel-1 interferograms revealing ~300 surface fractures around the Ridgecrest earthquake sequence region (13). (B) These fractures have been classified as prograde (red, 109 fractures), retrograde (blue, 60 fractures), and undetermined (black, 122 fractures) based on high-pass filtered (800 m) and decomposed (13) LOS deformation maps (upper and lower rows). The regions enclosed by the gray boxes cover exactly the same area as the red boxes in (A). (C and D) The north quadrant of the M_w 7.1 rupture has both prograde (A-A') and retrograde (B-B') fractures that occur in areas of positive (orange contours) and negative (green contours) Coulomb stress, respectively. Coulomb stress is computed with a receiver fault orientation aligned with the M_w 7.1 rupture (44° NW). Profile A-A' has a prominent right-lateral east-west signature but small up-down/south-north signature. Profile B-B' has a prominent left-lateral east-west signature as well as a substantial combined left-lateral south-north and/or downward up-down signature. Profile C-C' is transitional, right-lateral (prograde) at the eastern end and mainly up-down at the western end with slight right-lateral motion. (E and F) The south quadrant near the M_w 6.4 rupture has several right-lateral (retrograde) fractures (D-D' and E-E') as well as left-lateral (retrograde) fractures (F-F'). Both sets are in an area of negative Coulomb stress based on the respective fault orientations. Many of the fractures (D-D' and F-F') have prominent downward vertical signatures resulting from extensional stress. Note that the very straight lines in decomposed maps (C) and (D), which are marked by bold "PJ-s" and arrows, are not real fractures but radar phase jumps (PJ-s) across burst boundaries caused by the design of the Sentinel-1 radar and misregistration owing to azimuthal deformation from the earthquake.



shear modulus (G_l), embedded in a half space of higher shear modulus (G_h) (Fig. 2A). When a shear stress change ($\Delta\sigma_l$) is applied, a strain localization in the compliant zone appears as a step in the deformation profile (D) across the fault. From these parameters, we can estimate the ratio of the shear modulus in the

fault zone to the surrounding shear modulus (10, 24):

$$\frac{G_l}{G_h} = \frac{1}{\frac{DG_h}{w\Delta\sigma_l} + 1} \quad (1)$$

A shear modulus ratio of less than about 0.3 is

physically difficult to explain (25), whereas a very small contrast is equivalent to a very strong strain concentration, consistent with true fault slip. Compliant elastic strain is thus possible for larger modulus ratios.

We measured the amplitude and width of deformation for many profiles crossing each

Fig. 2. Response of a critically stressed fault zone to a change in stress from a nearby large earthquake.

(A) Schematic of a fault surrounded by a compliant zone of relatively low-shear modulus G_i in a half-space of normal shear modulus G_h .

(B) A stress change ($\Delta\sigma_i$) from a major nearby earthquake produces two types of surface deformation. Solid lines represent across-fault displacement profiles for the elastic half-space model shown in (A). Dashed lines represent the displacement

behavior for a compliant zone subject to a stress increase or decrease. (Left) A stress decrease with respect to the prevailing tectonic stress results in retrograde elastic deformation in the compliant zone. (Right) A stress increase will produce prograde deformation. Resulting fractures could thus be elastic deformation over the wide compliant zone or triggered frictional sliding over a narrow fault zone. (C) Example InSAR results from Fig. 1 corresponding to the two slip mechanisms illustrated in (B).

fracture (fig. S5) from the east-west deformation maps and projected the displacement onto the orientation of the fault to estimate the shear modulus reduction for each fracture (17) (Fig. 3). We found that the amplitudes of the measured prograde and retrograde displacements are similar, between 1 and 30 mm. However, the prograde deformation widths are systematically narrower (~100 m) than the retrograde deformation widths (200 to 500 m). This results in unreasonably small shear modulus ratios (less than 0.3) for the prograde deformations, compared with the shear modulus ratios of between 0.3 and 0.8 for the retrograde deformations. The best spatial resolution from Sentinel-1 Terrain Observation with Progressive Scan (TOPS) data is about 30 m, so prograde deformations are possibly narrower. From this analysis, we propose that the prograde deformations largely represent fault frictional slip, whereas the retrograde deformations are more likely due to elastic response of a compliant fault zone. These widths are much smaller than the ~1- to 2-km compliant fault zone width for retrograde displacements from the Hector Mine earthquake (10). Distribution of the mea-

sured and calculated parameters with respect to prograde and retrograde fractures present a composition of mechanism with analogy to a “ratchet” system, that is, the fault could deform in both directions, but true fault slip only occurs in the direction of prevailing tectonic stress.

Based on this analysis, we have developed a conceptual model for shallow displacement on faults in response to large ruptures (Fig. 2B and fig. S6). This region to the north of the ECSZ has a very high spatial density of conjugate faults that have slipped in the past in response to the regional tectonic stress (1). This interaction between the regional stress and earthquakes has been inferred to result in the faults being critically stressed, very close to failure by frictional sliding. The Ridgecrest earthquakes represent a major failure on two of these faults that subsequently changed the stress in the region. A Coulomb stress increment that is aligned with the preexisting tectonic stress increases the stress on the prograde faults to a state beyond their sliding strength, causing them to deform inelastically as frictional slip, especially at shallow depth where the overburden normal stress is low. A

Coulomb stress increment that is opposed to the preexisting tectonic stress lowers the stress on the retrograde faults, moving them away from failure, so they respond elastically over the wide damage zone of relatively low shear modulus. Between times of major earthquakes, the gradual increase in regional tectonic stress will bring the stress level of the retrograde faults closer to the critical level. The prograde faults could undergo aseismic slip or creep to maintain a critical stress level.

We examined two quadrants of the overall stress pattern where there is a transition between the prograde and retrograde faults and the changes in vertical deformation. The first quadrant is the region to the northeast of the epicenter of the M_w 7.1 event (Fig. 1, C and D). The dashed curved line at 0 MPa marks the transition of Coulomb stress along the M_w 7.1 rupture direction. One fault (A-A') east and north of this zone, where Coulomb stress is positive, has a clear and narrow right-lateral (prograde) deformation (east-west) and no vertical deformation (up-down), in agreement with the frictional slip model. West and south of the transition zone, there is a fault (B-B') that has a broad left-lateral (retrograde) deformation (east-west) and also a mixture of vertical (up-down) and left-lateral (south-north) deformation. Both deformation components are consistent with compliant zone fault slip behavior (10). Between these two faults, there is a fault (C-C') that has a transition from right-lateral slip and little vertical motion on the eastern end to slight right-lateral offset and prominent downward vertical motion on the western end inside the 0.2-MPa Coulomb stress change contour. This fault is a rare case where, with a small increase in Coulomb stress, this fault was able to maintain its continuum and not surpass its critical stress status to produce frictional failure. Retrograde fractures are also prominent south of the M_w 6.4 event (Fig. 1, E and F). There are two sets of retrograde fractures in this area: right-lateral fractures parallel to the M_w 6.4 rupture and left-lateral fractures perpendicular to the M_w 6.4 rupture. Both sets are in areas of negative Coulomb stress change for their respective fault orientations. A number of these faults also have downward vertical signatures related to extensional stress in this quadrant (fig. S1). One would need to develop a three-dimensional compliant fault model to better understand the combined strike-slip and vertical deformations seen in these areas.

We found substantial differences between the compliant fault zones imaged in our analysis and compliant zone images from the 1999 Hector Mine earthquake (9, 10). The typical widths of the deformation for the Roadman and Calico faults in the Hector Mine region were 1 to 2 km, whereas the retrograde deformation zones in the Ridgecrest region were

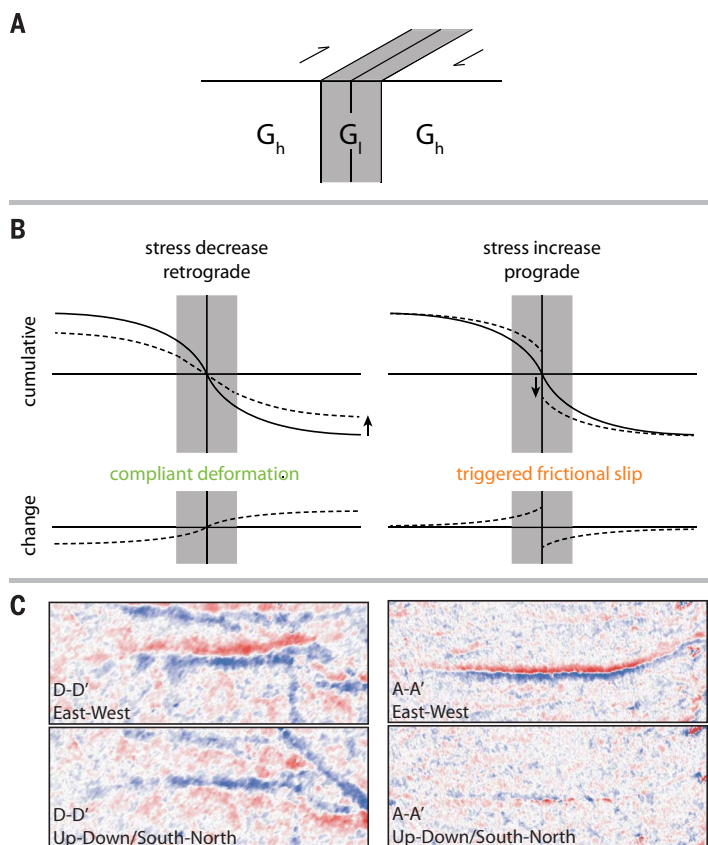
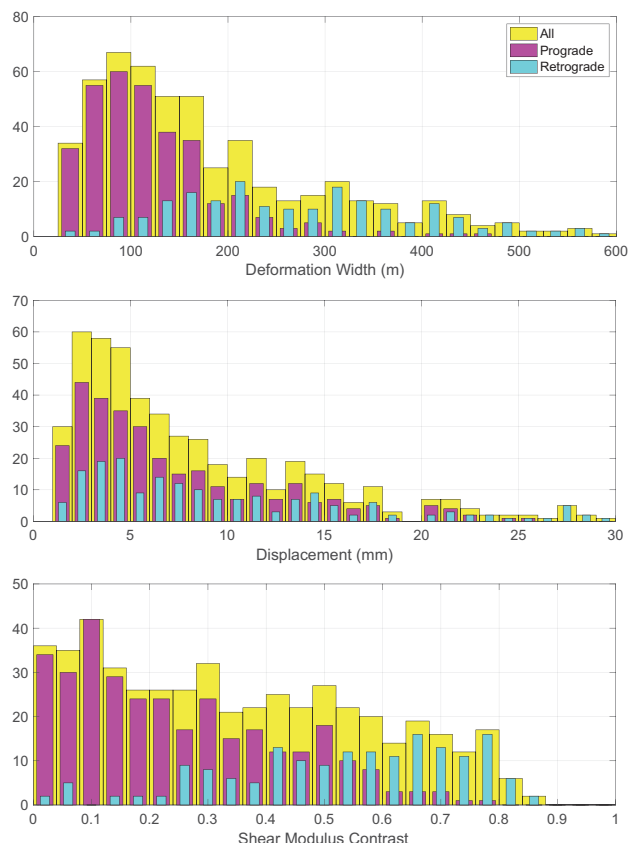


Fig. 3. Histograms of measured and computed parameters for the deformation profiles of the fractures. Show are measured deformation width (top), measured displacement amplitude (middle), and computed shear modulus contrast with Eq. 1 (bottom) for all fractures (yellow), prograde fractures (magenta), and retrograde fractures (cyan).



typically 200 to 500 m wide. Nevertheless, both areas have similar shear modulus reductions of 0.4 to 0.8. These are in agreement with the seismic velocity reduction observed on the San Andreas and San Jacinto faults (17). The difference in the widths of the damage zones may reflect the maturity of these faults (26), indicating that possible damage zones are well developed even at the faults' young age (27). Damage zone widths will increase with time as repeated rupture damage accumulates and becomes "fossilized" (26). By contrast, the amplitude of the shear modulus reduction may saturate at ~0.4, which explains the similarities between the two areas. The Rodman and Calico faults are relatively mature, having ~10 km of offset (28), whereas the small unnamed faults in the Ridgecrest area probably have much smaller offsets. Now that these faults have been mapped and characterized by Sentinel-1 InSAR, additional geological surveys will help to clarify their overall slip and thus maturity.

Compared with previous studies of deformations nearby large strike-slip earthquakes, the 2019 Ridgecrest earthquake sequence revealed an exceptional number of previously unmapped surface fractures. Whereas the prograde fractures might be new surface cracks that developed during the M_w 7.1 rupture, the retrograde compliant deformation must be associated with preexisting faults. These obser-

vations provided in situ records showing that very young and immature faults also have well-developed damage zones, which was previously thought to only occur for mature and repeatedly ruptured faults. Based on our analysis, all large earthquakes likely induce linear strain concentrations on preexisting faults. Stacking and phase gradient analysis of repeated images provides a new tool for discovering these faults, especially for areas with low strain accumulation. Such a large number of fractures were observed for the first time mainly because of the frequent, high-quality observations provided by the Sentinel-1 satellites, combined with the arid environment of the Mojave Desert region. These types of fractures may be ubiquitous to the ECSZ and many other diffuse strike-slip fault systems. Distributed as they are, these structures could account for the larger shallow slip deficit derived from the Ridgecrest earthquake sequence. A major implication of this result is that much of the "off fault" strain in the Mojave shear zone is due to permanent inelastic deformation on many small faults.

REFERENCES AND NOTES

1. Z. E. Ross *et al.*, *Science* **366**, 346–351 (2019).
2. D. J. Ponti *et al.*, *Seismol. Res. Lett.* **91**, 2942–2959 (2020).
3. P. L. Williams, S. F. McGill, K. E. Sieh, C. R. Allen, J. N. Louie, *Bull. Seismol. Soc. Am.* **78**, 1112–1122 (1988).
4. P. Bodin, R. Bilham, J. Behr, J. Gomborg, K. W. Hudnut, *Bull. Seismol. Soc. Am.* **84**, 806–816 (1994).
5. M. J. Rymer, *Bull. Seismol. Soc. Am.* **90**, 832–848 (2000).

6. T. Wright, E. Fielding, B. Parsons, *Geophys. Res. Lett.* **28**, 1079–1082 (2001).
7. X. Xu *et al.*, *Geophys. Res. Lett.* **45**, 10,259–10,268 (2018).
8. H. Goto, H. Tsutsumi, S. Toda, Y. Kumahara, *Earth Planets Space* **69**, 26 (2017).
9. D. T. Sandwell, L. Sichoix, D. Agnew, Y. Bock, J. B. Minster, *Geophys. Res. Lett.* **27**, 3101–3104 (2000).
10. Y. Fialko *et al.*, *Science* **297**, 1858–1862 (2002).
11. Y. Fialko, *J. Geophys. Res. Solid Earth* **109**, B03307 (2004).
12. A. A. Allam, Y. Ben-Zion, *Geophys. J. Int.* **190**, 1181–1196 (2012).
13. X. Xu, D. T. Sandwell, B. Smith-Konter, *Seismol. Res. Lett.* **91**, 1979–1985 (2020).
14. R. Torres *et al.*, *Remote Sens. Environ.* **120**, 9–24 (2012).
15. A. Rosenqvist *et al.*, *Remote Sens. Environ.* **155**, 3–12 (2014).
16. D. T. Sandwell, E. J. Price, *J. Geophys. Res. Solid Earth* **103**, 30183–30204 (1998).
17. Materials and methods are available as supplementary materials.
18. J. Townend, M. D. Zoback, *Geophys. Res. Lett.* **31**, L15S11 (2004).
19. X. Xu *et al.*, *Geophys. J. Int.* **204**, 1867–1886 (2016).
20. E. Klein *et al.*, *J. Geophys. Res. Solid Earth* **124**, 12189–12223 (2019).
21. C. Milliner, A. Donnellan, *Seismol. Res. Lett.* **91**, 1986–1997 (2020).
22. Y. Fialko, D. Sandwell, M. Simons, P. Rosen, *Nature* **435**, 295–299 (2005).
23. Y. Okada, *Bull. Seismol. Soc. Am.* **82**, 1018–1040 (1992).
24. P. Segall, *Earthquake and Volcano Deformation* (Princeton Univ. Press, 2010).
25. E. S. Cochran *et al.*, *Geology* **37**, 315–318 (2009).
26. J. F. Dolan, B. D. Haravitch, *Earth Planet. Sci. Lett.* **388**, 38–47 (2014).
27. H. M. Savage, E. E. Brodsky, *J. Geophys. Res. Solid Earth* **116**, B03405 (2011).
28. M. Oskin, L. Perg, D. Blumentritt, S. Mukhopadhyay, A. Iriondo, *J. Geophys. Res. Solid Earth* **112**, B03402 (2007).

ACKNOWLEDGMENTS

We thank T. Wright, K. Scharer, and an anonymous reviewer for their valuable suggestions and comments that strengthened this manuscript. **Funding:** This study was funded by the NASA Earth Surface and Interior Program (80NSSC19K1043), NASA Making Earth System Data Records for Use in Research Environments program (NNH17ZDA001N-MEASURES), the National Science Foundation (OAC-1834807 and EAR-1614875), and the Southern California Earthquake Center (SCEC) (19083). SCEC is funded by the NSF cooperative agreement EAR-1033462 and USGS cooperative agreement G12AC20038. **Author contributions:** X.X. analyzed InSAR images, performed source inversion and Coulomb stress analysis, computed the polarity and stress matching, performed profile analysis, calculated the histogram of parameters, worked on the conceptual model, and wrote the manuscript; D.T.S. analyzed InSAR images, performed profile analysis, worked on the conceptual model, and wrote the manuscript; L.A.W. initialized the Coulomb stress analysis and edited the manuscript; C.W.D.M. provided the optical imagery data to constrain the source model and edited the manuscript; B.R.S.-K. initialized the Coulomb stress analysis, worked on getting early InSAR acquisitions, and edited the manuscript; and P.F. and Y.B. analyzed and refined the GNSS offsets used in the source inversion. **Competing interests:** All authors declare no conflicts of interests. **Data and materials availability:** The Sentinel-1 images are copyright of the European Space Agency (ESA) and are freely available through the Sentinel-1 data hub, with data and orbital products archived and available at the Alaska Satellite Facility (ASF). The ALOS-2 images are copyright of the Japanese Aerospace and Exploration Agency (JAXA) and are provided under ALOS-2 RA6 research project 3071. GNSS offsets are processed and archived at Scripps Orbit and Permanent Array Center (SOPAC). All processed data are available from the authors on request.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/605/suppl/DC1
Materials and Methods
Figs. S1 to S6
Table S1
References (29–33)

4 June 2020; accepted 11 September 2020
10.1126/science.abd1690

MIGRATION

A complex phenotype in salmon controlled by a simple change in migratory timing

Neil F. Thompson^{1,2,3*}, Eric C. Anderson^{2,3,4*†}, Anthony J. Clemente^{2,3}, Matthew A. Campbell^{2,3,5}, Devon E. Pearse^{3,5}, James W. Hearsay⁶, Andrew P. Kinziger⁶, John Carlos Garza^{1,2,3*†}

Differentiation between ecotypes is usually presumed to be complex and polygenic. Seasonal patterns of life history in salmon are used to categorize them into ecotypes, which are often considered “distinct” animals. Using whole-genome sequencing and tribal fishery sampling of Chinook salmon, we show that a single, small genomic region is nearly perfectly associated with spawning migration timing but not with adiposity or sexual maturity, traits long perceived as central to salmon ecotypes. Distinct migration timing does not prevent interbreeding between ecotypes, which are the result of a simple, ancient polymorphism segregating within a diverse population. Our finding that a complex migratory phenotype results from a single gene region will facilitate conservation and restoration of this iconic fish.

Biodiversity exists on a continuum, but human perception of biological differences is often distorted by the cognitive benefits of categorization (1–4). Conspicuous phenotypic differences may serve as the basis for such categorization, but there can be a mismatch between the extent of perceived distinction and the biological basis that underpins it. In some cases, substantial phenotypic differences may be entirely environmental, with no genetic differences between individuals with alternative phenotypes (5, 6).

The seasonal timing of animal migrations is a particularly conspicuous trait, with indigenous peoples around the world using these migrations as the basis for customs related to subsistence and culture. Many human societies have historically depended upon the fishes in the family Salmonidae, which includes salmon, trout, and char, as a central component of their subsistence economies. Among the most well-known animal migrations are those made by anadromous salmonids, which travel up to 3000 km from the ocean to spawn in natal streams (7). In Chinook salmon (*Oncorhynchus tshawytscha*), the largest of the anadromous salmonids, multiple ecotypes exist and are characterized primarily by migration timing (8). Early-migrating fish are generally termed winter-run and spring-run ecotypes, whereas late-migrating fish include the fall-run and late-fall-run ecotypes (7). Early-migrating salmon return to fresh water up to 6 months before

late-migrating fish and hold during summer in cold-water pools until spawning. In addition, early-migrating fish have a suite of other prominent phenotypic differences, including smaller size, higher fat content, earlier spawning time, and the capacity to use different habitats (9–11), often at higher elevation, where they play a distinctive ecological role, conveying marine-derived nutrients and altering food webs. They are prized by fishers for their higher fat content and presence in fresh water when other salmon are not available. These traits have led many to regard the spring run as an entirely distinct type, or species, from the fall run (12). This sentiment mirrors the idea that migration timing is a “magic” trait that causes assortative mating, and potentially sympatric reproductive isolation, and which is also under disruptive selection (13).

Two broad themes have emerged in the study of seasonal animal migrations. First, successfully adopting new migration patterns requires a “syndrome” of physiological, morphological, and life-history traits to mediate the trade-offs involved with changes in migration (14, 15). Second, by exposing animals to new habitats, migration can promote reproductive isolation and speciation (16). Genomic studies accord with these two ideas. In birds, migration-linked genes have been implicated in divergence between subspecies with distinct migratory directions (17), and migratory forms of the same species can display different genomic and transcriptional patterns (18–20). In fishes, genomic variation associated with distinct migratory behaviors has been shown to be maintained in chromosomal inversions spanning hundreds of genes (21, 22).

Recently, a region on Chinook salmon chromosome 28 that contains two protein-coding genes, *GREBIL* and *ROCK1*, was found to be consistently associated with run-timing ecotypes in multiple drainages of North America (23–25). It has been suggested that this association is due to genetic adaptation for in-

creased fat storage to offset early migrators’ extended lack of feeding in fresh water (23). This migration phenotype has also been linked to reproductive maturity, with the adoption of “premature” and “mature” migrator nomenclature (23). As changes in migration have precipitated evolution in other key traits, it has been suggested that this trait variation is leading to reproductive isolation sufficiently profound that early- and late-migrating salmon should be categorized and managed as separate species (26).

We used whole-genome sequences, targeted genotyping assays, and individual-specific phenotype information from an indigenous fishery to elucidate the genetic architecture of ecotypic differentiation in Chinook salmon and determine the origin of phenotypic variation between ecotypes in these iconic fish. We focus on the Klamath River, where massive restoration efforts, including the removal of four mainstem dams, portend a renaissance for anadromous fishes. We also evaluate the role of genomic variation in ecotypic differentiation in the Sacramento River basin (Fig. 1), which harbors the greatest salmon run-type diversity known.

We resequenced the whole genomes of 160 fish from all ecotypes of Chinook salmon in the Klamath and Sacramento river basins (two of the largest rivers tributary to the Eastern Pacific Ocean), including fish from winter-, spring-, fall-, and late-fall-run ecotypes (Fig. 1, table S1, and data S1). Fish in these basins are not sister lineages evolutionarily, despite their geographic proximity (27, 28). Notably, in a combined analysis of these lineages, fixed differences between 64 fall-run and 64 spring-run fish were found only in a single ~140-kb region on chromosome 28 (Fig. 2A). Within this region, a smaller ~30-kb region, the “region of strongest association” (RoSA), included variants that were also fixed in the winter run, an early-migrating ecotype endemic to the Sacramento River (Fig. 2B).

Variation in the RoSA is organized in distinct, reciprocally monophyletic early- (E) and late- (L) migrating haplotype lineages (Fig. 3). Notably, the Sacramento basin possesses two divergent haplogroups within both the E and L lineages, whereas the Klamath basin has only one per lineage. Early-migrating, spring-summer-run Chinook salmon from the Upper Columbia River (Idaho, USA) are highly divergent from all other Chinook salmon (27–29) and carry a similar, but nonidentical, E-lineage haplotype in the RoSA (figs. S1 and S2) (25). This suggests that the E-lineage haplotype may be shared by all early-migrating Chinook salmon. We aligned the RoSA with the homologous region in its closest relative, coho salmon (*O. kisutch*), and estimate that the divergence between the E and L haplogroups is ~6% of that between coho and Chinook

¹Department of Ocean Sciences, University of California, Santa Cruz, Santa Cruz, CA 95064, USA. ²University of California, Santa Cruz, Institute of Marine Sciences, Santa Cruz, CA 95064, USA. ³Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, Santa Cruz, CA 95060, USA. ⁴Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO 80523, USA. ⁵Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, Santa Cruz, CA 95064, USA. ⁶Department of Fisheries Biology, Humboldt State University, Arcata, CA 95521, USA.

*These authors contributed equally to this work.

†Corresponding author. Email: eric.anderson@noaa.gov (E.C.A.); carlos.garza@noaa.gov (J.C.G.)

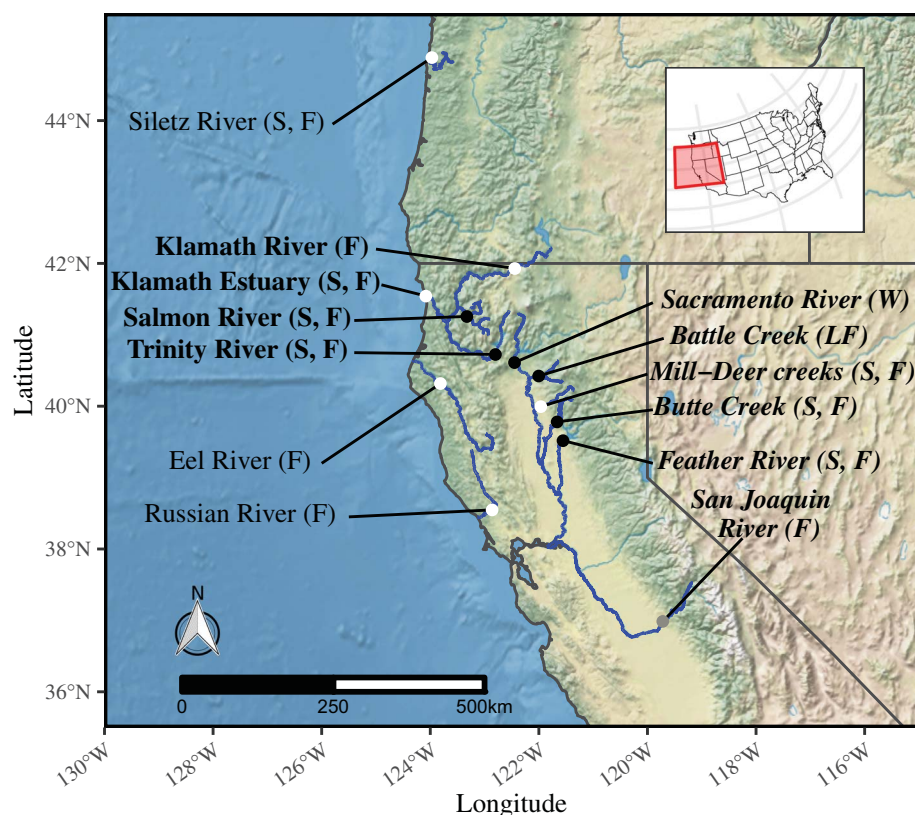


Fig. 1. Waterways from which Chinook salmon were sampled. Ecotypes are represented by letters (W, winter; S, spring; F, fall; LF = late-fall). Collections in bold are from the Klamath River basin, and names in italic-bold are from the Sacramento River basin. White circles indicate amplicon data; gray circles, whole-genome sequence data; and black circles, both data types.

salmon, indicating that the split occurred relatively recently within the Chinook salmon lineage (fig. S3).

Two of the fixed differences in the RoSA are nonsynonymous, single-nucleotide polymorphisms (SNPs) within *GREBIL* and show a near-perfect association between early- and late-migrating genotypes (table S2). *GREBIL* is a central regulator of vertebrate development, specifically affecting renal, gonadal, and inner ear organ systems (30, 31). Additionally, there are structural changes—short duplications between *GREBIL* and *ROCK1* (fig. S4) that might affect gene regulation and are strongly associated with distinct haplogroups (figs. S5 and S6). The structural differences might interact with the nonsynonymous SNPs or other variants to influence the phenotype. Evident within the RoSA is a block-like structure of allelic variation, high conservation, and linkage disequilibrium (Fig. 2B), but evaluation of sequences from this region did not yield evidence of a large inversion. The high conservation in this region may be attributable to its function as a small supergene, with multiple coadapted variants contributing to the phenotypic effects and large fitness differences associated with recombinant haplotypes.

We evaluated the effect of RoSA genotype on the phenotype of 502 adult Chinook salmon harvested by the Yurok Tribe in the Klamath River Estuary (table S1). We assayed eight SNPs (table S3 and data S2) that perfectly tag the E and L haplogroups (fig. S2) and compared the RoSA genotype to the phenotypic traits of freshwater entry date, fat content, and reproductive maturity. We found that the RoSA genotype predicted 85% of the variance in freshwater entry timing (table S4), with nearly disjunct freshwater entry windows among EE and LL genotypes (Fig. 4A). Heterozygotes enter with intermediate timing, on average, but overlap completely with the timing of homozygous fish. Thus, the RoSA genotype directly influences individual freshwater entry timing and, by extension, the initiation of spawning migration during the ocean phase. It may be that the RoSA influences individual response to photoperiod, as this is used by most organisms, including salmon (32), to detect seasonal changes, but other physiological processes are likely also involved.

By contrast, the RoSA genotype had no significant effect on maturation status or adiposity at freshwater entry [Fig. 4, B and C; likelihood ratio tests (LRT), table S5A], after accounting for

significant effects of sampling date, sex, and collection year (LRT, table S5, B and C). The strongest effect on maturation status and adiposity was due to sampling date, with partial correlations of 0.53 and -0.26 , respectively. Notably, all estuary-sampled fish had maturation-status values well below those of spawning fish (fig. S7), contrasting with the recent categorization of fall-run salmon as “mature migrators” (7, 23, 24).

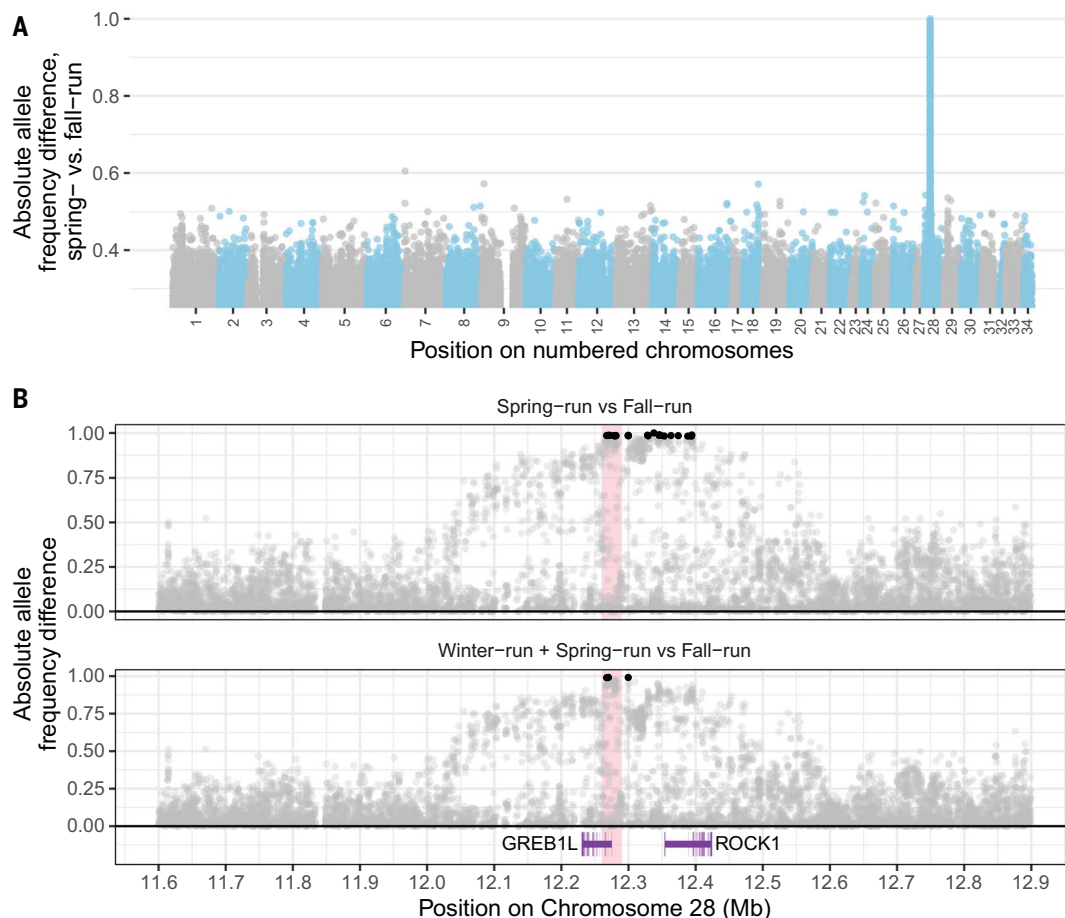
The RoSA genotype explained 67% of the variance in spawn timing at Trinity River Hatchery (table S6), where EE and LL individuals have a mean difference in spawn timing of 45 days (fig. S8). However, because the RoSA genotype does not affect maturation status at freshwater entry, this difference is likely a consequence of different freshwater residence durations, with warmer freshwater temperatures (33), relative to the ocean (fig. S9), accelerating maturation in early-entering fish during summer months. Water temperature is known to influence the rate and onset of final maturation in fishes, including salmonids (34, 35). Thus, although the reproductive physiology of spring- and fall-run fish is indeed distinct, validating indigenous knowledge, the RoSA directly influences freshwater entry timing, with other aspects of the spring-run “syndrome” (adiposity and maturity) attributable to differences in sampling date and environmental factors experienced as a consequence of differences in freshwater entry timing. We thus provide a counterexample to the widely held notion that differences in lipid metabolism and storage are central to migratory biology in vertebrates.

We next characterized the extent to which spawning behavior contributes to assortative mating and distinct populations in natural areas where both ecotypes persist, by evaluating differences in mating location and timing by RoSA genotype in the Salmon River (Klamath River tributary, which is mostly wilderness). The distribution of RoSA genotypes from 183 post-spawning carcasses indicates substantial overlap and opportunity for interbreeding among RoSA genotypes (figs. S10 and S11). This finding is bolstered by a large number of heterozygotes (table S7) and RoSA genotype frequencies close to those expected with random mating (fig. S12). As such, matings between heterozygotes must occur with some frequency, leading to full-sibling families that express both early- and late-migrating phenotypes.

We expanded our RoSA-marker survey to populations ranging from Coastal Oregon to the Sacramento basin (Fig. 1 and table S1). Heterozygous (EL) fish were widespread where early-migrating fish occur and suitable habitat for them exists (table S8 and data S3), and not only where hatchery propagation, or other anthropogenic influences, maintain them. However, we investigated whether spring-run and

Fig. 2. Allele frequency differences between ecotypes of Chinook salmon.

(A) Absolute value of allele frequency differences between 64 spring-run and 64 fall-run Chinook salmon. Each point is a genomic variant (differences <0.25 not shown). Colors alternate by chromosome. **(B)** Absolute value of allele frequency differences in a 1.3-Mb region around the peak in (A). Black points indicate nearly fixed (>0.98) differences. Vertical pink rectangle shows the RoSA. Violet horizontal bars at bottom show two nearby genes and exons within them (thin vertical bars).



fall-run fish were historically reproductively isolated via a simulation analysis of recombinant genotypes between previously reported SNPs in the imperfectly associated flanking region (23) and the RoSA (figs. S2 and S13 and table S3). We found that there is a long history of interbreeding between spring- and fall-run Chinook salmon in the Klamath River, consistent with earlier work (36), and that the high frequencies of recombinants (table S9 and data S4), necessarily generated in heterozygous fish, could not have arisen solely in the ~180 years since large-scale human manipulation of the watershed began (fig. S14 and table S10).

Disruptive selection between salmon ecotypes has been hypothesized to be strong, with heterozygotes particularly vulnerable (23, 24). Our data indicate that the RoSA has a partially dominant or additive inheritance pattern in the Klamath basin. Heterozygotes have migration timing that is skewed toward the early-migrating ecotype but overlaps entirely with the two homozygous classes. Moreover, we find that RoSA heterozygotes in the Sacramento basin are phenotypically more similar to late-migrating, than early-migrating, ecotypes for traits related to spawning (table S8), indicating that the dominance relation-

ship of the RoSA may be lineage specific and influenced by modifier loci. Unidentified modifier genes of smaller effect are known to influence within-season variation in migration timing of salmonids (37). Ultimately, it is clear that with temporally variable selection, fish with intermediate migration timing (heterozygotes) will periodically have equal or higher fitness than fish migrating earlier or later.

The importance of early-migrating fish to indigenous peoples and ecological functioning cannot be overstated, and early-migrating haplotypes are undoubtedly an important component of diversity in ecosystems and populations where they exist. Early-arriving fish are a critical early-season food source for indigenous peoples, and the early-migrating phenotype allows fish to exploit habitats that are less accessible to late-migrating fish (10), providing increased fitness and resilience to salmon populations, as well as important ecological effects, including deposition of marine-derived nutrients and altered food webs. In some cases, such as the upper Columbia River, early- and late-migrating lineages coexist with little gene flow, are not closely related, and evolved separately over long periods of allopatry (27–29). In general, however, fall- and spring-run Chinook salmon interbreed,

and the primary axis of genomic variation is associated with geography. As such, genome-wide adaptation to local environmental conditions is shared by sympatric spring- and fall-run ecotypes (27, 29, 36), and the gene flow between them provides critical connectivity for the maintenance of genetic diversity and long-term viability, consistent with a more inclusive definition of salmon populations (38–40).

The finding that the RoSA E-lineage haplogroup is conserved across Chinook salmon lineages and evolutionarily significant units (i.e., “species” under the U.S. Endangered Species Act) is a positive development for conservation, indicating that the most important diversity related to the evolution of ecotypes has not been lost with widespread extirpation of early-migrating salmon. As restoration efforts continue, the reestablishment of early-migrating populations will be facilitated by the evolutionary conservation and exchangeability of the E-lineage haplogroup. In addition, this highlights the importance of maintaining migratory opportunities and viable early-migrating source populations for recolonization.

In the Klamath River, the largest fish restoration project in history, removal of four dams that have blocked access to the upper basin

Fig. 3. Two hundred and two biallelic variants (columns) within the RoSA (chromosome 28, 12.26–12.29 Mb). Variants were from 202 haplotypes (rows) derived from whole-genome resequencing of 146 fish, with gold indicating the allele at highest frequency among spring run and blue the allele at lowest frequency among spring run. The local coalescent tree is at left, with the length of portion excised from the internal branches shown at top. Ecotype of the fish is in the first column right of the tree. The RoSA Zygosity column indicates whether the fish carried both an E- and an L-lineage haplotype (heterozygous) in the RoSA.

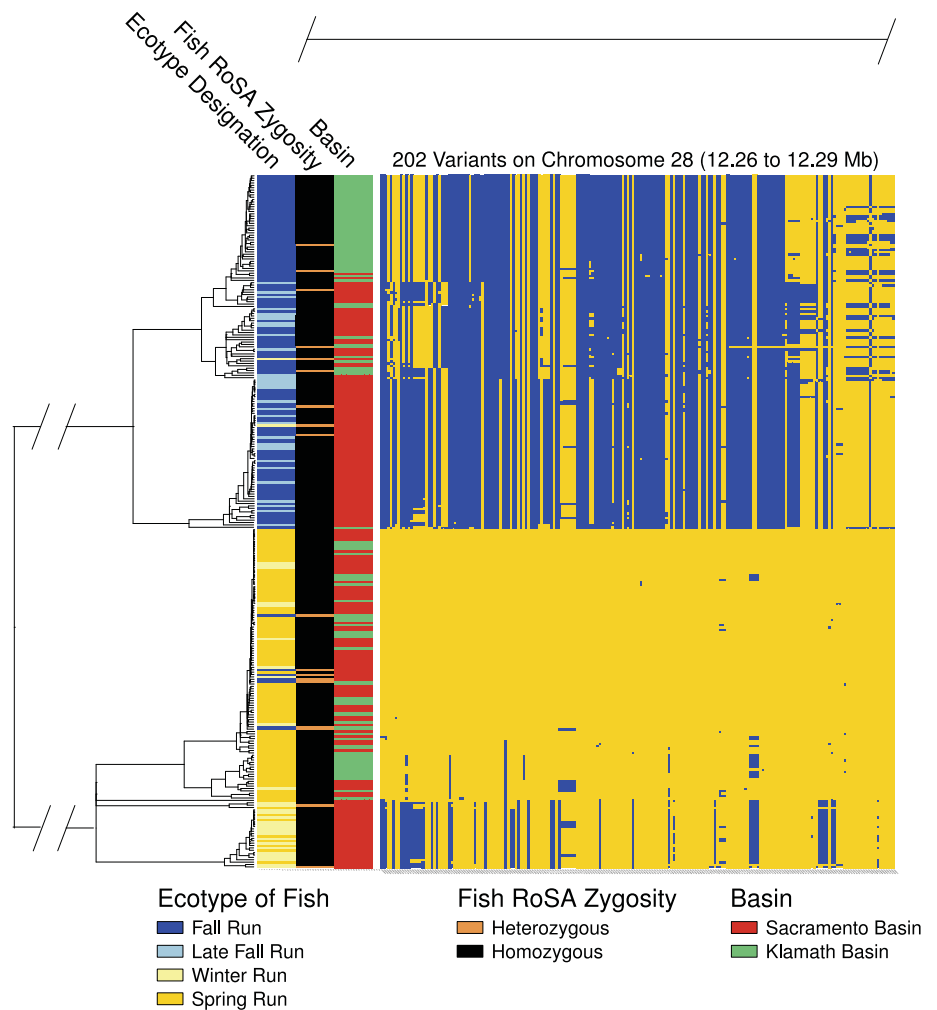
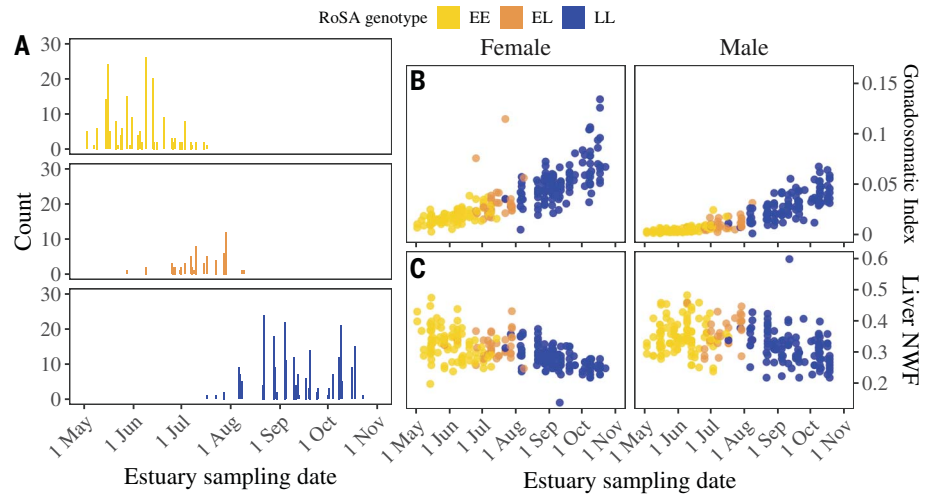


Fig. 4. Migration timing, physiological status, and RoSA genotype of Klamath River Chinook salmon. (A) Distribution of freshwater entry timing by day and RoSA genotype of 502 Chinook salmon sampled in the Klamath River estuary. (B) Maturation status, and (C) adiposity of estuary fish by sampling date and color coded by RoSA genotype. Data from 2009 and 2010 are combined.



(~22,600 km²) since 1912 (41) is imminent. In populations where the E haplotype occurs, it is highly likely that an early-run phenotype can emerge given suitable habitat. We found the E haplotype to be absent from adult returns to Iron Gate Dam (table S8), the current terminus

of anadromy in the Klamath River, indicating extirpation of the ancestral spring run from the upper Klamath basin. The lack of cool water required for summer survival in the mainstem Klamath River below Iron Gate Dam selects against both EE and EL genotypes during the

summer months and is likely responsible for this absence (33). However, our analysis of recombinants shows that descendants of the extirpated spring run persist below Iron Gate Dam, likely maintaining much of the genetic variation, outside of the RoSA, that resulted

from local adaptation to the upper Klamath basin. In addition, the E lineage remains abundant in the Klamath River basin (tables S8 and S11) but mostly as a consequence of the Trinity River Hatchery (36), which is ~450 river km from the dam removal site. As such, reconstituting a historically accurate, locally adapted, spring-run is possible by migration of E haplotypes into the population currently residing below Iron Gate Dam. Given the strong pattern of equilibrium between migration and genetic drift in the basin (40), E haplotypes will arrive in the upper Klamath basin through intrabasin migration, but this could be facilitated more rapidly by human-assisted translocation.

An association between ecotypic designation and a genomic region homologous to the RoSA has also been demonstrated in steelhead trout (*O. mykiss*), with similar ecological importance associated with the early-migrating ecotype (23, 42, 43). In addition, the three primary genera in the subfamily Salmoninae all have distinct early- and late-migrating forms that co-occur (7). We show that a complex migratory phenotype arises from a single gene region, which may facilitate its existence on multiple genetic backgrounds, and suggests that a similar, simple control of migratory traits may exist in other genera. Further research will elucidate how universally such apparent complexity and ecologically important phenotypic variation arises from a simple Mendelian polymorphism.

REFERENCES AND NOTES

1. J. Winawer *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7780–7785 (2007).
2. B. Park, M. Rothbart, *J. Pers. Soc. Psychol.* **42**, 1051–1068 (1982).
3. D. Roberson, J. R. Hanley, *Curr. Biol.* **17**, R605–R607 (2007).

4. J. W. Howard, M. Rothbart, *J. Pers. Soc. Psychol.* **38**, 301–310 (1980).
5. A. D. Bradshaw, in *Advances in Genetics* (Elsevier, 1965), vol. 13, pp. 115–155.
6. K. Gotthard, S. Nylin, *Oikos* **74**, 3–17 (1995).
7. T. P. Quinn, P. McGinnity, T. E. Reed, *Can. J. Fish. Aquat. Sci.* **73**, 1015–1030 (2016).
8. C. Groot, L. Margolis, *Pacific Salmon Life Histories* (UBC Press, 1991).
9. J. O. Snyder, *Salmon of the Klamath River, California* (California State Printing Office, 1931).
10. P. B. Moyle, *Inland Fishes of California: Revised and Expanded* (Univ. of California Press, 2002).
11. J. W. Hearsey, A. P. Kinziger, *Environ. Biol. Fishes* **98**, 413–423 (2015).
12. K. Langin, *Science* **360**, 590–592 (2018).
13. R. S. Taylor, V. L. Friesen, *Mol. Ecol.* **26**, 3330–3342 (2017).
14. H. Dingle, *J. Ornithol.* **147**, 212–220 (2006).
15. T. Ålerstam, A. Hedenström, S. Åkesson, *Oikos* **103**, 247–260 (2003).
16. S. P. Turbek, E. S. C. Scordato, R. J. Safran, *Trends Ecol. Evol.* **33**, 164–175 (2018).
17. K. Ruegg, E. C. Anderson, J. Boone, J. Pouls, T. B. Smith, *Mol. Ecol.* **23**, 4757–4769 (2014).
18. A. M. Fudickar *et al.*, *Biol. Lett.* **12**, 20160069 (2016).
19. R. A. Johnston, K. L. Paxton, F. R. Moore, R. K. Wayne, T. B. Smith, *Mol. Ecol.* **25**, 5680–5691 (2016).
20. M. Lundberg *et al.*, *Evol. Lett.* **1**, 155–168 (2017).
21. T. Kess *et al.*, *Sci. Adv.* **5**, eaav2461 (2019).
22. D. E. Pearse *et al.*, *Nat. Ecol. Evol.* **3**, 1731–1742 (2019).
23. D. J. Prince *et al.*, *Sci. Adv.* **3**, e1603198 (2017).
24. T. Q. Thompson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 177–186 (2019).
25. S. R. Narum, A. Di Genova, S. J. Micheletti, A. Maass, *Proc. Biol. Sci.* **285**, 20180935 (2018).
26. *Fed. Regist.* **83**, 8410–8414 (2018).
27. P. Moran *et al.*, *Can. J. Fish. Aquat. Sci.* **70**, 415–435 (2013).
28. A. J. Clemente, E. D. Crandall, J. C. Garza, E. C. Anderson, *Fish. Bull.* **112**, 112–130 (2014).
29. R. S. Waples, D. J. Teel, J. M. Myers, A. R. Marshall, *Evolution* **58**, 386–403 (2004).
30. P. D. Brophy *et al.*, *Genetics* **207**, 215–228 (2017).
31. I. Schrauwen *et al.*, *Hum. Genet.* **137**, 459–470 (2018).
32. T. P. Quinn, *The Behavior and Ecology of Pacific Salmon and Trout* (Univ. of Washington Press, 2018).
33. C. Z. Romberger, S. Gwozdz, “Performance of water temperature management on the Klamath and Trinity Rivers, 2017” (Arcata Fisheries Data Series DS 2018-59, U.S. Fish and Wildlife Service).
34. N. W. Pankhurst, H. R. King, *J. Fish Biol.* **76**, 69–85 (2010).
35. N. W. Pankhurst, M. J. R. Porter, *Fish Physiol. Biochem.* **28**, 385–389 (2003).
36. A. P. Kinziger, E. J. Loudenslager, D. G. Hankin, E. C. Anderson, J. C. Garza, *N. Am. J. Fish. Manage.* **28**, 1426–1438 (2008).
37. A. Abadía-Cardoso, E. C. Anderson, D. E. Pearse, J. C. Garza, *Mol. Ecol.* **22**, 4733–4746 (2013).
38. D. E. Schindler *et al.*, *Nature* **465**, 609–612 (2010).
39. S. M. Carlson, W. H. Satterthwaite, *Can. J. Fish. Aquat. Sci.* **68**, 1579–1589 (2011).
40. A. P. Kinziger, M. Hellmair, D. G. Hankin, J. C. Garza, *Trans. Am. Fish. Soc.* **142**, 1347–1357 (2013).
41. J. B. Hamilton *et al.*, *Oreg. Hist. Q.* **117**, 326–377 (2016).
42. J. E. Hess, J. S. Zandt, A. R. Matala, S. R. Narum, *Proc. Biol. Sci.* **283**, 20153064 (2016).
43. S. J. Micheletti, A. R. Matala, A. P. Matala, S. R. Narum, *Mol. Ecol.* **27**, 128–145 (2018).
44. E. C. Anderson *et al.*, 2020-chinook-salmon-migration-timing: Code/data repository for Thompson *et al.*, Version v1.2, Zenodo (2020); <http://doi.org/10.5281/zenodo.3836196>.

ACKNOWLEDGMENTS

We are extremely grateful to the Yurok Tribe for allowing us access to their tribal fishery. We thank the field biologists that sampled fish; C. Davis and M. Banks for Siletz River samples; E. Campbell, C. Columbus, and E. Correa for laboratory aid; R. Corbett-Detig for computational aid; E. Bjorkstedt and C. Edwards for temperature data; E. Palkovacs for helpful comments on the manuscript; and the UCLA IDRE for the Hoffman2 Shared Cluster. **Funding:** CIMEC (NOAA) award NA15OAR4320071; California Sea Grant/Delta Science Fellowship Program award R16AC00026; Hoopa Valley Tribal Fisheries Project N04351. **Author contributions:** N.F.T., E.C.A., A.J.C., M.A.C., D.E.P., A.P.K., and J.C.G. conceptualized the study. A.J.C. and M.A.C. curated molecular and whole-genome data. J.W.H. and A.P.K. curated estuary samples and associated data. N.F.T. and E.C.A. conducted formal analysis of the data. N.F.T., E.C.A., and J.C.G. wrote the manuscript, with input from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All genome sequence data are available from the NCBI Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/sra/PRJNA667732>). All other data and analysis code are available in GitHub, archived at Zenodo (44).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6516/609/suppl/DC1
Materials and Methods
Figs. S1 to S14
Tables S1 to S11
Data S1 to S5
References (45–97)

21 January 2020; accepted 18 August 2020
10.1126/science.aba9059



2020's Top Employers: Rapid response to COVID-19, diversity, and innovation

In a year like no other, 2020's Top Employers Survey not only highlighted the best working environments in pharma and biotech but also addressed some of the unique issues that arose this year, such as how to respond to a pandemic. By Alaina G. Levine

When an employee boasts that your business, **Vertex Pharmaceuticals**, is "the best company I've ever worked for," you tend to take note. This was just one of hundreds of comments from respondents to *Science's* 19th annual Top Employers Survey. Another observation concerning Vertex, a Boston-based firm that moved up four places from last year to No. 8, states that "Vertex puts its employees first. I am so proud to work here."

The Top Employers Survey was conducted by Cell Associates and Brighton Consulting. This year, the online survey took place from March 3 through May 3, 2020, and included approximately 7,600 respondents from across the world. Typically, the survey has highlighted pharma and biotech companies' commitment to innovation and progressive corporate culture as well as advanced technology and cutting-edge techniques, such as the use of CRISPR-Cas9 for genome editing and artificial intelligence, and machine learning for design, development, and manufacture of therapeutics and interventions. Not surprisingly, we saw these trends emerge again. But as we know, 2020 is unlike any other year. Given that the survey rolled out as quarantines were taking effect, respondents were able to provide a peek into the ambitious initiatives their companies were pursuing in response to COVID-19—which include everything from developing new work-at-home policies for family-focused employees to rapidly shifting corporate assets to support public health concerns and develop novel therapeutics.

For the companies that emerged in the top 20, remarks from respondents reflected their pride and gratitude in the fact that the organizations they represented had continuously invested in their well-being while still putting science and patients first. When employees see meaningful action by their employers that is designed to empower and support them in every way possible, they respond in kind: They produce their best work. And in the arena of pharma and biotech, that easily translates into better patient outcomes.

Gratitude is particularly strong among respondents whose corporations have appeared on the list before. "Being recognized is a great thing," says Hervé Hoppenot, CEO of **Incyte** (No. 2, advancing from No. 3 in 2019), a Delaware-based pharma company. "It means a lot to be on the list. Being at the cutting edge of science and having the best people want to work here [gives us] a sense of pride." John Frels, vice president of R&D at **Abbott**, a Chicago-based medical device and health care company that moved up to No. 14 this year from the No. 17 spot last year, notes the placement "validates what I have come to appreciate over my career: This is a company concerned about the long-term sustainability of delivering great value to our patients and customers, and it's a great place for scientists to apply their skills. We bring out the best in our scientists over the course of their career."

Having a work culture aligned with employee values is another important driver for the top employers and is referenced many times in the survey comments. Says one respondent, "Vertex innovates with speed and ferocity like no other company, while also putting a priority on culture that is amazingly open and supportive." An employee of **Syngenta**, the No. 4-ranked Swiss-based biotech firm that focuses on agrochemicals and seeds, notes what they consider to be their company's benefits: "Organizational culture, concern with the environment, concern with the well-being of employees." And a respondent referring to **Alnylam Pharmaceuticals**, headquartered in Cambridge, Massachusetts (No. 3), practically shouts their answer: "GREAT culture!! Collaborative! Inclusive! Exciting!"

DNA of top employers

What is the winning formulation that puts companies on the Top Employers list? Chief among the ingredients is an articulated mission of supporting scientists and science. Almost all of the top employers indicated that they are science- and patient-centered. And over and over again, the survey respondents echoed this philosophy.

"**Regeneron** is built around a science-first approach," says Drew Murphy, executive vice president of research at the American biotech organization based in Tarrytown, New York, which ranked No. 1 in this year's survey. "Unlike other

cont. >

POSTDOCTORAL TRAINING PROGRAM

PURSUE YOUR
PASSION FOR SCIENCE
IN A DYNAMIC
AND INNOVATIVE
ENVIRONMENT



Applications Accepted October 1 to December 1, 2020
Start dates flexible, but if possible, Summer 2021

- » Up to four years of formal postdoctoral training in a program co-Directed by **award-winning and published educators**, both of whom are working scientists at Regeneron **devoted to training postdoctoral fellows**
- » Exposure to the biotechnology industry's **most successful and innovative scientists**
- » **Weekly postdoctoral scientist meetings** to foster scientific rigor, creative thinking, critical reasoning, and career advancement
- » A **prestigious, well resourced, and highly driven** industry environment
- » Encouragement of **scholarly productivity** such as publications and conference presentations
- » Opportunities to pursue your own **cutting edge research ideas** in almost any area of biomedicine, including human genetics

LEARN MORE AT

www.regeneron.com/scienceeducation_professional

Regeneron is an equal opportunity employer. As such, applicants are considered for employment without regard to race, color, religion, sex, national origin, age, sexual orientation, marital status, disability, genetic information, military or veteran status, or any other category protected under applicable federal, state, or local law.

REGENERON
SCIENCE TO MEDICINE®

companies, our commercial people don't tell our researchers what to do. The scientists set the agenda. And if you do science the right way, you never really fail. You either succeed or learn something more valuable." The strategy is clearly working, as the attrition rate of its 8,600 staff (approximately half of whom are in R&D or administration) was less than half the industry average in each of the last 5 years. For example, Regeneron's 2019 turnover rate was 7.8% compared to an industry average of 18.7%, with turnover in its R&D organization ranking lowest of all employee groups. (The industry average is based on the Radford 2019 U.S. Workforce Trends Report for life sciences.)

In the life sciences sector, there is a symbiosis between science and patient priorities, and the top employers (and many survey respondents) emphasize this as a marker of a great company. This synergy manifests in multiple ways. **bluebird bio** (No. 20), for example, hosts Patient Days, in which scientists have the exceptional experience of interacting with those who directly benefit from their outputs and get a taste of the patient's journey. "You get a window into what it's like to be a patient dealing with the challenges we are trying to solve," says Philip Gregory, chief scientific officer of the Cambridge, Massachusetts-based firm that develops gene therapies for severe genetic disorders and cancer. "This is one way we connect the employees to the patient, so they can see why they are doing this—it grounds you and reminds you that you have a purpose beyond the one step you are aiming to do."

Abbott also places R&D in front, so that science serves as its fuel for constant growth, scaling, and sustainability. "There are ample opportunities to stay with the company," says Frels. "You can settle yourself into your career at Abbott, knowing that even if you move to different roles in the company, you can have confidence that what you will do will always be challenging and impactful in the long-term."

Access to professional development goes hand-in-hand with career advancement, of course, and the top employers are generous and proactive in designing and implementing programs that allow for skill building, networking, self-promotion, and leadership development. Abbott offers extensive training, including an engineering rotation program for new hires to learn about different divisions of the business. Its scientific employees are encouraged to produce individual development plans to map out their career, and there are plentiful opportunities for employees to move across departments, functions, and locations.

Language matters too. **bluebird** executives refer to their employees as "birds" and the firm itself as the "nest." They offer

Top Twenty Employers						
2020 Rank	2019 Rank	Employer (global headquarters)	Innovative leader in the industry	Treats employees with respect	Is socially responsible	Has loyal employees
1	2	Regeneron (Tarrytown, NY)	•		•	•
2	3	Incyte (Wilmington, DE)		•	•	•
3	1	Alnylam Pharmaceuticals (Cambridge, MA)	•	•	•	
4	10	Syngenta (Basel, Switzerland)	•		•	•
5	6	Biocon Limited (Bangalore, India)	•		•	•
6	4	Merck KGaA (Darmstadt, Germany)		•	•	•
7	20	Novo Nordisk (Bagsvaerd, Denmark)		•	•	•
8	12	Vertex Pharmaceuticals (Boston, MA)	•	•	•	
9	11	Moderna (Cambridge, MA)	•	•	•	
10	5	Spark Therapeutics (Philadelphia, PA)	•		•	•
11	8	Genentech (South San Francisco, CA)	•	•		•
12	13	AbbVie (North Chicago, IL)		•		•
13	16	GlaxoSmithKline (GSK) (London, United Kingdom)		•	•	•
14	17	Abbott (Abbott Park, IL)	•		•	•
15	14	Roche – excluding Genentech (Basel, Switzerland)	•		•	•
16	9	Eli Lilly and Company (Indianapolis, IN)		•		•
17	18	Pfizer (New York, NY)		•	•	•
18	–	Takeda Pharmaceutical Company Limited (Tokyo, Japan)		•	•	•
19	19	Novartis (Basel, Switzerland)	•		•	•
20	–	bluebird bio (Cambridge, MA)	•	•		•

The 20 companies with the best reputations as employers and the top three driving characteristics for each company, according to respondents in the 2020 survey undertaken for the Science/AAAS Custom Publishing Office.

The companies without a 2019 rank did not receive enough mentions to qualify or did not receive a high enough ranking from the 2019 survey.

The fifth attribute in the above table is "Has work culture values that are aligned with employees' personal values."

leadership development for all staff and have leadership coaches on-site to enable employees to go in the direction they want to. To facilitate better and more meaningful interpersonal messaging, the company uses a psychometric tool called Insights Discovery, which has a four-colored model to help people understand their personal style and preferences, including those related to communication. When you walk by someone's desk, you may see a stack of colored bricks on their cubical; the top brick signifies which communication behavior you lead with. For example, Cool Blue indicates a preference for data, structure and/or process. "It helps others to engage me, for example, in a way that is most effective. You understand who you are interacting with and their communication preference," says Gregory. "It creates a dialogue." And of course, better communication makes a better (and a top) company.

Flexibility is also important, and many companies actively demonstrate this trait, according to the survey. This is clear from the responses to COVID-19, as discussed below, but it is also seen in other ways, especially when it comes to harnessing the power of new and advancing technologies, such as data science. Vertex for example, leverages the kaleidoscope of data science benefits and applications across the enterprise, from human resources to legal to R&D. As David Altshuler, executive vice president, global research and chief scientific officer, explains, "We decided to build in an internal data science team and apply it broadly across the business. Each year the executive team picks several areas of the company to focus our data science efforts on, and we would assign it an innovative business leader. This is the secret sauce of Vertex—the urgency of making progress and seeing innovation and business go hand in hand: an alignment within the entire company."

Innovation mechanisms

One of the parameters that defines a top employer is its devotion to an innovation culture—and employees

cont. >



Creating the path to our vision

of a world where no life is limited by genetic disease.

There are no bounds to what we can do together.



Learn more about Spark Therapeutics at www.sparktx.com and through our social media channels on Twitter and LinkedIn.

© 2020 Spark Therapeutics, Inc. | N-UNBD-US-680007

notice innovation. In fact, one of the most common words survey respondents used to describe their employers was “innovative.”

Vertex, for example, has its VOICE Challenge, which engages employees so that “everyone is included in the innovation mission,” says Altshuler. This annual innovation tournament starts with identifying grand scientific and business challenges, which Altshuler describes as “things that would make a big difference to what we do.” He adds, “We invite the entire company to come up with ideas. Last year, we got 360 ideas.” Those ideas were suggested by 1,000 employees across all business units (making up about a third of the company). The top suggestions are assigned resources to develop, pitch, and design a program that advances them from conception to implementation. Some recent suggestions that have become reality include Vertex’s sickle cell disease and beta-thalassemia research programs, a mobile app and iPad kiosk system to help “Vertexians” find conference rooms and colleagues at their Boston headquarters, and a commuter bus program to reduce traffic congestion in Boston’s Seaport District.

“Innovation can come from everywhere,” says Natalie Hosea, site head for Takeda California and Drug Metabolism and Pharmacokinetics at **Takeda Pharmaceutical Company Limited** (No. 18). “Our scientists feel empowered to innovate in the space and feel comfortable with idea generation. We are a patient-first company: For every day it takes for an intervention to get to a patent, a patient is suffering. This motivates us.” The organization’s official innovation framework for research, Portfolio Entry, involves actively vetting ideas and testing hypotheses among its internal scientific community and external collaborators. Data is reviewed in a cross-functional manner, she adds, and feedback is provided as the R&D team reviews cases and narratives to decide what to invest in. The program is robust and fosters an important sense of organizational pride. And here, creativity begets creativity: Takeda focuses on four therapeutic areas, Hosea notes, but when an idea is generated that goes beyond those areas, the company aims high, actively pursuing unique licensing partnerships or other external arrangements—anything to get that solution to the bedside.

Novo Nordisk, an almost century-old Danish multinational pharmaceutical company with over 43,000 employees across the world and R&D centers in five nations (No. 7), recently initiated an R&D ideas challenge, which welcomes proposals for disruption from every employee. Recently, it received 500 proposals for transformational medications or diagnostics. “We were going to fund one, but we ended up funding the top five,” says Mads Krogsgaard Thomsen, chief scientific officer. “They get time off and work in an incubator environment to see if they can validate their idea. We give them the resources for do this for 6 months.” The company sees an appetite for repeating this process.

Biocon Limited (No. 5) takes a decidedly entrepreneurial approach to “impassioned innovation,” says Kiran Mazumdar-Shaw, founder and executive chairperson of this Bangalore,



* Driving characteristics are listed in descending order of impact on overall employer rankings.

** Shaded background indicates the characteristics in common for the 2 years.

India-based institution. Its victorious innovation program, Novel Biologics, acts like an incubator inside its R&D division, has regularly achieved key business and scientific benchmarks, including the creation of new assets within the portfolio of the company, and, not surprisingly, spin-off startups. In fact, one of the spin-offs has its own incubator. One example of their success with the Novel Biologics group is an immuno-oncology program focusing on development of novel bifunctional fusion antibodies, which is now housed in Biocon’s wholly owned subsidiary Bicara Therapeutics, based out of Boston.

Other companies, while still championing innovation, do not have formal programs to spur it. Regeneron’s leadership prefers to capitalize on organically generated discoveries. “Innovation is so rooted in the way we do everything, we don’t need an artificial mechanism to try and instill it,” says Murphy. “We don’t formalize this, and a lack of formality and being able to go off script allows us to pursue and pressure test ideas.” The culture at Regeneron is such that employees have autonomy to discuss hypotheses. “We don’t like people hiding the idea until they get all the data,” he says. “It’s like the Beatles’ lyric ‘Take a sad song and make it better’—you have to talk these things through. We encourage people to be generous with their ideas.”

COVID-19 responses

Not every industry gives its employees the privilege of being able to improve human health, and very few organizations are agile enough to grant their staff the opportunity to shift their focus and assets to confront an emerging plague. The pharmaceutical and biotech enterprises on this list are the notable exceptions. And those companies have quickly embraced the chance to serve humanity by fighting COVID-19.

Abbott has extensive experience in infectious disease diagnostics; it delivered the world’s first HIV test in the 1980s. Soon after SARS-CoV-2 was identified, Abbott’s scientists swung into action, initiating the fastest diagnostic product development campaign in the company’s history. “We leveraged next generation sequencing and informatics tools to rapidly design prototype tests,” explains Frels. “Our scientists collaborated quickly and effectively to help accelerate product development.” By the end of March, the teams had developed and launched laboratory and rapid point-of-care molecular diagnostic COVID-19 tests. This was followed quickly in mid-April with the launch of the first large-scale, high-throughput laboratory COVID-19 serology test. Since then, the teams have

cont. >

Science is resilient.
It can overcome diseases,
create cures and yes,
even beat pandemics.
It has the methodology
and the rigor
to withstand even
the most arduous scrutiny.
It keeps asking questions and
until there's a breakthrough,
it isn't done.
That's why, when the world
needs answers, we turn to science.
Because in the end
Science will win.



There's room to think big and be courageous. In fact, it's the norm. Join us.

<https://careers.pfizer.com>

top employers

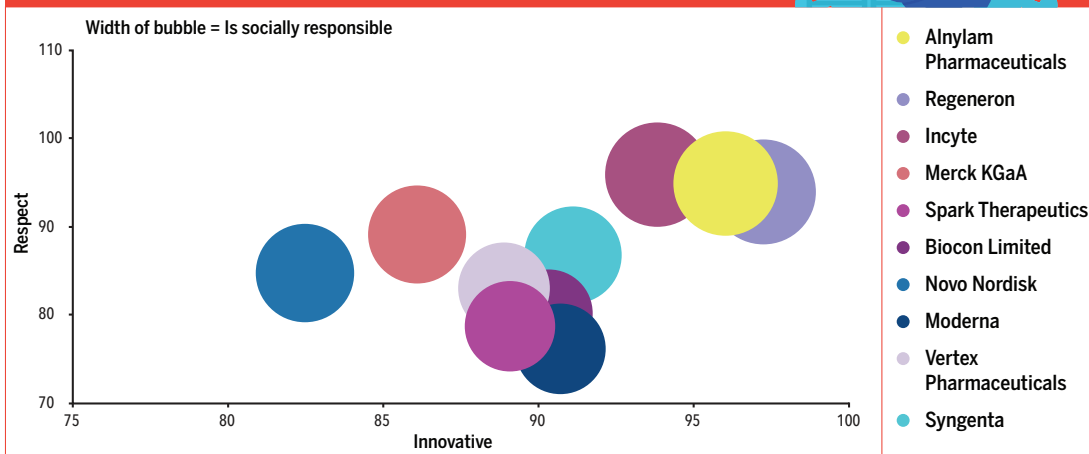
continued their work, developing and launching additional laboratory-based and rapid diagnostic tests. “We’ve had high-level management support and leadership through it all,” explains Frels, “and together with committed teams of staff from across Abbott working 24/7 shifts, we’ve made it happen.”

Incyte mobilized its resources to enable employees to work from home. “While working remotely after March 13th, we launched a new product and in addition realized that two of our established products could be helpful with the respiratory issues attached to COVID,” shares Hoppenot. “We had our teams on nights and weekends putting together a clinical program to send to the FDA [U.S. Food and Drug Administration] and get it approved. All of this was done remotely and was an enormous amount of work. Having an active role in the fight against COVID was important for all of us, and being able to do it while working remotely was even more motivating.”

Takeda reacted to COVID by focusing its efforts in five areas: research to address future pandemics; working with the CoVlg-19 Plasma Alliance to develop a plasma-derived hyperimmune globulin therapy for COVID; repurposing approved products or assets in development; data and information sharing; and establishing R&D partnerships. Takeda already has a culture that fosters external collaborations, so they were poised for this type of pivot. One of the company’s areas of expertise is in therapies derived from blood plasma. By May, the firm had joined with nine other biotechs to form the CoVlg-19 Plasma Alliance, described on its website as “an unprecedented partnership of the world’s leading plasma companies, spanning plasma collection, development, production, and distribution.” Furthermore, Takeda has taken a leadership role in the establishment of the COVID-19 R&D Alliance, which includes other top employers Alnylam Pharmaceuticals (No. 3), **AbbVie** (No. 12), **Pfizer** (No. 17), and **Novartis** (No. 19). One of the R&D Alliance’s projects is the development of an external-facing, data communication platform. “We are focusing on ways to share data and get it out to the public domain faster,” says Hosea. And in collaboration with this alliance, Takeda carefully examined their portfolio to see what approved products or assets in development could be pursued to treat COVID-19.

But COVID response didn’t just mean pumping out new medicines. At bluebird, it involved a lot of employee care. The company enacted an extra day off per month for every staff member, provided open Q and A sessions at open mics to converse with the entire leadership team on a regular basis, and organized support programs for “baby birds” (the children of employees) to help parents and caregivers—actions that were all in line with its core values.

COMPARISON OF TOP TEN COMPANIES BY TOP THREE CHARACTERISTICS



Comparison of the top 10 companies on the basis of 3 of the top drivers (scored out of 100):
Is an innovative leader in the industry (x-axis), Treats employees with respect (y-axis), and Is socially responsible (bubble width).

Diversity and social justice

While these companies have waged war against a microscopic virus, a macroscopic issue has come to light, as social justice conversations in the United States have reinforced many organizations’ commitment to provide nurturing, inclusive spaces for all employees. Diversity, equity, and inclusion (DEI) programs are closely tied to company cultures. Scientists look for such programs as signs of a place where they can thrive, where their inputs are seen, and where their voices are heard. Over and over again, survey respondents mentioned their gratitude to employers for building programs that buoy these critical efforts.

But there is still work to be done. “There is a lack of diversity in this industry,” admits Hoppenot. “Racism and science do not go together. If you are to be successful in science, we cannot have racism.” Hoppenot, like many leaders of Top Employer companies, is committed to having the difficult but necessary discussions about race, inclusivity, and diversity, and is already making changes in the company to advance this mission. The firm’s leadership is reaching out to organizations to assist it with improving the recruiting and mentoring of underrepresented minorities, especially African Americans, in its employee ranks. “We have always had this as a top priority, but we realized we needed to include more people in the discussion. To that end, our Inclusion Team is actively working to identify and implement initiatives that increase inclusion and also provide development opportunities for our employees,” says Paula Swain, Incyte’s executive vice president of human resources. “We don’t want to look at this as a moment in time. Investing in inclusion, mentoring, development, and retention will be part of what we will do as an organization now and in the future.”

bluebird bio is one of several organizations that has instituted DEI practices throughout its structure. “Our philosophy can be summarized as ‘all birds fly further together.’ The diversity of the ‘flock’ is something we measure. We stand for an environment where everyone can be their best selves and know they belong. Diversity and equity and inclusion allow us to dream boldly,” says Gregory. He explains that the company has three “domains of action and accountability”—inclusive business practices, **cont.**

Featured participants

Abbott
www.abbott.com**AbbVie**
www.abbvie.com**Alnylam Pharmaceuticals**
www.alnylam.com**Biocon Limited**
www.biocon.com**bluebird bio**
www.bluebirdbio.com**Incyte**
www.incyte.com**Novartis**
www.novartis.com**Novo Nordisk**
www.novonordisk.com**Pfizer**
www.pfizer.com**Regeneron**
www.regeneron.com**Syngenta**
www.syngenta.com**Takeda Pharmaceutical Company Limited**
www.takeda.com**Vertex Pharmaceuticals**
www.vrtx.com

inclusive workforce culture, and workforce diversity—and adds, “Our core values are connected to our five nonnegotiables: We challenge our colleagues to be authentic, courageous, humble, caring, and transparent.” The DEI initiatives include top-down investments in hiring, retention, and development as well as employee resource groups, such as those that support individuals who are LGBTQ, Black, Latinx, Asian, women, veterans, disabled, and parents and caregivers. The company has a DEI officer on its senior management team and has developed specific programs to support underrepresented employees and foster an inclusive culture.

Regeneron is also willing to scrutinize its own diversity practices. “We believe that diversity of employees is as important as diversity of ideas. We are proud of a diverse workforce in terms of immigrants. But we look to do more. We can do better,” says Murphy. One example of the company’s dedication to this philosophy is its increase in recruitment activities at Historically Black Colleges and Universities. “There is more talent. We want to be truly reflective of the U.S. and the world. Biotech is not as diverse as it could be.”

Although the MeToo movement has ignited support for gender diversity among several of the top employers, Biocon stands out because Mazumdar-Shaw has always aimed to promote gender parity. Of the company’s 12,000 employees, half are in R&D, and over a third of its scientific employees are female. “As a woman scientist, I have been driven to make sure the company is supportive of women scientists,” says Mazumdar-Shaw, who serves as executive chairman. “I wanted this to be a company where women scientists feel comfortable and excited to go to work.” As the pandemic anchored her employees to their residences, she launched a listening tour to dialogue with them and find out how they were managing. She noticed that working from home could be leveraged—for the benefit of the staff. “Working from home is a boon to our women scientists,” she says. “To offer better work-life balance, we will give them the opportunity to work from home 2 days a week when things normalize post-COVID.”

Building sustainable organizations, one human at a time

Top employers continue to look for ways to express their commitment to their communities. Of key importance are environmental and sustainability concerns. At Syngenta, for example, climate change is not just a consideration—it is the impetus for innovation to support farmers.



DEMOGRAPHICS

GENDER:

49% Male, 45% Female, 6% No response

EXPERIENCE:

72% have 10 or more years work experience

HIGHEST DEGREE EARNED:

31% Doctorate, 32% Master’s, 31% Bachelor’s, 6% Other

COMPANY TYPE:

More than 9 out of 10 work in private industry: 25% Biotech, 42% Biopharma, 28% Pharma

NATURE OF WORK:

12% Basic Research, 16% Applied Research, 25% Development, 8% Production, 11% QA/QC or Regulatory Affairs, 9% Executive, 19% Other

GEOGRAPHY:

67% North America, 19% Europe, 10% Asia/Pacific Rim, 4% Rest of world

“We consider what we do as contributing to the world’s food security and helping agriculture to protect the planet,” says Gusui Wu, head of global seeds research. “The world faces significant environmental and nutritional challenges, which are magnified in the developing world and recently underscored by COVID.” Detailed, rigorous decisions about how to assist food growers in an environmentally friendly avenue is fortified through data science that pervades every touchpoint of the organization. But data doesn’t run the show, humans do. “From multiple surveys we have done over the years, there is a genuine feeling from our employees that we are proud of what we do because we are contributing to the global food supply. By 2050, there will be a global population of 10 billion, so the world needs to significantly increase its food supply,” he adds. “Employees know that our continuous work is needed by our customers. Farmers depend on the products and technologies we develop. Our employees are a big part of the ag economy.”

Indeed, the future of pharma and biotech industries, and for that matter, science, is decidedly human. And our top employers celebrate this.

“Employees drive our business,” says Murphy, noting that his company has been on the Top Employers list at either No.1 or No. 2 over the last ten years. “It’s really important to us. Organizations don’t make drugs, people do.”

To read this article online and to view top employer profiles, visit ScienceCareers.org/TopEmployers.

Alaina G. Levine is a science writer, science careers consultant, professional speaker, and author of Networking for Nerds (Wiley, 2015).



Zhengzhou University is one of the key universities of the national “211 Project”, and of “World 1st University Project”, as well as of universities co-sponsored by the Henan Provincial Government and the Ministry of Education. It was born in 10th, July 2000 when the former ZZU, Zhengzhou University of Technology (ZUT) and Henan Medical University (HMU) merged into one, and now developed into a comprehensive university with 12 major disciplines, namely liberal arts, science, engineering and medicine, etc. It has over 50,000 full-time undergraduates, more than 23,000 full-time postgraduates and nearly 2,500 international students. Meanwhile, there are 30 first-level disciplines authorized to confer doctorate degrees, 3 professional doctorate programs, 28 postdoctoral research stations; 3 “first-class” disciplines and 6 national key (being cultivated) disciplines. ZZU is home to 12 national scientific and research platforms including state key laboratories, engineering research centers, engineering laboratories, etc. Among a total of more than 5800 faculty members, there are 15 academicians of Chinese Academy of Science and Chinese Academy of Engineering, 2 members of Chinese Academy of Social Sciences and 4 overseas academicians as well as 62 elected into national talent

programs (projects). They have formed a strong team with academicians and academic masters as the head, distinguished young scholars and Yangtze River scholars as academic leaders, and outstanding PhDs as core talents.

Rooted in the broad and profound culture of the Central Plains of China and nurtured by blended diversified cultures and disciplines, ZZU people have developed the virtues of tolerance, lenience and can-do spirit and pursued ZZU’s mission and spirit of seeking truth and being responsible, thus forming typical school ethos-- Perseverance, Trustworthiness, Benevolence, Tolerance, Modesty, Prudence, and Critical Thinking and Diligence. Standing at this new starting line, ZZU decided where to go and laid out a three-step strategy to make it to the list of world-class comprehensive & research-oriented universities by the middle of this century. To achieve high-quality development, the university has made efforts to strengthen itself through talents. With the idea of “top-notch talents breakthrough, team integration, excellent scholars cultivation and structural optimization”, ZZU implemented Distinguished Professor Program to attract more academic masters, and Excellent Scholar Cultivation Program to create an ideal environment for the growth of talents.

ZZU, your stage to make a difference!

■ Recruiting Position

Distinguished Professor, Top talents

A competitive remuneration package and scientific research services referring to the standards of world-class universities will be provided for the right candidates.

■ How to apply

Please send full curriculum vitae (please note the position and discipline you are applying for) to

rcb@zzu.edu.cn.



中北大学
NORTH UNIVERSITY OF CHINA

Welcome to Join Us

Brief Introduction

Sharing the only garden-like campus in Taiyuan, a beautiful historic city of over 2,500 years old in Shanxi Province, North University of China (NUC) is a multi-disciplinary teaching and research university including engineering, science, liberal arts, economics, management, law, art and education. NUC consists of 1 national key discipline, 15 provincial key advantage disciplines and 6 centers for post-doctoral studies, and is authorized by the Academic Degrees Committee of the State Council to offer 24 PhD programs in 7 first-level disciplines, 64 Master programs in 23 first-level disciplines and 85 Bachelor programs.

Up to now, NUC has achieved tremendous accomplishments in a great number of research projects, some of which have won high-level awards, including 5 Second Prizes of National Award for Science & Technology and more than 100 Provincial Awards for Science & Technology.

Recruitment

1. Postdoctoral scholar/fellow holding formal research positions in overseas universities, or Graduate student with PhD of overseas universities.
2. Distinguished talents with different titles of National-Level and Provincial-Level Scholars, Known Associated Professor of universities home and abroad and Doctors with remarkable achievements.

Payment and Welfare

Generous Settling-down Allowance, Research Start-Up Funds, On-duty Allowance and ample Housing will be offered.

Contact

Tel: 86+03513924993
Fax: 86+03513924993
Email: rsglk@nuc.edu.cn
Website: www.nuc.edu.cn

Who's the top employer for 2020?

Science Careers' annual survey reveals the top companies in biotech & pharma voted on by *Science* readers.

Read the article and employer profiles at sciencecareers.org/topemployers



Science 2020 TOP EMPLOYER



The Zoological Research Museum Alexander Koenig seeks a

Head for the Centre of Molecular Biodiversity Research as a joint appointment of a chair in Molecular Biodiversity Research (W3) of the Museum and the University of Bonn

The Faculty of Mathematics and Natural Sciences of the Rheinische Friedrich-Wilhelms-Universität Bonn, Department of Biology, and the Zoological Research Museum Alexander Koenig (ZFMK) - Leibniz Institute of Animal Biodiversity are looking for a head of the Centre for Molecular Biodiversity Research (successor to Prof. Bernhard Misof) at the Zoological Research Museum Alexander Koenig (Bonn), jointly appointed as Chair of Molecular Biodiversity Research (W3) by the University of Bonn and ZFMK.

The ZFMK is a foundation under public law and member of the Leibniz Association. As an institution, it contributes to the research and conservation of global biodiversity, conducts taxonomic and molecular biodiversity research, documents and analyzes evolutionary and ecological biodiversity change and engages in public communication on biodiversity change and its potential causes. The ZFMK is an internationally renowned research institute for terrestrial and freshwater zoology and houses the Centre for Molecular Biodiversity Research, which is unique in Germany.

The ZFMK is looking for a strategically thinking and integrative scientist with international standing as head of the center who will further develop molecular biodiversity research as a bridge between basic research and applied biodiversity research at an internationally excellent level.

Applicants should have leadership experience and must meet the requirements for a professorship (habilitation or habilitation-equivalent performance).

Scientific excellence and expertise should be evident, for example, in the fields of:

- phylogenomics and its algorithms,
- comparative/evolutionary genomics and its method development,
- bioinformatics,
- species-based molecular or computational biodiversity research.

We expect

- applicants to demonstrate successful international networking,
- substantial acquisition of third-party funding,
- and the development of a future strategy of the center including the molecular lab and HPC infrastructure.

It is desirable that applicants are familiar with collection-based science.

Seven working groups with different expertise are cooperating in the Centre for Molecular Biodiversity Research. These range from bioinformatics, method development in phylogenomics and comparative genomics, comparative genomics of vertebrates and insects, evolutionary genomics, to molecular laboratory method development. In addition, the center operates the central molecular laboratory and the HPC cluster of the ZFMK. The center is well connected and enjoys an excellent international reputation (<https://www.zfmk.de/en/zmb>). The new head of the center will have the opportunity to carry out research in an outstanding research environment, and cooperate with a young, dynamic and motivated team.

The University of Bonn, as one of the few Excellence Universities in Germany, offers an outstanding intellectual research environment. The Department of Biology at the University of Bonn and especially the Department of Zoology cooperate intensively with the ZFMK. The common research focus lies on the interaction between environment, phenotype and genome as a basis for an understanding of evolutionary as well as ecological biodiversity change. The W3 Professorship Molecular Biodiversity Research will be integrated in the Institute of Evolutionary Biology and Ecology; the professorship will be filled for an unlimited period of time; the appointment will be based on the leave of absence model ("Jülich model"). The reduced teaching commitment of 2 SWS should be provided in the international Master's programme OEP Biology. The membership status of a professor can be granted. Further information can be found at www.biologie.uni-bonn.de and www.oep-bio.uni-bonn.de. The working place will be the ZFMK.

The requirements for employment are based on the standards for filling scientific leadership positions in the Leibniz Association and on § 36 of the Higher Education Act of North Rhine-Westphalia.

The Leibniz Association and the University of Bonn are committed to diversity and equal opportunities. The University of Bonn is certified as a family-friendly university and has a dual career service. The ZFMK is also certified as a family-friendly institute and cooperates with the Dual Career Network of the Science Campus Bonn. The aim is to increase the proportion of women in areas in which women are underrepresented and to particularly promote their careers. Therefore, women with relevant qualifications are strongly encouraged to apply. Applications are treated in accordance with the State Equal Opportunities Act.

The application of suitable candidates with proven severe disabilities and persons of equal status is particularly welcome.

Applications in English with the usual documents (CV, list of publications, teaching experience, details of third-party funding, copies of university certificates and diplomas) and a concept for the development of the Centre for Molecular Biodiversity Research and for the cooperation with the university are requested in electronic form by **6th of December 2020** to the Chair of the Division of Biology, Endenicher Allee 11 - 13, 53115 Bonn, Germany (email: biologie@uni-bonn.de).

The call for proposals is available on the university website:

<https://www.uni-bonn.de/die-universitaet/stellenangebote/professuren>



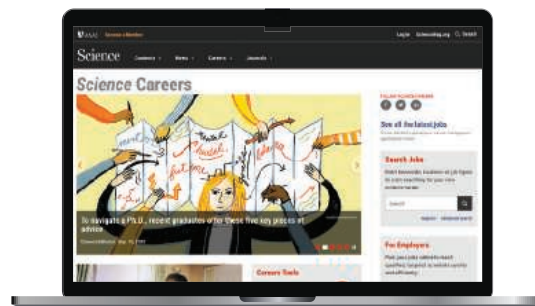
Science Careers helps you advance your career. Learn how !

- Register for a free online account on **ScienceCareers.org**.
- Search hundreds of job postings and find your perfect job.
- Sign up to receive e-mail alerts about job postings that match your criteria.
- Upload your resume into our database and connect with employers.
- Watch one of our many webinars on different career topics such as job searching, networking, and more.
- Download our career booklets, including Career Basics, Careers Beyond the Bench, and Developing Your Skills.
- Complete an interactive, personalized career plan at "my IDP."
- Visit our Employer Profiles to learn more about prospective employers.
- Read relevant career advice articles from our library of thousands.

Visit **ScienceCareers.org**
today — all resources are free

ScienceCareers

FROM THE JOURNAL SCIENCE 



SCIENCECAREERS.ORG



ICGEB International Centre for Genetic Engineering and Biotechnology | Developing Knowledge

Applications are invited for positions of **Group Leaders/Scientists** at International Centre for Genetic Engineering and Biotechnology (ICGEB), New Delhi Component.

ICGEB is an international intergovernmental organization with components at Trieste (Italy), New Delhi (India) and Cape Town (South Africa). ICGEB is a Centre of Excellence for research and training in biotechnology addressing issues of the developing world and works under United Nations Common System. ICGEB, New Delhi has about two dozen independent groups working in three broad domains; Molecular Medicine, Plant Biology and Integrative Biology (**Further information: <http://www.icgeb.org>**).

Incumbents can be of any nationality and must have an excellent post Ph.D. record of research/ publications in any area of modern biology and biotechnology. While highly competent candidates under any of the research mandates of ICGEB, New Delhi will be considered, expertise in areas of plant biology with direct applications in agriculture and virology concerning new and emerging diseases are currently preferred. Ability and commitment to build a dynamic program of international stature is essential. Incumbent must have significant accomplishments commensurate with their career stage and should have proven ability to compete for extramural national/international funding and head a successful research laboratory. Evidence of independent research achievements, ability to guide Ph.D. students and teach Ph.D. level courses as well as post-Ph.D. level training programmes will be an asset. Applicants should submit their curriculum vitae, a statement of research interests and names and addresses of three referees to **The Director, ICGEB, Aruna Asaf Ali Marg, New Delhi- 110 067, INDIA**. Additionally, applicants should submit their full CV by completing the ICGEB's Personal History Form available online at: <https://www.icgeb.org/about-us/work-with-us/>, and address it to email id: icgeb.director@gmail.com.

Science Careers

FROM THE JOURNAL SCIENCE

Confused about
your next career move?



**Download Free Career
Advice Booklets!**

ScienceCareers.org/booklets



**CASE WESTERN RESERVE
UNIVERSITY** EST. 1826

TENURE TRACK/TENURED FACULTY POSITION IN GASTROINTESTINAL CANCER RESEARCH

Case Western Reserve University
School of Medicine

Case Comprehensive Cancer Center

The Gastrointestinal (GI) Cancer Genetics Program at The Case Comprehensive Cancer Center (<https://case.edu/cancer/>), an NCI-designated Comprehensive Cancer Center at CWRU, with affiliates University Hospitals Cleveland Medical Center and Cleveland Clinic, invites applications for tenure track/tenured faculty positions at the level of Assistant, Associate, or Full Professor. The program has a nationally recognized GI cancer translational research faculty that is currently supported by multiple team awards that include: GI SPORE (NIH Specialized Program in Research Excellence in GI Cancers, BETRNet (NIH Barrett's Esophagus Translational Research Network), EDNR (NIH Early Detection Research Network), and SU2C (Stand Up 2 Cancer) colorectal cancer dream team award. Our areas of strength include colorectal and esophageal cancers. Applicants are being sought that have basic/translational GI cancer research expertise in one of the following areas: molecular pathology, cancer genetics, bioinformatics, cancer immunology, cancer metabolism or signaling. Qualified individuals should have a doctoral degree (PhD or MD) with established expertise in the specific cancer to be studied. Applicants for Assistant Professor must have a track record of outstanding cancer research and publications and potential for extramural funding. Candidates applying for consideration at a senior rank must possess national (Associate Professor) or international (Professor) reputations in collaborative cancer research, a distinguished record of publication and funding, a commitment to mentoring, teaching, and leadership, and they must fulfill other qualifications necessary for a tenured appointment at CWRU.

Please send curriculum vitae and a cover letter outlining your research interests electronically to: Zhenghe John Wang, PhD, Co-leader, GI Cancer Genetics Program, Case Comprehensive Cancer Center, c/o cancersearch@case.edu. After initial review, you will be asked for a list of three or more references. Please include "GI Cancer Faculty Search" in the subject line.

In employment, as in education, Case Western Reserve University is committed to Equal Opportunity and Diversity. Women, veterans, members of underrepresented minority groups, and individuals with disabilities are encouraged to apply.

Case Western Reserve University provides reasonable accommodations to applicants with disabilities. Applicants requiring a reasonable accommodation for any part of the application and hiring process should contact the Office of Inclusion, Diversity and Equal Opportunity at 216-368-8877 to request a reasonable accommodation. Determinations as to granting reasonable accommodations for any applicant will be made on a case-by-case basis.

FIND YOUR HAPPIER PLACE.



Find your next job at [ScienceCareers.org](https://www.sciencecareers.org)

There's scientific proof that when you're happy with what you do, you're better at what you do. Access career opportunities, see who's hiring and take advantage of our proprietary career-search tools. Get tailored job alerts, post your resume and manage your applications all in one place: [sciencecareers.org](https://www.sciencecareers.org)

ScienceCareers

FROM THE JOURNAL SCIENCE  AAAS

myIDP:

A career plan customized
for you, by you.



For your career in science, there's only one Science

Features in myIDP include:

- Exercises to help you examine your skills, interests, and values.
- A list of 20 scientific career paths with a prediction of which ones best fit your skills and interests.
- A tool for setting strategic goals for the coming year, with optional reminders to keep you on track.
- Articles and resources to guide you through the process.
- Options to save materials online and print them for further review and discussion.
- Ability to select which portion of your IDP you wish to share with advisors, mentors, or others.
- A certificate of completion for users that finish myIDP.



Visit the website and start
planning today!
myIDP.sciencecareers.org



In partnership with:



University of California
San Francisco



BURROUGHS
WELLCOME
FUND



FASEB
Federation of American Societies
for Experimental Biology



OPEN RANK FACULTY POSITION IN LIFE SCIENCES

Building on a successful multiyear initiative to hire, support, and retain leading research scientists with a strong commitment to promoting the success of underrepresented students, the Division of Life Sciences in the UCLA College of Letters and Sciences announces an open rank **faculty position** with a tenure-track, Academic Senate appointment in one of the departments of Life Sciences (www.lifesciences.ucla.edu). The departments of Ecology and Evolutionary Biology (EEB); Integrative Biology and Physiology (IBP); Microbiology, Immunology, and Molecular Genetics (MIMG); Molecular, Cell, and Developmental Biology (MCDB); Institute for Society and Genetics (ISG), and Psychology are all partnering in this search, as are many associated research institutes. Candidates must have a PhD in a field relevant to one of the above departments and a history of mentoring acumen in addition to an outstanding record of scholarly publications and success obtaining funding. The successful candidate will be expected to initiate or continue an active independent research program and continue participating in mentoring activities with particular focus on students and trainees from underrepresented groups. Service and teaching expectations will not exceed those of any other tenure track faculty position. Faculty appointment at the level of Assistant, Associate, or Full Professor will be commensurate with experience and academic standing. We highly encourage applications from women and individuals who are from groups underrepresented in the sciences. UCLA offers competitive salaries, research set-up funds, and recruitment allowances as well as a thriving community of scholars committed to mutually supporting one another.

Application packages should be submitted online through <https://recruit.apo.ucla.edu/JPF05899> and include the following documents: **1)** a curriculum vita; **2)** a research statement; **3)** a statement of contributions to equity, diversity, and inclusion with particular attention to formal and informal mentoring activities and detailed plans for continuing such activities in the future; **4)** a statement of teaching interests; and **5)** a cover letter that includes the names of three referees who can be contacted for letters. Each of the five items should be submitted as a standalone document. Review of applications will begin on **November 13, 2020** and continue until the position is filled. Both inquiries about the position and nominations of potential candidates should be sent to search committee chair: Professor Gina Poe (ginapoe@ucla.edu).

The University of California is an Equal Opportunity/Affirmative Action Employer. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, sexual orientation, gender identity, national origin, disability, age, or protected veteran status. For the complete University of California nondiscrimination and affirmative action policy see: UC Nondiscrimination & Affirmative Action Policy. (<http://policy.ucop.edu/doc/4000376/NondiscrimAffirmAct>).

By Anonymous

The things I wish I could say

Dear Ph.D. adviser: I just finished editing my paper based on the comments from our two colleagues. Reading what they wrote felt like walking through hell (though I don't believe that hell or heaven exists). Yes, it felt that bad. It took me 2 days to get over that horrible feeling before I could focus on editing the paper. Then I asked myself why their comments made me feel so bad. Does my writing suck that much? No. I know I have room to improve, but I have received positive responses from other colleagues about my manuscripts in the past. Am I incapable of taking feedback? No. Comments from another professor, two postdocs, and you, dear adviser, do not cause that sharp pain. Were the comments profane or abusive? No. But in my head, I could hear our colleagues making their points in the sarcastic, disrespectful tone they use when they speak to me. The comments were a trigger, waking other painful memories I have tried to bury.

Over the years I have worked with these colleagues, they have made me feel less intelligent and less human than they are. One told me I needed to read more. The other told me I was stupid on many occasions, saying "I don't believe you" until I showed him papers that stated almost exactly what I had said. Then there was the time when he disparaged one of my ideas so loudly that another colleague put on headphones so she didn't have to hear it.

I can't help thinking they treat me this way because of my gender and race. I'm a Thai woman who speaks English with a strong accent. They seem to think of me as an exotic animal who happens to be able to speak English.

Dear adviser, you and I don't always agree, but you never treat me that way. When I talk with you, I feel well-informed. You tell me I'm intelligent. You say you believe in me.

Yet I'll never be brave enough to tell you how these colleagues make me feel. You're so close to them, and I'm afraid you would try to protect them and justify their behavior. You would probably tell me they are kind and empathetic in their own ways. I'm afraid that you would tell me to tolerate the sexism, racism, and microaggressions. I'm afraid that, if I speak up, you might not want me in your lab anymore. And I'm afraid that, if you respond that way, I won't want you to be my adviser anymore.

I'll never be brave enough to say it to them, either. They would reject my reality and say I am too sensitive, as one of them has already told me. But even though these



**"I need more than a thicker skin.
I need a spiky one, like
a cactus, to reclaim my dignity."**

colleagues dismiss my experience, my feelings are valid.

Dear adviser, you once told me I need to grow a thicker skin. But why is it my responsibility to change my skin? Why should I have to invest my time, money, and energy to heal myself from their behaviors? I already went to a therapist and found a good self-help book to deal with inescapable toxic co-workers. But why isn't it their responsibility to learn that their behavior is unacceptable?

I love my work and I want to finish my Ph.D. with grace. I just wish I didn't have to sacrifice my energy and mental health to deal with this discrimination. I just wish I didn't feel like I was in a meanness competition, where the meanest person gets the trophy. I just wish my colleagues treated me as an equal. Dear adviser, I wish all of this not

only for me, but also for other nonwhite young women who want to pursue their scientific dreams.

Lastly, I will take your advice about changing my skin. But I think, to survive all of this, I need more than a thicker skin. I need a spiky one, like a cactus, to reclaim my dignity. I will no longer let anyone step on me. Although I am not ready to share my feelings with you or our colleagues directly, I plan to change my approach going forward. I will empower myself to say no to unacceptable behavior—without becoming another contestant in the meanness competition. It's hard, and it will take practice. But for my mental health, it's worth it. ■

The author is a Ph.D. student in the southwestern United States.
Send your career story to SciCareerEditor@aaas.org.