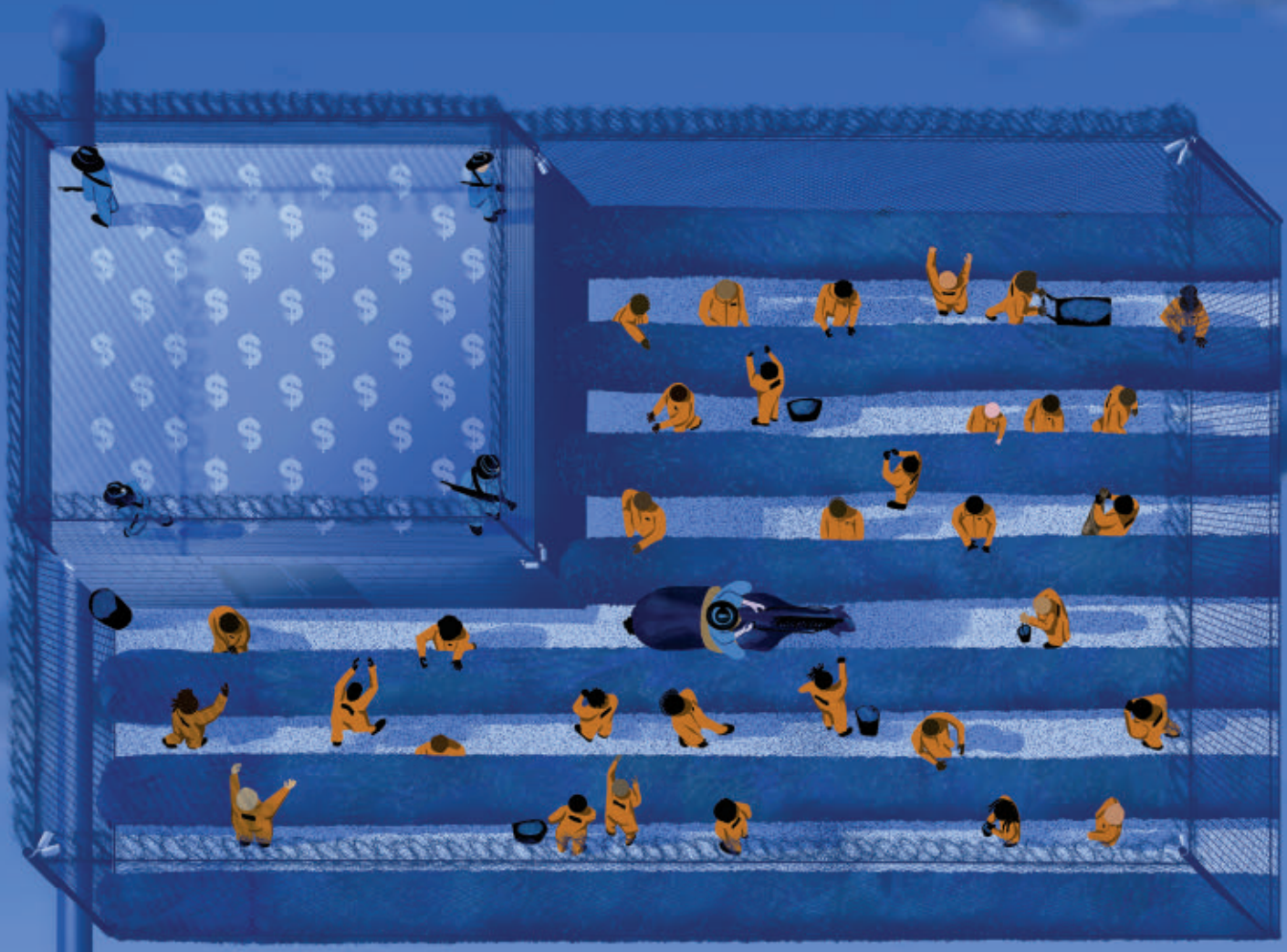
**Should a dog's sniff send a suspect to prison?** p. 251

**Controlling phase selectivity in zeolite synthesis** pp. 257 & 308

**Doctoring delivery decisions in the clinic** pp. 260 & 324

# Science

$15
15 OCTOBER 2021
SPECIAL ISSUE
science.org
AAAS



# CRIMINAL INJUSTICE

Mass incarceration in the United States p. 270

BII | Prize for Science | Innovation

# Apply for the new BII & Science Prize for Innovation today

Behind every life-changing solution is an entrepreneurial scientist–a creative mind who proved an idea in the lab and dared to carry it out in the world.

To encourage more scientists to translate their research, BioInnovation Institute (BII) & Science present a new annual award.

Our three winners will have their essays published in Science magazine and will be invited into BII's entrepreneurial ecosystem. In addition, the Grand Prize winner will receive a prize of USD 25,000 and each runner-up will receive USD 10,000 at a grand award show celebration in Copenhagen, Denmark.

The call for applications has just opened.
Apply before November 1, 2021.
**www.bii.dk/scienceprize**

Presented by BII & Science

BII BioInnovation Institute Science

Apply before November 1, 2021
www.bii.dk/scienceprize

# CONTENTS

15 OCTOBER 2021 • VOLUME 374 • ISSUE 6565

SPECIAL SECTION

## Mass incarceration


**INTRODUCTION**

**270** An outlier of injustice

**PERSPECTIVE**

**272** Policing social unrest and collective violence *E. Hinton*

**POLICY FORUM**

**274** The corporate shadow in democratic policing *E. E. Joh*

**REVIEWS**

**277** Assessing mass incarceration's effects on families *H. Lee and C. Wildeman*

PODCAST

**282** Exclusion and exploitation: The incarceration of Black Americans from slavery to the present *C. Muller*

**286** Toward an understanding of structural racism: Implications for criminal justice *J. M. Rucker and J. A. Richeson*

**291** The predatory dimensions of criminal justice *J. Page and J. Soss*


**ON THE COVER**

Mass incarceration in the United States has deep structural roots, with ties to race, slavery, labor, and money. In the words of the illustrator, "This illustration shines a light on the structural role of the prison system and how deeply embedded it is in the fabric of this country." For more on this topic, see page 270. *Illustration: Adrian Brandon*

SEE ALSO EDITORIAL p. 237; NEWS STORY p. 251

ILLUSTRATION: ADRIAN BRANDON

# NEWS


**IN BRIEF**

**242** News at a glance

**IN DEPTH**

**245 WHO gives first malaria vaccine the green light**
Widespread distribution in Africa set to begin despite lingering questions about efficacy *By G. Vogel*

**246 Massive road project threatens New Guinea's biodiversity**
The 4000-kilometer Trans-Papua Highway could turn the island's western half into an "El Dorado for investors" *By D. Rochmyaningsih*

**248 USDA revokes license of leading chinchilla supplier**
Hearing researchers may be left scrambling for animals *By M. Wadman*

**248 Biden restores key monuments**
Move reinstates protections for fossil sites, marine zone *By A. Reese*

**249 China embraces global biodiversity push—cautiously**
Xi's silence on 30% set-aside target disappoints advocates *By D. Normile*

**250 New observatory to bring soil up from the deep**
Ecotron will move 3-meter cores into the laboratory for global change experiments *By E. Stokstad*

**FEATURES**

**251 The sniff test**
Courtroom testimony about dogs detecting dead bodies keeps sending people to prison—even without physical evidence. Critics say the science is lacking *By P. A. Smith*

MASS INCARCERATION SECTION p. 270; PODCAST


# INSIGHTS


**PERSPECTIVES**

**256 The benefits of disturbance**
Physical disruption of an ecosystem by sea otters affects the genetic diversity of their habitat *By J. Roman*

REPORT p. 333

**257 No more trial and error for zeolites**
A data-intensive approach scours known organics for synthesizing targeted zeolites
*By W. Chaikittisilp and T. Okubo*
RESEARCH ARTICLE p. 308

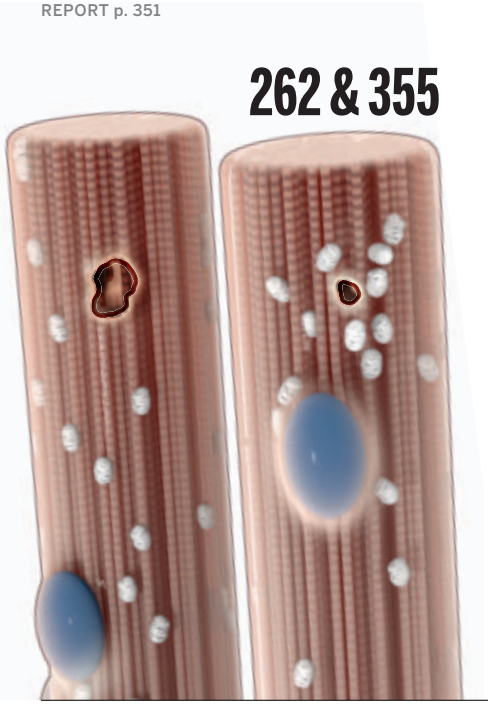
**259 Gulf Stream and Kuroshio synchronization**
The mid-latitude jet stream ties together the two major western boundary currents
*By P. Cessi*
REPORT p. 341

**260 Physicians' flawed heuristics in the delivery room**
Delivery complications can influence later decision-making *By M. Li and H. Colby*
RESEARCH ARTICLE p. 324

**262 Resealing and rebuilding injured muscle**
Reconstructing an injured muscle involves nuclear and mitochondrial contributions
*By E. M. McNally and A. R. Demonbreun*
REPORT p. 355

**263 A cell cycle regulator branches out**
A key $G_1$ kinase targets transcription to drive cell cycle commitment
*By R. P. Fisher*
REPORT p. 347

**264 Not all stress is bad for your heart**
Shear stress forces instruct valve formation during cardiac development
*By R. Jain and J. A. Epstein*
REPORT p. 351




## BOOKS *ET AL.*

**266 The history and future of fire**
More prescribed burns and less fossil fuel combustion could help rebalance our relationship with fire
*By M. E. Hannibal*

**267 Toward a holistic model of Alzheimer's**
An overemphasis on amyloid has stymied progress on the disease, alleges a neuroscientist *By T. Spires-Jones*

## LETTERS

**268 Bird-friendly buildings for China's cities**
*By H. Yang* et al.

**268 Community stewardship of China's national parks**
*By Y. Feng* et al.

**269 Fences undermine biodiversity targets**
*By J. Sun* et al.

**269 Technical Comment abstracts**

# RESEARCH

## IN BRIEF

**296** From *Science* and other journals

## RESEARCH ARTICLES

**299 Immunology**
Mitochondrial translation is required for sustained killing by cytotoxic T cells
*M. Lisci* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABE9977

**300 Neuroscience**
Astrocyte $Ca^{2+}$-evoked ATP release regulates myelinated axon excitability and conduction speed *J. Lezmy* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABH2858

**301 Catalysis**
Univariate classification of phosphine ligation state and reactivity in cross-coupling catalysis *S. H. Newman-Stonebraker* et al.

**308 Zeolite synthesis**
A priori control of zeolite phase competition and intergrowth with high-throughput simulations *D. Schwalbe-Koda* et al.
PERSPECTIVE p. 257

**316 Neuroscience**
Modulation of tactile feedback for the execution of dexterous movement *J. M. Conner* et al.

**324 Health economics**
Heuristics in the delivery room *M. Singh*
PERSPECTIVE p. 260

## REPORTS

**330 Exoplanets**
A compositional link between rocky exoplanets and their host stars
*V. Adibekyan* et al.

**333 Trophic cascades**
Physical disturbance by recovering sea otter populations increases eelgrass genetic diversity *E. Foster* et al.
PERSPECTIVE p. 256

**336 Physics**
Electrically tunable Feshbach resonances in twisted bilayer semiconductors
*I. Schwartz* et al.

**341 Ocean circulation**
The Gulf Stream and Kuroshio Current are synchronized *T. Kohyama* et al.
PERSPECTIVE p. 259

**347 Cell cycle**
$G_1$ cyclin–Cdk promotes cell cycle entry through localized phosphorylation of RNA polymerase II *M. Kõivomägi* et al.
PERSPECTIVE p. 263

**351 Developmental biology**
Bioelectric signaling and the control of cardiac cell identity in response to mechanical forces *H. Fukui* et al.
PERSPECTIVE p. 264

**355 Muscle repair**
Muscle repair after physiological damage relies on nuclear migration for cellular reconstruction *W. Roman* et al.
PERSPECTIVE p. 262

## DEPARTMENTS

**237 Editorial**
Analyzing mass incarceration *By Sean Joe*
MASS INCARCERATION SECTION p. 270

**366 Working Life**
Getting personal *By Sakshi Ghai*

New Products ..... 360
*Science* Careers ..... 361

ILLUSTRATION: V. ALTOUNIAN/*SCIENCE*

**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 © 2021 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 (51 issues): $2148; Foreign postage extra: Air assist delivery: $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.science.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.

EDITORIAL

# Analyzing mass incarceration

With almost 2 million people in confinement, the United States locks up more people per capita than any other nation. Understanding the reasons and then forging a path to reduce mass incarceration in America will require better research and analyses of the government policies and spending that sustain the US carceral system.

Can the largest prison population in the world be attributed to greater criminal behavior by Americans? Not likely—the US crime rate is comparable to those in many Western industrial nations. Although the national violent crime rate increased during the crack-cocaine epidemic, peaking in 1993, the rate has steadily declined to a low not seen since the 1970s. Yet the incarcerated population has increased by an estimated 700%.

In 2014, the National Research Council reported that this phenomenon is fueled by an increase in the likelihood of imprisonment and of longer prison sentences. This was stoked by the "get tough" policies enacted since the 1980s, such as mandatory minimum sentences, a reduction in early release from prison mandatory sentences for drug-related crimes, life sentences without the possibility of parole, and the "three-strikes" law that increases the prison sentences of persons convicted of a felony who have been previously convicted of two or more serious crimes. Since 1984, the number of people serving life sentences has nearly quintupled, even though serious violent crime has been declining for the past 20 years.

> "Yet the incarcerated population has increased by an estimated 700%."

These government policies have led to only minor gains in crime reduction. Notably, the majority of incarcerated people in the United States are not charged with violent crimes. Every year, an estimated 13 million misdemeanor charges sweep masses of Americans into the criminal justice system for behaviors as benign as jaywalking, sitting on a sidewalk, or technical parole violations. Such low-level offenses account for over a quarter of the daily jail population nationally and even more in some US states and counties.

Misdemeanor arrrests may not sound like a big deal, but they carry serious personal, social, and financial costs, especially for defendants. The earnings lost each year by the roughly 14% of the US population with a misdemeanor criminal record is about $238 billion. There are also monetary costs to society for a "mass misdemeanor" system that pays for processing these court cases and the resulting incarceration. The US carceral system operates at the scale of billions of dollars and is financed mainly by the government and banks. To better understand this sustained investment, it is necessary to assess the economy that the carceral system upholds. The direct economic beneficiaries are largely the rural towns where prisons are built, those employed by these prisons, and the state-owned and private correctional industries that manage prisons and prisoner services. So stakeholders have a vested economic interest in building more jails and expanding the market for prisoner services.

The policies that have increased the "mass misdemeanor" system have also resulted in the "mass criminalization" of mostly Blacks, Latinos, and the poor for nonviolent misdemeanor violations and drug possession. Blacks make up 13% of US residents yet are 40% of the incarcerated population. Generally, criminal laws themselves in America are written to be color-blind and class-blind. Nevertheless, the way that policing is conducted and judicial decisions are made can result in the laws being unequally applied. For example, according to an American Civil Liberties Union report in 2013, even though Blacks and whites have comparable rates of marijuana use, Blacks were almost four times more likely to be arrested for marijuana possession. Compounding the imbalance, minority communities are more routinely surveilled and minority individuals are more often unjustly stopped, arrested, and abused by police for what appear to be misdemeanor crimes.

There is no simple solution to these problems. A giant step forward would be to better understand how policy and spending sustain the US carceral economy. One would think that insights could be quickly gained given advances in digital technology and the wealth of critical data from many government agencies involved in the justice system. Unfortunately, the data are fragmented across these agencies, ripe with bias, and not designed to help policy-makers or the public understand what is going on. Future discussions on the criminal justice system must address these data problems so that evidence-based reforms can truly end mass incarceration.

**–Sean Joe**

**Sean Joe**
is the Benjamin E. Youngdahl Professor of Social Development at the George Warren Brown School of Social Work, Washington University in St. Louis, MO, USA.
sjoe@wustl.edu

PHOTO: JOE ANGELES/WASHINGTON UNIVERSITY

10.1126/science.abm7812

# Sparking a new vision for collaborative research

For NOMIS Fellow Jan Pencik, a biochemist from the Czech Republic, the decision to pursue his postdoc at the Salk Institute for Biological Studies–the world-renowned research enterprise perched on a bluff above the Pacific Ocean in La Jolla, California–was an easy one. "It wasn't just the view," he jokes. "It has a certain scientific vibe." That vibe is one that has allowed him to flourish there in only 2½ short years. As a research fellow in cancer immunotherapy at the Institute's NOMIS Center for Immunobiology and Microbial Pathogenesis, Pencik is interested in tracking signaling molecules involved in resistance to immunotherapy in lung cancer. This requires close collaboration with immunotherapy experts. The NOMIS Center Postdoctoral Fellowship Program provides him with the unique opportunity to do just that. "This means a lot because I can interact with NOMIS Center leaders who are immunologists," says Pencik. "The interaction and the support of the Foundation is the key. It is helping me to drive this research."

Lidia Jiménez, a pharmaceutical scientist and molecular biologist from Spain, is another NOMIS Fellow who has greatly benefited from the interdisciplinary approach that is core to the fellowship, the NOMIS Center, and the Salk Institute. After receiving her Ph.D. focusing on macrophages, the sentinel cells that engulf and digest pathogens and dead cells from our bodies, she wanted to study a family of receptors called Tyro3, Axl and Mertk (TAM) to better understand their underlying mechanisms for clearing away dead cells, the accumulation of which has long been associated with autoimmune diseases such as lupus, she says. At the NOMIS Center, she is able to hear from experts and share her passion with them. "If I need advice about certain topics, I can ask a doctor here who can help me better understand not only macrophages, but also new techniques and other cell types that I am less familiar with," she says. When she needs new experimental protocols, tools, or reagents, she has easy access to team members who can assist her.

## Mentoring the next generation of scientists

Founded in 1960 by Jonas Salk, the Salk Institute is an independent, nonprofit organization dedicated to basic research in the biological sciences and is one of the foremost institutions of its kind. Through its collaborative approach, the Institute aims to drive scientific breakthroughs and actively mentor the next generation of researchers and scientists. Sharing these values, the NOMIS Foundation–a private Swiss foundation whose vision is to "create a spark" in the world of science by supporting insight-driven research–partnered with the Institute in 2008 to establish the NOMIS Center. There, research is underway to illuminate new ways to fight infectious disease and develop long-term immunity as well as to harness our immune system to fight cancer, neurodegeneration, and autoimmunity, and to promote health and suppress inflammation via the body's microbiome. A natural evolution of the Salk Institute–NOMIS partnership, the NOMIS Center Postdoctoral Fellowship Program was established in 2017.

## A symbiotic partnership

The NOMIS Center is the perfect headquarters for these highly promising investigations in health and immunity. "Establishing the NOMIS Center was fabulous timing," says Susan Kaech, professor and director of the NOMIS Center at Salk. The Institute was founded after Salk developed the polio vaccine, and is fundamentally rooted in advancing research to improve health and immunity. But as the Institute grew and evolved, several of the original pioneering scientists in that field retired, explains Kaech. "The NOMIS Foundation appreciated this opportunity to help rejuvenate and reestablish research areas in immunology and infectious disease as an integral part of the Salk Institute."

Kaech, who left Yale University for her current position 4 years ago, has seen her own research advance in this vibrant community of scholars. "You are entrenched and immersed with diverse scientists studying all different areas of biology. This makes it easier to cross over into new fields, to find those intersections for new ways to think about how our immune system functions."

> "You are entrenched and immersed with diverse scientists studying all different areas of biology. This makes it easier to cross over into new fields, to find those intersections for new ways to think about how our immune system functions."
>
> Susan Kaech

Susan Kaech
Director of the NOMIS Center at Salk

## Tapping into the brain trust

Andre Mu, a NOMIS Fellow from Australia, likens the NOMIS Center and the Salk Institute to a "brain trust." A computational biologist who previously studied subsurface microbial ecology, he pivoted for his postdoc into analyzing the microbial communities in the body and how they influence illness. The collegial atmosphere, combined with a collective sense of everyone wanting to help and support each other, was a game changer. "I transitioned from environmental sciences to infectious diseases. Thanks to the helpful research community and my mentors, I can apply my microbiome sciences toward understanding the intricate details of the immune system, host metabolic responses, and the likely microbial host for key metabolites," he says. "It gives a deeper perspective to my research."

Indeed, the NOMIS fellowships at the Salk Institute are about advancing high-risk/high-reward science through seemingly atypical and creative collaborations. "It's great because it gives you access to a network of people to discuss your research with and potentially achieve a breakthrough in the fight against cancer," says Pencik. Adds Kaech, "Because Salk is a small institution, barriers to collaboration with other scientists don't exist like they do at larger institutions."

NOMIS Fellow Dan Chen, a biologist from China, is studying the mechanisms of resistance to immunotherapy in patient-derived xenograft and glioblastoma models, with the goal of bolstering the immune system to fight tumors. Her research is highly collaborative and ambitious and holds potential treatment options for patients with glioma, a type of brain tumor. Given glioma's location in the skull, it is very difficult to treat and "no current immunotherapy strategy has been effective against glioblastoma," she says. Leveraging in vivo models, Chen and her team inject tumor cells into the brains of mice and check their immune response. One of the novelties in her approach revolves around $CD4^+$ T cells, a type of cell that plays an important role in the immune system, particularly the adaptive immune system. Chen demonstrated that $CD4^+$ T cells work with another immune cell type, microglia, which typically serve as a first line of defense against disease in the central nervous system; she is using her NOMIS fellowship to explore how $CD4^+$ T cells and microglia support each other in an antitumor state. Her work absolutely depends on engaging multiple experts across the Center and Institute.

Andre Mu

Lidia Jiménez

## Building community through shared experience

An interdisciplinary colloquium has helped Chen to build these critical bridges. Every other Friday, the NOMIS Center hosts a seminar series in which staff are invited to share their research journeys and openly ask for guidance and feedback on everything from the scientific approach and research techniques to their presentation and storytelling skills. "The seminars give us a good opportunity and platform to show our work," says Chen. And with that platform comes potential transformation. In her case, she gleaned insight from experts on glioma and cancer immunology, which gave her new ideas and perspectives that influenced her work for the better. The seminars exemplify a spirit of mutual assistance. "We share and help each other," she says. "It is quite a good platform for young scientists to communicate with each other and talk about science with people from different fields." Adds Jiménez, "Each time I present, I get great feedback that helps me to have fresh ideas and design new experiments to answer new questions. I couldn't do it alone."

Another example of the benefits of interdisciplinary cooperation comes from Pencik. As he designed his research agenda, the NOMIS Fellow thought he needed to examine immunotherapy opportunities in vivo. However, "putting cancer cells in mice is a big challenge, because most of the tumor cells are attacked and killed by the immune system," he says. "But to study immunotherapy, we are trying to develop effective immunotherapy models without rejection of tumor cells." A postdoctoral fellow working in one of the labs in the NOMIS Center gave a talk and shared how they used a direct CRISPR/Cas9 gene-editing system. It sparked a new idea, and suddenly Pencik saw a completely different workflow and novel set of techniques that would help him selectively knock out the genes important for immunotherapy resistance. "I used that technique and improved it and was able to make better cancer immunotherapy models," he says. Could he have providentially stumbled upon this method himself? Perhaps. But in this case, the answer was literally down the hall from him.

## Breaking down borders, building bridges

Another special feature of the NOMIS fellowships is that they are not dictated by national boundaries–these fellowships at Salk are open to one and all. Since residency is not a factor, the program is especially diversified and serves a greater purpose in bringing together the brightest minds, which makes science all the more powerful. Fellows are selected from the pool of existing postdocs at the Institute. They submit a proposal, and then it is reviewed by an interdepartmental panel; the top three make the cut. And the endowment that funds the NOMIS fellowships ensures long-term security.

But the NOMIS Center has an another mechanism to spur exciting research: strong alliances with and participation from the Salk development officers. They play a key role in the innovation ecosystem there, says Kaech. Not only do they identify granting and philanthropic opportunities, but they are involved in managing research programs. At Salk, "the power is in the scientists to pursue some of the most difficult problems facing humanity, and to carry this out, we rely entirely on grants and private philanthropy," says Kaech. Salk supports investigator-driven research initiatives by allocating fundraising resources to support philanthropy, community outreach, and the advancement of scientific communications. It's very much an "all-hands" environment and has led to some notable accomplishments, including raising more than $80 million in private philanthropy for plant biology research. For example, Salk's Harnessing Plants Initiative aims to develop scalable approaches to combatting climate change by optimizing the ability of common crop plants to capture and store excess carbon from the atmosphere in the soil. "This is an example of what happens when you put together a bunch of great brains to tackle hard global problems," notes Kaech.

Right: Jan Pencik
Left: Dan Chen

## Future focused

The open nature of the Foundation's fellowships, combined with exceptional assistance and support from the Salk team, positions these scientists for success by essentially granting them freedom: freedom to be creative, audacious, and daring in their endeavors, and freedom from burdens typical of postdoc appointments. Chen's fellowship has enabled her to participate in conferences and workshops and to travel to other institutions for collaborations that she might not have had the chance to be involved with otherwise.

Pencik is encouraged by the results he has garnered so far and is eager to take his research to the next level–translating it into treatments for patients.

Through his network at Salk, he is considering how he could partner with physicians and industry to develop drugs. He is not sure he would have gotten the inspiration to do so in another circumstance.

"It is always challenging to get funding for high-risk science, and even more so for basic research," says Jiménez. "It is very important to fund basic research, because then we can find new theoretical approaches that can lead to drugs for health and wellness." And it goes beyond funding: "The NOMIS fellowship helped me to feel supported and trusted, and this makes me stronger as a scientist."

NOMIS FOUNDATION | FELLOWS

salk®

## NOMIS Center Postdoctoral Fellowship Program at the Salk Institute

NOMIS fellowships at the Salk Institute support emerging early-career researchers who demonstrate the potential to lead groundbreaking, high-risk basic research. The fellowships provide postdocs with the freedom to explore big unanswered questions and work on high-risk interdisciplinary research with the support of outstanding scientific mentorship. The NOMIS Center typically appoints three postdoctoral fellows annually to work on health and immunity research. Learn more at www.salk.edu/science/research-centers/nomis-center/fellows.

NOMIS & Science
YOUNG EXPLORER AWARD

## NOMIS & *Science* Young Explorer Award

In 2021, the NOMIS Foundation and *Science*/AAAS launched the NOMIS & *Science* Young Explorer Award, which recognizes bold early-career researchers who ask fundamental questions at the intersection of the life and social sciences. The focus is on researchers who take risks to address relevant and exciting questions with creative approaches. The next annual call for applications will open in the spring. For more information, visit www.science.org/content/page/nomis-science-young-explorer-award.

> “It’s kind of weird. But it’s also nice.”
>
> **Suzanne Simard,** author of a best-selling book about how trees communicate, to CBC, about the many emails and touch of fame she received when her research was mentioned in the hit TV show *Ted Lasso*.

## IN BRIEF

*Edited by* **Jeffrey Brainard**

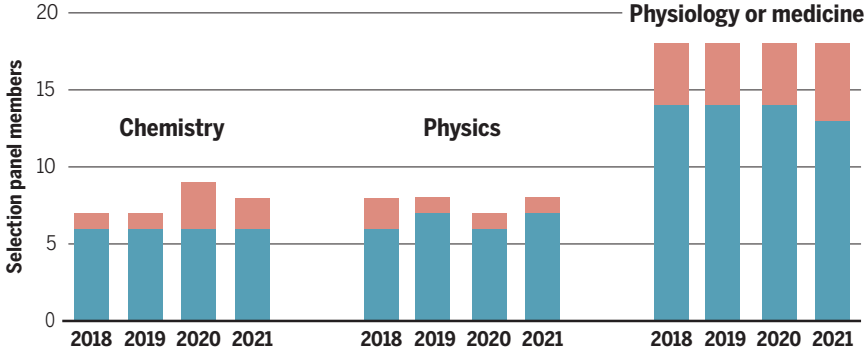
NOBEL PRIZES

# All-male Nobels raise gender equity issue

This year’s all-male slate of winners of the science Nobel Prizes—physics, chemistry, and physiology or medicine—disappointed but didn’t shock many in the scientific community. It’s in line with most of the Nobel Foundation’s 121-year history: In only 18 of those years has a woman been among the science laureates. But it also comes despite a push for more women to be nominated, which resulted in a doubling of the fraction of nominees for the science prizes in recent years who are women, according to members of two award-granting committees who shared their internal numbers with *Science*. This year, female scientists made up 13% of the nominees for the physiology or medicine prize and 7%–8% for the chemistry prize. The uptick in female nominees is encouraging, but the numbers remain low relative to the fraction of women in those fields. Some observers say more women should also be appointed to the committees that select prize winners.

**Nobel selection panels: mostly a male affair**

Over the past 4 years, relatively few women have served on the selection committees for science Nobel Prizes. ● Women ● Men




## A Nobel for organic catalysts

NOBEL PRIZES | This year’s Nobel Prize in Chemistry honors the discovery of small organic catalysts that mimic the work of conventional ones to develop new drugs and other compounds more efficiently. Benjamin List of the Max Planck Institute for Coal Research and David MacMillan of Princeton University share the award for discovering the “organocatalysts,” which speed intricate reactions that build complex functional molecules. Before List’s and MacMillan’s work, many catalysts contained metals, which can be toxic, or were large biological enzymes, which can be difficult to use outside cells. The dozens of pharmaceuticals developed using organocatalysts include paroxetine, used to treat anxiety and depression, and the influenza drug oseltamivir (Tamiflu).

## Natural studies in economics win

NOBEL PRIZES | Three pioneers in the use of natural experiments to determine causal relationships in economic activity won the Nobel Memorial Prize in Economic Sciences this week. Controlled studies of such behaviors and outcomes are often impossible or unethical. But in the 1990s, David Card of the University of California, Berkeley, showed he could probe questions such as whether increasing the minimum wage leads to higher unemployment by comparing neighboring communities—one that raised its minimum wage and another that did not. (The wage increase did not increase unemployment.) Joshua Angrist of the Massachusetts Institute of Technology and Guido Imbens of Stanford University shared the other half of the $1.14 million prize for methodological work on the conditions needed to make such studies valid.

## Third dose of China vaccines urged

COVID-19 | Two widely used COVID-19 vaccines made in China require a third dose to provide solid protection—but only in people ages 60 and older, a vaccine advisory panel of the World Health Organization (WHO) said this week. About 2 billion doses of the Sinopharm and Sinovac Biotech shots, both based on an inactivated version of the entire virus, have been administered in China, and another 1 billion have been exported. The Strategic Advisory Group of Experts on Immunization, which based its recommendation on real-world evidence of vaccine effectiveness, acknowledged a third dose will further strain global supply of vaccines and stressed that countries should first give two doses to the unvaccinated. The group insists that a third shot for the Chinese vaccines is not a “booster” shot—which WHO has said most people don’t need yet—but an “additional” dose to complete the “primary series.”

## Foreign influence probes tax NSF

RESEARCH SECURITY | The National Science Foundation’s in-house watchdog says the agency can’t keep up with the soaring number of allegations that scientists have ignored rules requiring them to disclose support from China when seeking NSF funding.

CREDITS: (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) NOBEL COMMITTEES

The goliath grouper attracts divers to Florida waters, but renewed fishing could leave fewer to see.

PHOTO: MICHAEL PATRICK O'NEILL/SCIENCE SOURCE

FISHERIES SCIENCE

## Overruling scientists, Florida board authorizes fishing of the vulnerable goliath grouper

A Florida commission voted last week to again allow fishing of the Atlantic goliath grouper, a storied species that the state had protected from any harvests since 1990 to allow it to recover from overfishing. The panel acted despite objections from fisheries scientists that the fishing will halt the population's fledgling recovery and lacks any scientific justification. More than 90 fisheries scientists in the United States and other countries endorsed a letter in May to the Florida Fish and Wildlife Conservation Commission recommending against the move. Beginning in 2023, it will allow harvesting of up to 200 juveniles annually in most of the state's waters, which extend up to 14.4 kilometers offshore. The International Union for Conservation of Nature lists the species (*Epinephelus itajara*) as vulnerable, and the U.S. government bans fishing it in federal waters. With a life span of up to 37 years, it can grow to more than 2 meters and 450 kilograms.

Allison Lerner, NSF's inspector general, whose office investigates such allegations, told Congress last week that tracking foreign influences "now makes up 63% of our caseload" and that "we don't have the resources we need" to deal with a problem she says wasn't on her agenda until 4 years ago. Lerner also revealed that NSF so far has recovered $7.9 million from 23 grantees at 21 institutions who were found guilty of breaking NSF's rules on disclosure. The agency also imposed other sanctions, including making violators ineligible to receive federal funding or to serve as a reviewer on grant applications. All but one of the cases resolved have involved an award to a scientist with links to China, and an analysis by *Science* of publicly available data suggests that as many as 80 cases remain open.

## Biden to restore environment rule

ENVIRONMENT | Fulfilling a campaign promise, the Biden administration last week said it will reverse changes by former President Donald Trump to federal rules for reviewing the environmental impact of major development projects. Builders of highways, dams, and other publicly funded projects will again have to consider their projects' cumulative, long-term impacts on ecosystems, climate, and people living nearby, the White House said. The Biden administration is also expanding opportunities for public comment on the projects. Trump argued his changes were needed to speed project approvals, but the Biden White House says they gutted safeguards. In a related move last week, 23 federal agencies, including NASA and the Department of Energy, released plans for ensuring their operations can "adapt to and are increasingly resilient to climate change impacts." The public can comment on the agency plans until 6 November.

## U.S. official broke journal policy

PUBLICATION ETHICS | A top journal retracted an article last week because it was edited by a high-profile scientist now working in the White House who had ties with the article's co-authors. Jane Lubchenco, a marine biologist appointed in March as deputy director for climate and the environment at the White House Office of Science and Technology Policy, had co-authored a related paper with authors of the now-retracted article, one of whom is her brother-in-law. The retracted article, published in October 2020 by the *Proceedings of the National Academy of Sciences* (*PNAS*), concluded that expanding marine preserves could improve fisheries harvests. *PNAS*'s retraction notice, first reported by Retraction Watch, said Lubchenco's ties violated its policies; the notice also cited an error in the article's data. Its authors plan to submit the article to another journal. Lubchenco, who led the U.S. National Oceanic and Atmospheric

BY THE NUMBERS

**120,000**

U.S. children who lost care when a parent or grandparent died from COVID-19, according to statistical modeling. More than half the children are Black or Hispanic. (*Pediatrics*)

Administration from 2009 to 2013 and is a former president of AAAS, *Science*'s publisher, provided a statement: "I greatly regret the poor decision I made quite some time ago in agreeing to edit this paper."

## U.S. panel revises aspirin guidance

BIOMEDICINE | People age 60 or older should not start taking a daily baby aspirin to prevent heart disease, an influential federal advisory panel said this week, backing away from its previous recommendation. New research has bolstered evidence that aspirin's risks, such as gastrointestinal and brain bleeding, outweigh its preventative benefits, the panel said. People between ages 40 and 59 at high risk for a first heart attack or stroke should consult with their physicians about whether to start taking aspirin as a preventive, the U.S. Public Health Service's Preventive Services Task Force concluded in a draft recommendation. The guidance, which is expected to be finalized after a comment period, updates the task force's 2016 report endorsing preventative low-dose aspirin for at-risk people ages 50 to 59 and advising that those between ages 60 and 69 decide for themselves.

## Reactor leak linked to staffing

NUCLEAR SAFETY | A shortage of experienced staff, exacerbated by the pandemic, contributed to a 3 February release of radiation from the small research reactor at the National Institute of Standards and Technology (NIST), according to a report last month. On 4 January, workers refueling the 52-year-old reactor apparently failed to latch into place one of 30 fuel rods, and flowing cooling water pushed it out of place, investigators found. When the reactor restarted 1 month later, the rod overheated. NIST said monitoring at the boundary of its 234-hectare campus indicated that the maximum dose a member of the public could have received from the leak was less than 0.5 millirem, or 1/200th of the regulatory limit. NIST also reported that 10 staff members were exposed to radiation levels about what a person would receive during a computerized tomography scan. Poor training and a lack of experience contributed to the accident, investigators found. In 2011, nine of 11 operators had more than 20 years of experience, compared with three of 22 now.

## Chimp refuge gives vaccine

ANIMAL WELFARE | Project Chimps, a U.S. animal sanctuary in Georgia, this month started to administer an experimental vaccine to its 77 chimpanzees to protect them from COVID-19. Zoos, conservatories, and aquariums in the United States have administered the vaccine to their animals. But Project Chimps is the first U.S. sanctuary accredited to house great apes retired from research labs, the pet trade, and the entertainment industry to do so. Five other accredited great ape sanctuaries—including the largest, Chimp Haven in Louisiana—are taking a wait-and-see stance about the animal-only vaccine, made by the manufacturer Zoetis; they want to learn more about its safety and efficacy, and believe the current risk of their chimps contracting COVID-19 is low. To protect the animals, many staff members at those sanctuaries and at Project Chimps have themselves been vaccinated against COVID-19.

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

THREE QS

## Drawings celebrate diverse scientists' 'superpowers'

Ask children around the world to draw a scientist, and many sketch an older white man. CodeMakers, an education nonprofit in South Africa, is working to break down such stereotypes in a new way: disseminating trading cards and posters (above) that depict real, local scientists as Marvel Comics–style superheroes, hoping to stimulate children's imagination and ambition. The project, Super Scientists, has donated these printed materials to schools throughout Africa. Founder Justin Yarrow, an educator, says the United States is second only to South Africa as a source of traffic to the project website. A longer version of this interview with Yarrow is available at scim.ag/ScienceSuperhero.

**Q: What's your goal? Why superheroes?**
**A:** In South Africa [and elsewhere], the demographics of established science don't reflect the demographics of the country. We want to have young Black people see these amazing up-and-coming scientists. We wanted to show off the power of scientists—astronomers' ability to see back in time, or biotechnologists' ability to create new forms of life. Superpowers are a language that young people understand.

**Q: How did you choose the 49 scientists now featured?**
**A:** [We wanted] scientists in different fields, so we have paleoscientists, oceanographers, astronomers. Another component is making sure they are representative of the kids we work with. Our most recent scientist, Mpho Kgoadi, aka Cosmic Dawn, is a disabled astrophysics Ph.D. student at the University of the Witwatersrand. We also are profiling scientists from Kenya, Uganda, Ghana, Rwanda, Zambia, and Botswana. South Africa is 80% Black and 8% white, [as are] the Super Scientists. I'm proud to say that they're 65% women.

**Q: What impact have you seen?**
**A:** We hear anecdotal things. We've heard from a University of Cape Town oceanography Ph.D. student—Kolisa Sinyanya, aka Nitro—that she's received phone calls from people in tears, telling how when they were growing up, they thought only white men could be scientists. Now they can show these materials to their kids: this powerful, strong image of a scientist.

ILLUSTRATION: CLYDE BEECH FOR CODEMAKERS/SUPER SCIENTISTS

Aron Kaleso, a health officer, administers the RTS,S malaria vaccine at Bereu Health Center in southern Malawi, one of three African countries included in its pilot rollout.

PHOTO: THOKO CHIKONDI

GLOBAL HEALTH

# WHO gives first malaria vaccine the green light

## Widespread distribution in Africa set to begin despite lingering questions about efficacy

*By* **Gretchen Vogel**

The world is about to get its first malaria vaccine, after a 35-year quest. In a momentous decision, the World Health Organization (WHO) on 6 October recommended the wide rollout of RTS,S, produced by GlaxoSmithKline (GSK), to protect children in Africa.

Data from a pilot rollout involving more than 800,000 children in three countries convinced a panel of malaria and vaccine experts advising WHO that the vaccine, also called Mosquirix, is safe and effective, and should be offered to children in African regions that have moderate or high malaria transmission. (The vaccine targets the malaria parasite prevalent in Africa, *Plasmodium falciparum*.)

The announcement "is an historic moment in the fight against malaria," says Corine Karema, a member of WHO's advisory group for the pilot rollout and former director of Rwanda's National Malaria Control Programme. It opens the way for countries to decide how to use the vaccine, the first ever to be approved for a parasitic disease, as part of their malaria control programs. It also allows funders to pledge financial support for ramping up vaccine production and distribution.

RTS,S is far from perfect: It requires four doses, only provides roughly 30% protection against severe malaria, and has a price tag that could be hard to afford. Clinical trials also raised possible safety concerns, and some researchers caution that studies so far may have missed some of its downsides. But malaria kills an estimated 260,000 children in Africa each year, a number that was falling rapidly between 2004 and 2015 but has since leveled off. "We need new tools to get malaria under control," WHO Director-General Tedros Adhanom Ghebreyesus said at a press conference announcing the decision.

RTS,S, developed in the 1980s by scientists at a predecessor to GSK, contains a piece of a *P. falciparum* protein linked to a protein from the hepatitis B virus, added to trigger a stronger immune response. The vaccine is designed to block the parasite's ability to infect the liver and mature there.

In 2003, RTS,S was the first malaria vaccine to enter large trials. The initial results were encouraging, but not outstanding: It cut the risk of clinical malaria by about one-third. The vaccine worked better when given to children starting at 5 months of age than in younger babies, which means it can't piggyback on the standard infant immunization schedule. More concerning, there were hints it might increase the risk of developing cerebral malaria or catching meningitis. Another analysis of the data concluded overall mortality increased in girls who received the vaccine compared with controls.

Still, the data were promising enough for the European Medicines Agency to approve RTS,S for use in children 6 weeks to 17 months old in July 2015. But a few months later, WHO's vaccine advisory panel decided the safety concerns and logistical hurdles needed more study and recommended a pilot rollout to better understand the vaccine's real-world impact. In response, WHO established the Malaria Vaccine Implementation Program, which began to administer the vaccine to children in selected regions in Ghana, Malawi, and Kenya in 2019.

Initial data from the first 2 years of the program, presented to WHO's panel last week, showed the logistics hurdles could be surmounted: Between 62% and 67% of eligible children in the rollout regions received the first three doses. (The fourth dose is administered 12 to 18 months after the third dose, so most children have not yet received it.) The vaccine worked about as well as in earlier studies, lowering rates of hospitalization for

severe malaria by about 30%. No increase in meningitis or overall mortality was observed. Surveys showed vaccinations reached up to 75% of children who don't sleep under a bed net—a common measure for controlling the disease. That shows it can complement existing tools, says Abdoulaye Djimdé, a malaria expert at the University of Bamako in Mali.

The WHO panel also considered data from a trial, published last month in *The New England Journal of Medicine*, in which Djimdé and his colleagues tested the vaccine both alone and in combination with regular doses of antimalarial drugs, given prophylactically just before the rainy season in Mali and Burkina Faso. Over the 3-year study period, the vaccine was as effective as the prophylaxis, and children who received both had a roughly 70% lower risk of severe malaria than those who only received one or the other.

Taken together, the data convinced WHO's advisory panel. But Christine Stabell Benn, who studies vaccine impacts at the University of Southern Denmark, cautions that protection wanes significantly after 6 months, so the data from the first 2 years of the pilot are "zooming in on the period of maximal benefit and minimal harm" from possible side effects. Those could include rebound infections, in which children develop malaria after vaccine protection wears off.

Stabell Benn also notes that follow-up time in the pilot rollout was too short to measure whether the vaccine saved lives. The data showed a 7% decrease in mortality from any cause, but that was not statistically significant. "They set out to detect a 10% reduction in all-cause mortality. They haven't achieved that," she says. "And that's during the period it's supposed to be having the biggest effect."

The vaccine's cost also could prove a drawback. The estimated $5 per dose is a bargain compared with many vaccines used in rich countries, but it still means countries will need to carefully consider how the vaccine fits in with other, less expensive malaria prevention tools, says Catherine Pitt, who studies the economics of malaria at the London School of Hygiene & Tropical Medicine. The approved rollout is "fantastic," Pitt says, but "the budget doesn't seem to be getting bigger, so hard choices have to be made."

WHO leaders say studies of the best use of the vaccine will continue. "We'd all like a magic bullet," says Dyann Wirth, a malaria researcher at Harvard University and chair of WHO's Malaria Policy Advisory Group. "But this is a very complex disease that has evolved ways to evade the immune system," which makes vaccine development especially challenging. "We're not saying this is the end," Wirth says. "I'm hoping this is the beginning of a renaissance of vaccine development in the malaria field." ■

PHOTO: ULET IFANSASTI/GREENPEACE

CONSERVATION

# Massive road project threatens New Guinea's biodiversity

## The 4000-kilometer Trans-Papua Highway could turn the island's western half into an "El Dorado for investors"

*By* **Dyna Rochmyaningsih**

Seventeen years ago, botanist Miriam van Heist of Wageningen University and a team of local collaborators set out to explore plant diversity in the Foja Mountains in Western New Guinea. It was an uncharted realm: Of the 487 species they found, nearly one-third had never been described anywhere. "The results illustrate how much remains unknown about New Guinea's vegetation," the authors wrote in a 2010 paper. More such studies were needed, they concluded, "so that conservation can be planned in advance of impending threats—otherwise we may lose species and never know what we have lost."

Those threats have now arrived, according to a study based on satellite imagery and maps from local governments that was published last month in *Biological Conservation*. Between 2001 and 2019, it found, 2% of Western New Guinea's old-growth forests were cleared to make room for palm oil plantations, roads, mines, and towns and cities. And the pace of development is set to accelerate, the researchers concluded. The Trans-Papua Highway, a road system under construction that connects major urban areas, could speed up development dramatically.

Western New Guinea, also known as Indonesian New Guinea or Tanah Papua, is still largely untouched, with a bewildering variety of landscapes—from mangroves, savannas, and lowland forests to alpine grassland and tropical glaciers—and a welter of endemic species, including 29 bird of paradise species. But now, "These forests are at a critical turning point for the future because Papua is the new El Dorado for investors in land development and in extractive industries," says landscape ecologist David Gaveau, the study's lead author. Gaveau, who was deported from Indonesia in 2020 after publishing estimates of the damage from Indonesia's 2019 wildfires that far exceeded the Indonesian government's numbers (*Science*, 14 February 2020, p. 722), now runs the Tree Map, a France-based company that analyzes the loss of rainforests.

Indonesia's government incorporated Western New Guinea, formerly known as Irian Jaya, in 1969, 8 years after the Papuans formally declared independence from the Dutch. (The eastern half of the island is part

The Trans-Papua Highway, still under construction, cuts through a pristine rainforest in Papua province in Indonesia.

of a separate country, Papua New Guinea.) About the size of Iraq, it has a huge ethnic diversity, with an estimated 200 Indigenous tribes and as least as many languages.

It is also by far Indonesia's poorest region. Its two provinces, Papua and West Papua, have poverty rates of 27% and 22%, respectively, according to the country's statistical agency. The Indonesian government promises the Trans-Papua Highway will help alleviate that poverty and distribute infrastructure development more fairly. Snaking almost 4000 kilometers across the island, the highway connects the coastal cities of Sorong, Jayapura, and Merauke (see map, right).

But it is driving conflict as well as development. The mining and palm oil companies it attracts have cleared the land of Indigenous Papuans, often without their consent, which has sparked anger, protests, and violent attacks. Two people were killed and machinery was set on fire during the most recent one, on 8 September in Yahukimo regency. A similar attack in 2018 killed 20 people. The West Papua National Liberation Army—the military arm of the Free Papua Movement, a separatist group—has claimed responsibility for the actions. The attacks have delayed construction; so far only about half of the route has been paved.

The road's most controversial stretch, 190 kilometers long, cuts right through Lorentz National Park, a World Heritage Site and one of the largest protected areas in Southeast Asia. In August, UNESCO called on Indonesia to close this section to traffic while the government draws up a plan to mitigate the road's impact, as it is obliged to do under UNESCO rules.

But the highway will have important impacts elsewhere as well. The study by Gaveau and his colleagues showed a clear correlation between road construction and the massive expansion of industrial plantations in Merauke and Boven Digoel and rapid growth of Kenyam and Dekai, two towns in the heart of Papua. The highway has also attracted artisanal gold miners, who often extract gold by dissolving it in mercury, which they then release into the environment.

The two semiautonomous provinces have pledged to preserve their natural riches. In the 2018 Manokwari Declaration, they aimed to set aside at least 70% of their lands for protection while improving the livelihoods of Indigenous communities. "That is an ambitious target," says Yubelince Runtuboi, a forest policy researcher at the University of Papua New Guinea. Many plantation concessions were granted before the declaration was made, and they may be difficult to revoke. And local governments are keen to develop new villages and towns in their quest for economic development, Runtuboi says. As a result, she wrote in a 2018 paper, Papua's lowland forests "are likely to come under immense pressure."

The rapid deforestation elsewhere in Indonesia, the world's largest producer of palm oil, does not bode well. A model developed by Gaveau and his co-authors showed that under the worst-case scenario, Western New Guinea would lose about 4.5 million hectares of forest, an area the size of Denmark, by 2036. "Given the large amounts of flat lowland forests, and the suitable climate for palm oil, Papua could easily become the next agricultural frontier for years to come," Gaveau says.

To what extent the Trans-Papua Highway will alleviate poverty, meanwhile, is unclear. A study by Cahyo Pamungkas, a sociologist at the Indonesian Institute of Sciences, showed it has primarily helped migrants from elsewhere in Indonesia distribute commodities to rural areas, sometimes outcompeting local goods. Ordinary pigs raised by migrants are replacing Wamena pigs, a type of wild boar bred locally, for example. The better the surface of the road, the more logging takes place nearby, Pamungkas found. The highway "benefits the capitalists more," he says.

Papua's mineral riches—including Grasberg, the world's largest gold mining pit—have barely benefited the local population either, critics argue. The development of a new gold mine close to the Trans-Papua Highway, called the Wabu Block, is now the topic of heated debate and protests. (Activists recently alleged that Luhut Pandjaitan, a retired general and Indonesia's coordinating minister for maritime affairs and investment, has business interests in the mine, a charge Pandjaitan denies.)

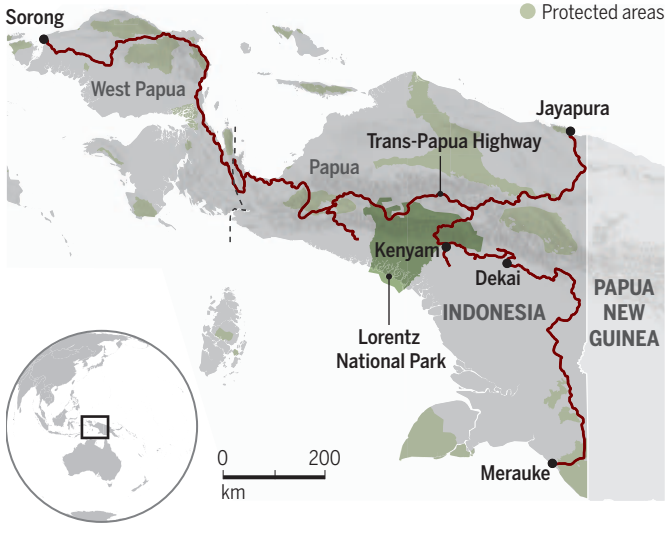
Still, the picture is not black and white, says Sophie Chao, an anthropologist at the University of Sydney who spent 2 years living with the Marind people in Merauke. The new road has led to the destruction of forests to make room for oil palm plantations, including sacred places that were cleared without the Marind's consent, she says. But the road also "brings new people, new ideas, new commodities to the villages," Chao says, as well as humanitarian aid and jobs. "It's a mix of hope and anxiety," she says.

Balancing the positive and negative impacts requires more dialogue with Indigenous people than there has been so far, Gaveau and his colleagues say in their new paper. Theo Hesegem, who heads the Advocacy Network for Upholding Law and Human Rights of Papua Central Highlands, agrees. "It's no use if they build the infrastructure without listening to us," Hesegem says. "The Trans-Papua Highway will not be finished." ■

Dyna Rochmyaningsih is a science journalist in Palembang, Indonesia.

## On the road to development

The Trans-Papua Highway, which connects three coastal centers in Western New Guinea, has attracted plantations and mining and spurred growth of towns and cities. A 190-kilometer segment cuts through Lorentz National Park.




CREDITS: (GRAPHIC) K. FRANKLIN/*SCIENCE*; (DATA) D. GAVEAU *ET AL.*, *BIOLOGICAL CONSERVATION*, 261, (2021) 10.1016/J.BIOCON.2021.109225

ANIMAL RESEARCH

# USDA revokes license of leading chinchilla supplier

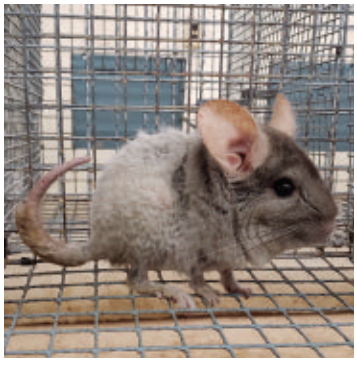
## Hearing researchers may be left scrambling for animals

By **Meredith Wadman**

In a rare decision, an administrative law judge in the U.S. Department of Agriculture (USDA) this week permanently revoked the license of a leading U.S. supplier of chinchillas for research. Judge Jill Clifton called the record of animal welfare violations at Moulton Chinchilla Ranch (MCR) in Chatfield, Minnesota, "absolutely astounding" and pronounced its proprietor, Dan Moulton, "unfit" to hold a license. She also fined Moulton $18,000.

Her decision, the first time since 2016 that USDA has revoked the license of a research-related animal facility, culminates a hearing during which USDA lawyers argued that Moulton willfully violated the Animal Welfare Act (AWA) for years, leaving sick and injured animals without veterinary care and housing them in decrepit, filthy conditions.

Moulton, a lawyer in Rochester, Minnesota, did not respond to requests for comment. He told the judge in his closing statement: "I am not taking violations loosely. I have tried to be responsive to the USDA inspectors. I have cooperated with them."



A chinchilla with tail fur missing, allegedly photographed at Moulton Chinchilla Ranch in 2020.

He could appeal the decision within USDA or, if that fails, in the U.S. court system. But it seems likely to mark the end for a large facility that has been a go-to source for U.S. researchers whose work relies on the docile South American rodents. Chinchillas' relatively large ears, which are anatomically similar to humans', make them a prime model for studies of hearing and ear infections.

"Many institutions across the country are already feeling the pinch of insufficient chinchillas to explore diseases affecting hearing. The loss of any single provider further impedes the research opportunities," says Ronald Banks, a veterinarian and director of comparative medicine at the University of Oklahoma Health Sciences Center, which stopped using Moulton chinchillas in 2018 because of their poor quality.

In 2018, USDA, which is charged with enforcing the AWA, filed a formal complaint against Moulton alleging 213 violations of the act between 2013 and 2017. He categorically denied the allegations. That led to the current hearing.

Aaron Rhyner, a veterinarian who is an assistant director in field operations at USDA, told the court last week: "I have probably been associated with several thousand inspection reports and several thousand facilities. … This is the worst compliance history I have seen."

In May, the Moulton ranch housed 748 chinchillas, when another USDA inspection documented sick, untreated animals; rusted, broken feeders; and ammonia—from accumulated excrement—that burned inspectors' throats and eyes.

In 2019, U.S. researchers used 1250 chinchillas, according to USDA data. Moulton testified during the hearing that he was the sole supplier of research chinchillas. But Rhyner challenged that. "There [are] currently 85 facilities licensed with us that have chinchillas," he said.

Veterinarian Christoph Mans and his team at the University of Wisconsin, Madison, recently published a paper using animals from a small Illinois supplier. But, he says, "Sourcing chinchillas has become increasingly challenging."

Earlier this year, Moulton was also the subject of a criminal investigation under Minnesota's animal cruelty law, but county attorneys decided to leave the matter in the hands of USDA.

Russ Mead, an animal law attorney at Lewis & Clark Law School, deplored MCR's long history of AWA violations and what he called USDA's glacial pace. "It took 213 violations to bring about this license revocation. The system is broken," he said. ■

PALEONTOLOGY

# Biden restores key monuments

## Move reinstates protections for fossil sites, marine zone

By **April Reese**

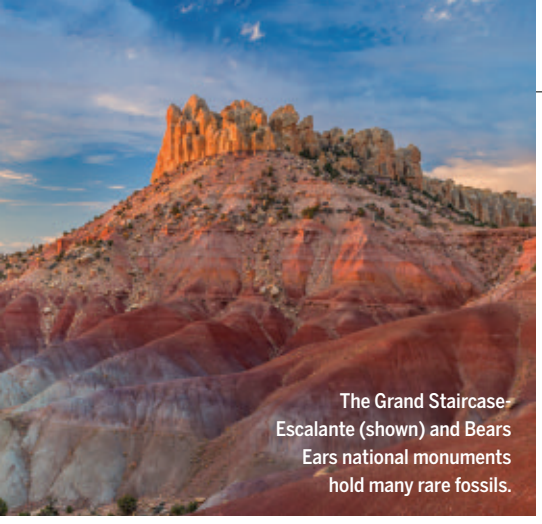
In a long-awaited decision, President Joe Biden last week restored Utah's fossil-rich Bears Ears and Grand Staircase-Escalante national monuments to some of their previous boundaries, 4 years after former President Donald Trump made drastic cuts to both. Secretary of the Interior Deb Haaland, a member of the Pueblo of Laguna in New Mexico, was visibly moved at the proclamation signing ceremony she attended with Biden. "Thank you, Mr. President for protecting the homelands of our ancestors," she said, her voice cracking.

The boundary reversals garnered praise from paleontologists, conservationists, and Southwestern tribes that trace their ancestry to the Bears Ears region. Together, they had persuaded former President Barack Obama to protect 547,000 hectares of the area. In establishing the monument in 2016, Obama cited the significance of Bears Ears to multiple tribes and its rich paleontological and archaeological record. Obama also expanded the Grand Staircase-Escalante National Monument, created in 1996.

The Trump administration, however, reduced the size of Bears Ears by 85%, leaving 82,000 hectares split into two separate units, and cut the 768,902-hectare Grand Staircase-Escalante by almost half. It did add 4532 new hectares to Bears Ears, which are retained in the Biden designation.

The monuments' colorful buttes and badlands harbor rare fossils that chronicle hundreds of millions of years of Earth's history,

PHOTOS: (LEFT TO RIGHT) PETA; JACK DYKINGA/MINDEN PICTURES



The Grand Staircase-Escalante (shown) and Bears Ears national monuments hold many rare fossils.

and some key sites, such as the Valley of the Gods, were left outside the boundaries after Trump's cuts. Those areas, researchers wrote last year in a review of the monument's scientific value, could provide "unique insights" into important paleontological periods, including the rise of dinosaurs some 200 million years ago, at the end of the Triassic. Bears Ears is thought to hold the world's most abundant cache of fossils from that period.

"It's very good to see those sites back within the boundaries," says Robert Gay, an independent paleontologist who has studied Triassic fossils in Bears Ears for years. Biden's move, Gay adds, means some research projects on monument resources will again be eligible for federal funding. But he had hoped the Biden administration would enlarge the monument to include an understudied area with known paleontological sites. "I'm a little disappointed."

Biden also restored protections to Northeast Canyons and Seamounts Marine National Monument, off the coast of New England. Obama created the monument, the only one in the U.S. Atlantic, in 2016, but in 2020 Trump issued an order to lift restrictions on commercial fishing there. Commercial fishing for lobster and red crab in the 1.3-million-hectare area will now be phased out by 2023.

Although supporters of the expanded monuments were quick to declare Biden's proclamations the last word on the sites, his decision, like those of his predecessors, is likely to be challenged in court. Many Utah lawmakers have opposed the creation or expansion of the Bears Ears and Grand Staircase-Escalante monuments, arguing that doing so through the Antiquities Act of 1906 amounts to presidential overreach.

Steve Bloch, legal director of the Southern Utah Wilderness Alliance, is doubtful a new legal challenge will succeed. "They already litigated and lost" previous cases challenging the monuments, he notes. ■

April Reese is a journalist in Santa Fe, New Mexico.

CONSERVATION

# China embraces global biodiversity push—cautiously

## Xi's silence on 30% set-aside target disappoints advocates

*By* **Dennis Normile**

China took advantage of an international stage this week to show it's serious about protecting the planet's threatened species. But what it had to offer left some conservation advocates disappointed.

In a 12 October video message to the 15th meeting of the Conference of the Parties to the United Nations Convention on Biological Diversity, President Xi Jinping announced a new fund to support biodiversity protection in developing countries, which China will seed with a donation of $232 million. The move "should jumpstart an urgently needed conversation on biodiversity finance," says Li Shuo, a senior policy adviser for Greenpeace East Asia in Beijing. But Xi didn't mention a goal some other countries have embraced: setting aside 30% of the world's land and sea areas for conservation. That goal "needs support from China" to be adopted, Li says.

> **"China wants to be seen as a leader" in global conservation.**
>
> **Alice Hughes**, Xishuangbanna Tropical Botanical Garden

Expectations were high because "China wants to be seen as a leader in this area," says Alice Hughes, a conservation biologist at the Xishuangbanna Tropical Botanical Garden in southwestern China. China also recently pledged to achieve carbon neutrality by 2060 and to end Chinese financing for overseas coal plants.

The meeting, originally scheduled for 2020 but repeatedly postponed because of the pandemic, has been split in two, with preliminary sessions online this week and actual negotiations next spring, when delegates hope to gather in person in Kunming, in southwestern China. On 13 October, after *Science* went to press, delegates were expected to adopt the Kunming Declaration, calling on the world to approve an ambitious biodiversity "framework" that is still being negotiated.

The latest draft of the framework lists 31 targets and milestones to reach by 2030. In addition to the 30% set-aside, it calls for halving the introduction of new invasive species, eliminating plastic pollution, and increasing financial support for conservation efforts in developing countries by $200 billion.

The world fell short of the less ambitious "Aichi targets," set at the convention meeting in 2010 in Japan. But Stuart Pimm, a conservation scientist at Duke University, thinks the draft Kunming goals are realistic. Over the past 30 years, "we have dramatically increased how much of the planet we've protected," he says.

China has made significant progress protecting its own biodiversity. It has 18% of its land area under protection and scored a notable achievement in 2016 when the International Union for Conservation of Nature upgraded the status of the iconic giant panda from "endangered" to "vulnerable." Protecting panda habitat benefited other species as well, Pimm says, including the takin, a goatlike animal considered vulnerable, and endangered snub-nosed monkeys.

But China has made less progress battling invasive species, and whereas air quality in the country has improved, "water pollution is not getting any better," says ecologist Johannes Knops of Xi'an Jiaotong-Liverpool University. Pollution, boat strikes, and fishing bycatches all likely contributed to the decline of the baiji, a river dolphin found only in the Yangtze River watershed that hasn't been spotted since 2002 and may be extinct.

This week's meeting is "like the starting gun in a race," says Georgina Chandler, an international policy specialist at the United Kingdom's Royal Society for the Protection of Birds. "We've got 6 months to finish the negotiations." The 30% conservation target could survive the discussions. But Xu Haigen of the Nanjing Institute of Environmental Sciences cautions that the numerical goal is not enough. "Areas should be carefully chosen for maximal impact on biodiversity conservation and not just designated for the purpose of boosting percentages."

And China's support for the 30% goal will be key, Li says. The country "will need to make up its mind before next spring." ■

With reporting by Bian Huihui.

EARTH SCIENCE

# New observatory to bring soil up from the deep

## Ecotron will move 3-meter cores into the laboratory for global change experiments

*By* **Erik Stokstad**

Just a meter or two down, below the topsoil that nurtures crops, is a little known part of the ecosystem that may be critical to the planet's climate future. But this deep soil is surprisingly hard to study. It helps to know the right backhoe operator, and even then extracting samples without disturbing their structure or inhabitants is tricky. "The deeper you go, the harder it is," says Daniel Richter, a soil scientist at Duke University.

Last month, the U.S. National Science Foundation announced funding for a new $19 million research facility, called the Deep Soil Ecotron, that aims to make studying this frontier easier. The initial design for the lab, to be built over the next 5 years at the University of Idaho, calls for 24 richly instrumented soil columns topped with airtight chambers for vegetation. These ecosystems-in-a-lab, or ecounits, will allow researchers to manipulate environmental conditions down to 3 meters. Surprises are assured. "It's kind of like when people launched the first deep-sea submarine," says Zachary Kayler, a co–principal investigator (co-PI) and biogeochemist at the University of Idaho. "The possibilities are endless."

Ecotron scientists and visitors examine columns of soil in this artist's concept.

Ultimately, researchers hope to use the ecotron facility to study a wide range of questions, including how the deep soil might release carbon and accelerate climate change, how soil microbes and plants interact, and how torrential summer rains and hard winter freezes influence the birth and growth of soil. The center could also become a testbed for new sensors that reveal what's happening in deep soil in the real world. Planners hope to attract researchers from across the United States and host multiple experiments at once. "This facility has potential that's incredibly exciting," says Sharon Billings, an ecosystem scientist at the University of Kansas, Lawrence.

About a dozen ecotron facilities exist, nearly all in Europe. But they allow researchers to tinker with just one or two soil conditions, usually precipitation. In contrast, the new U.S. lab will allow researchers to manipulate a host of factors throughout the entire columns, including temperature, moisture, and carbon dioxide concentrations (which influence rock weathering and soil formation). They will be able to simulate the upwelling of groundwater and perhaps even the freeze-thaw cycles that can speed soil development.

The ecotron also opens the door to studying deep life, including invertebrates, says Nico Eisenhauer, an ecologist at the German Centre for Integrative Biodiversity Research who advises the U.S. ecotron. Little is known about the activity of these organisms, he says, but their hidden behavior could be monitored in the new ecounits. Acoustic sensors, for instance, can track earthworm activity, whereas buried tubes with small entrances collect and count tiny invertebrates that stumble in.

Researchers who study soil carbon note that topsoil—typically the first 30 centimeters—has been well studied because it is crucial for agriculture. Scientists have assumed this shallow, root- and microbe-rich layer holds most soil carbon. But a few studies of the deeper, vaster subsoils suggest they might contain at least as much carbon, some of it carried down by roots. "It's the forgotten half of soil organic carbon," says soil biogeochemist Michael Schmidt of the University of Zurich.

Studying how that carbon might respond to future climate change is a key goal for the ecotron, which can simulate changing conditions "with a degree of control I've never seen before," says Emma Aronson, a soil scientist at the University of California, Riverside. Deep soils are expected to warm as Earth heats up, and field experiments suggest that could cause soil to release carbon into the atmosphere 30% to 50% faster than today, accelerating global warming. Precipitation shifts caused by climate change will also have an impact. For example, wetting of desert soils, especially those that have been polluted with reactive nitrogen, can lead to much greater emissions of nitrous oxide, a potent greenhouse gas.

Idaho is "really a cool place" to locate the ecotron lab because it boasts a diverse range of soil types, says the project's PI, microbial ecologist Michael Strickland of the University of Idaho. He and colleagues are still mulling how many and which types of soils to place into the columns. "Capturing dynamics in different types of soils will be key," says soil biogeochemist Asmeret Asefaw Berhe, who is awaiting Senate confirmation to lead the Department of Energy's Office of Science.

One challenge will be extracting and moving the selected soils. Field crews must disturb them as little as possible, Richter says, because their structure influences the movement of gases, water, and particles. Strickland hopes new equipment from the company building the facility will be able to extract long cores intact. But no matter how tender the touch, researchers expect the samples to have severed roots and surfaces that will be exposed to oxygen for the first time in centuries or millennia. It could take years for the extracted soil to return to equilibrium, says Pete Smith, a soil scientist at the University of Aberdeen.

Despite the unknowns, excitement is already building among soil scientists. "The fun will really start," says co-PI Rodrigo Vargas, an ecosystem ecologist at the University of Delaware, Newark, "once it's built." ■

IMAGE: SECOND BAY STUDIOS

# THE SNIFF TEST

## Courtroom testimony about dogs detecting dead bodies keeps sending people to prison—even without physical evidence. Critics say the science is lacking

*By* **Peter Andrey Smith**

PHOTO: DIGITAL FIRST MEDIA GROUP/*OAKLAND TRIBUNE* VIA GETTY IMAGES

Dogs are renowned for their sense of smell, a reputation that carries increasing weight in court.

Mark Redwine kept mementos of his youngest son in the cab of his truck. The boy's name, Dylan, is tattooed on his left shoulder. He distributed missing person posters at truck stops. In November 2012, his 13-year-old son had arrived to spend Thanksgiving at Redwine's cabin, near the San Juan National Forest in southwestern Colorado. Redwine told an FBI investigator that he left Dylan asleep on the couch when he went to town. When he returned, the boy was gone.

More than 6 months later, hikers found skeletal remains on a trail high in the desert-clear air, about 16 kilometers from Redwine's cabin. Authorities confirmed the bones belonged to Dylan, but the local coroner could not determine cause of death. To go missing is not a crime, but Redwine's ex-wife, Elaine Hall, and the couple's oldest son came to suspect Redwine did some-

Carren Corcoran Gummin's testimony about the "indications" provided by her German shepherd, Molly, helped convict Mark Redwine of murder this summer.

PHOTO: JERRY MCBRIDE/*THE DURANGO HERALD*

thing to make Dylan disappear. Redwine flatly denied any involvement, but police soon homed in on him.

On 5 August 2013, the investigation took a dramatic turn when Carren Corcoran arrived in Durango, Colorado, with her German shepherd, Molly. At the time, Corcoran worked for the Madison, Wisconsin, police department. In her off hours, she ran Canine Search Solutions, a private firm that specialized in using dogs trained to sniff out cadavers to search for missing people. By her count, she had worked 265 cases in 30-odd years. Molly, or "Bitty," as Corcoran called the dog, was unusually small for her breed. Molly had started out as "a timid little puppy that was given up to a rescue," Corcoran later wrote, but under her tutelage the dog became "a legend in the field of human remains detection."

Corcoran and Molly met local police at a La Plata County Sheriff's Office warehouse known as Area 51. Inside, Corcoran let Molly off leash. "Go find," she said. Nose down, Molly sniffed at the concrete floor. Corcoran understood the dog's behavior to mean the area was all clear—free of odors. Then she asked investigators to bring in the evidence, three paper packages containing clothes Mark Redwine had allegedly worn at the time of Dylan's disappearance. Corcoran again commanded Molly to search. This time, her dog went up to the bags, gave a deep, audible sniff, looked Corcoran in the eye, and sat down next to two of them. Because Corcoran had trained Molly to sit when the dog detected the odor of human remains, those "indications" suggested the odor of a dead body lingered on jeans, a pair of sneakers, and a work shirt.

Next the team went to Redwine's cabin, which was unoccupied. Molly sat down in a nearby stream and indicated at five other locations outside the home, according to Corcoran. Then she saw her dog looking up at a window. An investigator climbed into the house and opened the door for Corcoran and Molly. Inside, the dog made its indication—a sit—near a wood stove, in a bedroom, on a brown sofa, by a washing machine, and elsewhere. In February 2014, police called Corcoran back and had Molly sniff Redwine's Dodge pickup truck. Again, Corcoran said, her dog detected the odor of human remains, more than a year after the boy's disappearance.

> **"It's not enough to say I have this amazing expert. ... You have to explain exactly what their method is."**
>
> **Binyamin Blum**, University of California Hastings College of the Law

Nearly 4 years later, Redwine, driving a semitrailer, pulled into a parking lot in Bellingham, Washington. Police officers approached and arrested him on suspicion of killing his youngest son. In a body camera video of the encounter, released to a local TV station, Redwine looks disheveled and confused. "If you guys can, can you please explain to me what's going on?" State prosecutors in Colorado persuaded a grand jury to indict Redwine on some of the most serious charges the state could pursue: child abuse and second-degree murder. His trial began in July and the prosecution's narrative was built, in no small part, on commonly held assumptions about the behavior of a scent-detecting dog.

Dogs have been celebrated since antiquity for their ability to sniff a particular odor and lead humans to its source. But the domesticated canine's transformation into crime-fighting companion emerged much more recently, as U.S. police launched K-9 training programs and a thriving cottage industry of private firms, which often aid law enforcement, emerged. Today, police use dogs to track fugitives and search for missing persons, obtain probable cause (that is, legal justification to get a search warrant), and find substances, particularly illegal drugs. In what are known as scent lineups, agencies use trained canines to match evidence collected at a crime scene to the scent of a suspect or body. Increas-

ingly, testimony from dog handlers has also served as direct evidence of guilt—accepted in lieu of an actual corpse, drug stash, or other physical evidence of a crime.

Yet critics worry that the criminal legal system has embraced a technique profoundly lacking in scientific validation. Dog-sniff evidence has led to wrongful convictions, and studies show human biases skew animal behavior. Almost no published research indicates just what dogs detect or how they do it. Defendants and their lawyers can't cross-examine a dog, which means the accused cannot scrutinize the evidence or readily confront their accusers, a right enshrined in the U.S. Constitution.

"It's not enough to say I have this amazing expert with an incredible nose who can distinguish between scents," says Binyamin Blum, an evidence scholar at the University of California (UC) Hastings College of the Law, who contends such testimony short-circuits the safeguards in place to discriminate between junk science and real science. "You have to explain exactly what their method is."

**IN 1990**, about the time Corcoran joined the police force in Madison, she read a book, *So That Others May Live: Caroline Hebard & Her Search-and-Rescue Dogs*, describing one woman's passion for finding missing people. She was intrigued, but her bosses were not initially interested in deploying K-9s. So Corcoran began to train German shepherds on her own time. Between 1999 and 2001, she went to five seminars hosted by Canine Solutions Inc. (CSI) and taught by Sandra Anderson, a trainer with an uncanny knack for turning up evidence where others found nothing. (Sometimes no evidence existed: The FBI accused Anderson of planting bones and other remains for her dogs to find, and she went to prison in 2004.) In 2005, Corcoran founded Canine Search Solutions, attracting volunteers who shared a love for dogs, the outdoors, and in the words of one colleague, "solving mysteries and bringing closure to families."

Corcoran's biggest break came in 2004 when she cracked a cold case. A Wisconsin woman had gone missing in 1976. Nearly 4 decades later, Corcoran's dog Cleo indicated in 14 locations where the woman's ex-husband, the chief suspect, allegedly killed her and then tried to hide the body. No remains were found, and a judge barred Corcoran from testifying in the trial. Nonetheless, police said the suspect confessed and admitted to dumping the body at the locations Cleo flagged. Soon, other prosecutors were asking Corcoran to testify in trials. In court transcripts, she described dogs as "terrible poker players," incapable of deceit. As for Molly, she told one court last year, "I would put my house up against her indications."

Her career arc mirrors a broader shift in the United States. The number of police dogs swelled from 2000 in 1989 to as many as 50,000 in 2010. After the 2013 Supreme Court ruling in *Florida* v. *Harris*, which found drug-detection dogs reliable as long as the dog-and-handler team was properly trained and certified, judges increasingly allowed expert witnesses to testify about dog-sniff evidence.

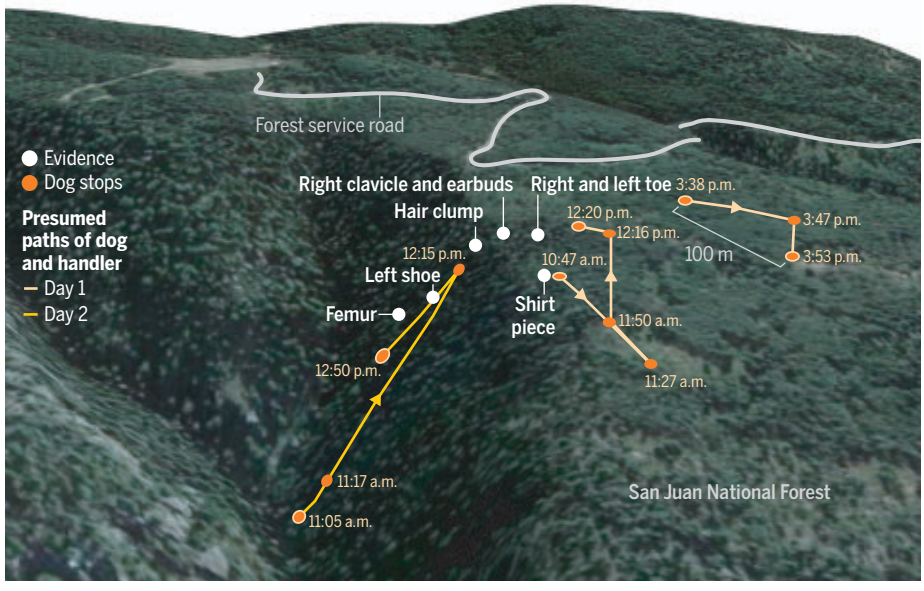
As with the 2004 cold case, the charges against Redwine hinged on Molly's ability to detect the odor of dead bodies where no human remains could be found. "The best way I can describe it is someone pops popcorn in your house," Corcoran explained to one Wisconsin jury. "You come in. You can smell the popcorn that's been popped, but there is no popcorn left." Suspects could hide a body but, as she saw it, no one could outrun a dog's nose.

The notion has plausible roots. Behind a dog's leathery, wet nose lies a cavernous labyrinth of scroll-shaped chambers called ethmoturbinates lined with some 200 million olfactory receptors, encoded by an estimated 2.5 times as many genes as in humans. In recent years, researchers studying canine cognition have shown pet dogs can sniff out minute quantities of odorants, such as the odor of their owner's T-shirt after it has been worn.

In the best known study on what is sometimes called residual odor—one that Corcoran has cited in court—German researchers placed carpet squares underneath two recently deceased corpses and then presented those squares to three trained Malinois. The 2007 study showed the dogs accurately distinguished those squares from negative controls, though critics contend the findings were muddied by possible contamination and poor experimental design.

## Follow the evidence?

Molly, a dog trained to find cadavers, stopped at multiple locations near Mark Redwine's cabin in Colorado, supposedly after detecting the odor of human remains. No remains were found at any of the locations, and the GPS coordinates do not appear to match where investigators previously collected physical evidence.




CREDITS: (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) LA PLATA COUNTY SHERIFF'S OFFICE; CARREN CORCORAN; *PEOPLE OF THE STATE OF COLORADO V. MARK REDWINE* (2017)

But what odors would be left 7 months or more after Dylan's alleged murder? Although forensic anthropologists have identified hundreds of compounds associated with cadavers, little evidence suggests all dead bodies immediately start to give off a standard, identifiable odor signature. John McGann, a neuroscientist at Rutgers University, New Brunswick, who wrote an influential review paper in *Science* arguing that poor human olfaction is a myth, says, "Even in cases where you say, 'OK, a dog can follow a scent trail through the woods and find a person.' Yes, they can. But what are they smelling? We don't know. Right? We just literally don't really know what exactly they're smelling."

**IN THE DECADE** leading up to Mark Redwine's trial, Corcoran had amassed an impeccable track record. Prosecutors had called her to testify in six criminal cases.

Each time judges allowed her testimony to be introduced before a jury, the suspects were found guilty, and unexplained deaths became criminally punishable homicides.

On 27 June 2019, almost 7 years after Dylan disappeared, Dana Delger and M. Chris Fabricant, attorneys from the Innocence Project's strategic litigation department, arrived at the La Plata County Court in Durango. Their goal was to challenge Corcoran's technique and persuade Judge Jeffrey Wilson not to admit the dog-sniff evidence.

The number of police dogs in the United States grew from 2000 in 1989 to as many as 50,000 in 2010.

Fabricant, hired by the Innocence Project in 2012, had assembled a portfolio of questionable forensic techniques. By the time of Mark Redwine's trial, dog-sniff evidence was high on that list. For one thing, dogs were prone to error. In data from Australia, police recorded 10,211 alerts made by drug-sniffing dogs; 74% of the hits turned out to be false alarms. But Fabricant was especially concerned about instances where canines provided direct evidence.

In one case, a Michigan man named D'Andre Lane told police his daughter had been kidnapped during a carjacking. But after an expert claimed his dogs detected the odor of decomposing remains in Lane's car, a jury found him guilty of murder. In another, an Arizona woman was sentenced to life in prison for murdering her daughter and throwing the body in the trash. Prosecutors called two handlers who claimed their dogs detected the scent of death in a dumpster. A body was never found in either case, but the testimony helped seal the suspect's fate.

Those cases suggested a larger pattern. A dog helped turn the loose threads in the prosecution's case into solid narrative rope. "The cadaver dog is bridging the gap between pretty plausible suspect and guilty beyond a reasonable doubt," Delger says, emphasizing that she could not comment on any particular case.

What happened in trials without any testimony about dogs was also telling. In late 2019, Fabricant and Delger helped a Wisconsin attorney representing a man accused of killing his ex-girlfriend and dumping her body in a lake. Corcoran and a colleague claimed dogs detected the odor of human remains inside the suspect's home, but the defense convinced a judge that the testimony did not meet the legal standard for scientific evidence. The suspect was found not guilty—ending Corcoran's winning streak.

No comprehensive database exists, but according to the National Registry of Exonerations, a project hosted by the University of Michigan Law School, at least 17 innocent people have been freed after dog-sniff evidence erroneously sent them to prison. Other convictions have been upheld—even after errors came to light. In 2000, Anderson, who trained Corcoran at CSI, helped persuade a jury in Madison to convict a Tanzanian immigrant accused of murdering his cousin. In court, Anderson paraded her Doberman pinscher in front of a jury, claiming the dog sniffed blood inside the suspect's home. Even though Anderson was later jailed for faking evidence, a judge upheld the life sentence, saying the discredited handler's testimony was "harmless." In a letter from prison, Peter Kupaza, the man convicted of the murder, told *Science* he "did not commit this crime [and] that is 100% fact."

**BY 2019**, when Fabricant and Delger submitted motions in Colorado, they had honed their argument. The Innocence Project attorneys did not question whether dogs could detect odors imperceptible to humans. Rather, they argued, "What is at issue is what has never been proven with any degree of scientific reliability: the ability of a dog to detect the residual scent of a particular object, including human remains, at a specific location days, weeks, months, or even more than a year after that object was removed." A sniff could be used to corroborate, but they argued a dog's indications alone should not be used to prove a person's guilt.

Fabricant and Delger argued that a dog's behavior may reflect a handler's expectations, pointing to a 2011 study in *Animal Cognition* by Lisa Lit, then at UC Davis. Lit found handlers "cued" dogs into making false indications. In one test, Lit showed up each morning with evidence bags containing cannabis and gunpowder, explaining to 18 teams that those target scents might be present inside a church. No target odors were present, and yet dogs positively indicated 85% of the time, handlers said, suggesting the dogs served as loyal companions first and objective scent detectors second.

Similar issues came to light more recently in letters criticizing a 2018 study published in *Forensic Science International*. The experiment was designed to test the long-standing belief that dogs could "mantrail" unfamiliar suspects by their odor. In the study, dogs correctly trailed people 82% of the time. But earlier this year, several critics pointed out methodo-

PHOTO: AFRICA STUDIO/SHUTTERSTOCK

logical shortcomings stemming from handler and experimenter bias, and the journal's editors added a cautionary "expression of concern."

Such biases may be impossible for dog handlers to avoid, Delger and Fabricant argued in their motion to exclude handler testimony in the Redwine case. A dog is "a companion animal with a vested interest in pleasing its master," they argued. Moreover, dogs are astute students of deliberate human signals and unintentional cues, paying close attention to people's faces, particularly our eyes.

Corcoran, who went by Carren Corcoran Gummin during the Redwine trial, stands by her work. In an interview with *Science*, she compared her critics to flat Earthers, saying she had "spent the last 20 years being criticized by people who do not understand how cadaver dogs work." She saw her work as one tool to extract information where there otherwise was nothing. "There's so many missing people that need to be found. You know, it's just sad. And that's really what it's all about: If we can recover somebody, so their family can know the ending. That's really what it should be about."

At the pretrial hearing in Durango, Delger, who took the lead in cross-examining Corcoran, asked whether dogs could reliably indicate in areas where no remains turned up. She asked the handler: "So there might be odor and you haven't recovered remains?"

"Correct," Corcoran said.

"But you might also have a false positive?"

"That's absolutely positive—or possible," Corcoran said.

Delger asked Corcoran whether Molly's indications on Redwine's property could be independently confirmed. "There's no test other than the dog that can tell us whether odors are there?"

"The testing that we conduct is simply: 'Is there odor there or not?'" Corcoran said. Molly indicated, she said, which put the odor of a dead body, at some point in time, in contact with the suspect's property.

Wilson allowed the testimony, citing an earlier ruling in Colorado that found that although dog tracking was not scientific, the criminal legal system accepted the technique as "common knowledge." The judge also noted that Corcoran would be a star prosecution witness.

**ON 6 JULY**, the 11th day of Redwine's trial, the prosecution called Corcoran to the stand. She described how Molly made more than 40 trained final indications, sitting near Redwine's clothes in Area 51, in and around the defendant's house, and in 12 places near where bones were found. That all suggested, as Corcoran told jurors, "She detected the odor of human remains."

Redwine's defense team raised the lack of a scientific basis for dog-sniff evidence and the possibility that dogs reflect explicit prejudices, implicit biases, or simply the very human desire to fill in gaps in our perception. But research suggests many jurors simply don't see it that way. In 2019, Lit teamed up with Itiel Dror, a cognitive neuroscientist from London who studies expert decision-making. They asked 554 mock jurors to read a story about a dog detecting the odor of drugs on a suspect's car even though a later search turned up cash but no evidence of illegal activity. Lit says the scenario was meant to sound preposterous. Nonetheless, two-thirds of jurors surveyed said they would have found the suspect guilty.

> **"The cadaver dog is bridging the gap between pretty plausible suspect and guilty beyond a reasonable doubt."**
>
> **Dana Delger**, Innocence Project

None of the three major accrediting organizations that certify cadaver dogs tests their ability to detect residual odor, and little published research shows dogs are capable of the feat. "I don't even know how you would even attempt to do that," says one of Corcoran's mentors, an official at an accrediting organization. Dror, by contrast, says a rigorous experiment would not be hard. "It's not rocket science. It's very simple." He suggests putting out cadavers for 1 week, removing the remains, and then asking handlers—who knew nothing about the experimental setup—to conduct a search.

Still, Dror says he understands why people put their faith in untested claims. People want justice. Humans seek cognitive closure. "You sleep better at night—the detectives and the juror do—if they find who did it and do not leave an unsolved crime," Dror told *Science*. "Here the dog comes and tells them."

**IN THE DAYS BEFORE** Corcoran's testimony, the evidence that Redwine murdered his son remained overwhelmingly circumstantial. The prosecution called Dylan's girlfriend, who tearfully remembered him saying he did not want to visit his father. Dylan's older brother hinted at a possible motive, saying Dylan was going to confront his father about photos allegedly showing Redwine eating what appeared to be feces out of a diaper. (As a *New York Post* headline put it, "'Compromising' fetish pics led Colorado dad to kill teen son, prosecutors claim.") The boy's skull contained apparent fractures, and though the possibility that cattle or scavenging wildlife damaged the bones after death was hard to rule out, the prosecution called on an expert who said a sharp object was responsible. No weapon was ever found.

The only direct evidence putting a dead body inside the suspect's home and his pickup came from Corcoran's testimony. She had trouble remembering details from a search 8 years earlier, and she broke down in tears when prosecutors displayed a photo of Molly, who died in 2018, on a screen behind the stand. None of the GPS coordinates Corcoran recorded on the mountain matched the location of any skeletal remains (see graphic, p. 253). But that did not necessarily prove Molly wrong, she said, because odor could potentially travel down the mountain in streams and rivers.

As the trial wound down, the defense called another expert, James Ha, a researcher at the University of Washington, Seattle, who studies applied animal behavior but—as prosecutors got him to admit—has little firsthand experience with cadaver dogs. Ha questioned whether odors could linger for more than 48 hours because they are volatile and likely to dissipate. Given the lack of published research and the difficulty in ruling out false positives, he said, scientists generally did not accept residual odor as a valid concept. Ha argued that without tangible evidence to validate an indication, a dog alone could not yield any information. "We don't know how to get inside a dog's mind," Ha told the jury. "They're not machines."

Redwine chose not to testify. Six months after his son's disappearance, he wrote in a statement to local media, "I absolutely do not know where Dylan is." His attorneys tried to paint him as a loving father who must have felt guilty for losing his son on his watch. In a case built on circumstantial evidence, the trial pitted his claims of innocence against a dog's telling indications. Twelve people were asked to decide whose story to believe.

In late July, after more than 6 hours of deliberation, the jury found Redwine guilty—an outcome that cemented Corcoran's reputation for convincing jurors that dogs just know what the rest of us could not. ■

This story was supported by a grant from the Massachusetts Institute of Technology Knight Science Journalism fellowship.

# INSIGHTS

PERSPECTIVES

ECOLOGY

# The benefits of disturbance

## Physical disruption of an ecosystem by sea otters affects the genetic diversity of their habitat

*By* **Joe Roman**

Of the many legacies of the Atomic Age—the Manhattan Project, the bombings of Hiroshima and Nagasaki, and the nuclear arms race—one of the brighter outcomes is the return of sea otters to Southeast Alaska and British Columbia, a project that was funded by the U.S. Atomic Energy Commission. The otters' return over the past 50 years is a conservation success story, and in part because of the well-documented history of their reintroduction, they have become a flagship species for conservation biology, ecology, and ecological economics. On page 333 of this issue, Foster *et al.* (*1*) suggest that otter restoration can enhance the genetic diversity of their habitat. The disturbance of seagrass meadows by foraging otters could increase the adaptive potential of these grasses, making meadows more resilient and resistant to perturbations.

How did the Atomic Energy Commission get involved with sea otters? Before the 1971 underground detonation of the *Cannikin* warhead on Amchitka Island, Alaska—the largest nuclear test to date—there were concerns that the explosion would affect the many otters around the island. The commission funded the relocation of several hundred sea otters from the waters around the test site to Southeast Alaska and British Columbia, where the otters had long been absent. (Many of the otters that were left behind died.)

As a graduate student, James Estes worked on Amchitka and the surrounding islands before and after the blast. Early in his career, he predicted that the sea otter's influence on coastal ecosystems would prove to be "strong and far reaching" (*2*). Otters had been extirpated from much of their range during intense commercial harvests in the 18th and 19th centuries. As their range expanded in the 1960s and 1970s after the relocation efforts, and their habitats became better understood, sea otters would become iconic species in ecology; one of the most familiar trophic cascades is the relationship between sea otters, sea urchins, and kelp forests. Estes and his colleagues would later show that this relationship—otters eat urchins, urchins decline, kelp increases—has widespread impacts, influencing the diet of bald eagles and glaucous-winged gulls, fish abundance, and even the carbon cycle (*3*, *4*).

Many studies focus on kelp forests, but sea otters reside in other coastal habitats, including seagrass meadows, where they feed on rock crabs and clams, often creating forag-

Gund Institute for Environment, University of Vermont, Burlington, VT 05477, USA. Email: romanjoe@gmail.com

PHOTO: SUZI ESZTERHAS/MINDEN PICTURES

Sea otters dig for clams in eelgrass meadows. The disturbance from these foraging pits promotes genetic diversity and perhaps ecological resilience.

ing pits in the eelgrass. Foster *et al.* looked at the impact of this activity on eelgrass genetic diversity by comparing seagrass ecosystems with no otters (pre-disturbance) to those where otters had been reintroduced (post-disturbance). In some areas, otters had been established for decades; in others, they had arrived only recently. The authors found that allelic richness (the number of alleles) of eelgrass was highest in areas where otters had resided for more than 20 years. Foster *et al.* argue that allelic richness is of central importance here because it is integral to resilience and the evolution of clonal organisms. Genotypic diversity (the proportion of different genotypes) was also higher in these established areas. Why? The primary mode of expansion for seagrasses is clonal growth, but sea otter digging stimulates sexual reproduction in the flowering plant and opens up patches for the seeds to settle.

There were no differences in heterozygosity, which Foster *et al.* say is of limited value in clonal species. Other geneticists might disagree: heterozygosity (having two different alleles at a locus) is often considered to be one of the gold standards in measuring genetic diversity. Future work will be helpful here, to examine pre- and post-perturbation patterns of sea otters and seagrasses over time, reveal the ecological and genetic mechanisms at play, and look for these patterns in other ecosystems. Even though disturbance underpins the diversity and dynamics of many of the world's ecosystems, its influence on genetic diversity remains poorly appreciated (*5*).

The amount of disturbance that was observed in the seagrass meadows was relatively low: About 5% of the seafloor had foraging pits, although the pits are ephemeral, and visual observations likely underestimate the amount of activity. The intermediate disturbance hypothesis, presented in 1978 on the basis of evidence from tropical forests and coral reefs, is well known to ecologists, although it remains debated (*6*). Early studies showed that intermediate disturbance could increase species diversity (*7*), but research on the effects of disturbance on genetic diversity are relatively rare. The disturbance of the bacterium *Pseudomonas fluorescens* in a laboratory setting caused competition among genotypes, maintaining higher genetic diversity at intermediate levels of disturbance (*8*). It was recently reported that moderate fire disturbance increases genetic diversity in butterflies, which could have implications for fire management in the western United States (*9*).

The role of sea otters in eelgrass meadows provides insights into how genetic diversity can change when species interactions are lost—and how these patterns can be restored. Without otters, genetic diversity appears to decline; when they return, so does diversity. In addition to enhancing diversity, these top predators also provide other ecological benefits. Elkhorn Slough in central California is subject to high nutrient loading from agriculture, which can degrade eelgrass beds. Sea otters feed on predatory rock crabs in this area, and as the crabs decline, sea slugs and other grazers thrive, reducing algae and promoting the growth and expansion of eelgrass (*10*).

The findings of Foster *et al.* come at an important time. The need to protect and monitor genetic diversity has often been overlooked in global commitments such as the Convention on Biological Diversity (*11*). With the Fifteenth Meeting of the Conference of the Parties to the Convention on Biological Diversity, or COP 15, gathering this year, policy-makers should commit to actions that will avert the loss of genetic diversity and advance monitoring, going well beyond agricultural species and their relatives (*12*). Research on the effects of disturbance and species interactions on genetic diversity could help inform the restoration efforts of trophic rewilding (*13*).

The role of sea otters and other animals in biogeochemical processes such as the carbon cycle has gained attention in recent years (*14*). Animals can also enhance the productivity and resilience of coastal ecosystems (*15*) and might be crucial allies in protecting coastlines and storing and sequestering carbon. During the United Nations Decade on Ecosystem Restoration (2021–2030), more attention should be paid to rebuilding animal populations and the key ecological functions they provide—with all the urgency and resources of a war effort, but without the nuclear bombs. ■

## REFERENCES AND NOTES


1. E. Foster *et al.*, *Science* **374**, 333 (2021).
2. J. A. Estes, *Serendipity: An Ecologist's Quest to Understand Nature* (Univ. California Press, 2020).
3. J. A. Estes *et al.*, *Science* **333**, 301 (2011).
4. C. C. Wilmers, J. A. Estes, M. Edwards, K. L. Laidre, B. Konar, *Front. Ecol. Environ.* **10**, 409 (2012).
5. S. C. Banks *et al.*, *Trends Ecol. Evol.* **28**, 670 (2013).
6. J. W. Fox, *Trends Ecol. Evol.* **28**, 86 (2013).
7. J. H. Connell, *Science* **199**, 1302 (1978).
8. A. Buckling, R. Kassen, G. Bell, P. B. Rainey, *Nature* **408**, 961 (2000).
9. D. Gates, B. Jackson, S. D. Schoville, *J. Hered.* **112**, 367 (2021).
10. B. B. Hughes *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 15313 (2013).
11. S. Hoban *et al.*, *Bioscience* **71**, 964 (2021).
12. S. Hoban *et al.*, *Biol. Conserv.* **248**, 108654 (2020).
13. E. S. Bakker, J.-C. Svenning, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **373**, 20170432 (2018).
14. O. J. Schmitz *et al.*, *Science* **362**, eaar3213 (2018).
15. N. A. J. Graham *et al.*, *Nature* **559**, 250 (2018).


10.1126/science.abm2257

MATERIALS SCIENCE

# No more trial and error for zeolites

## A data-intensive approach scours known organics for synthesizing targeted zeolites

*By* **Watcharop Chaikittisilp**[1] *and* **Tatsuya Okubo**[2]

Zeolites, a class of crystalline materials with microscopic pores, have industrial uses ranging from adsorbents and ion-exchangers for gas separations and water treatments to petroleum, petrochemical, fine-chemical, and automotive catalysts (*1*). Zeolites can form naturally or artificially and have more than 250 different structures identified thus far according to the Structure Commission of the International Zeolite Association (*2*), with millions more theoretically possible structures yet to be found (*3*). A priori design of zeolite synthesis is a long-standing challenge because it is a trial-and-error process that lacks universal design principles. On page 308 of this issue, Schwalbe-Koda *et al.* (*4*) report a data-intensive approach to selecting organic structure-directing agents (OSDAs) that can control phase selectivity in zeolite synthesis. Their protocol combines literature mining, high-throughput simulations, and human–computer interaction to examine known OSDAs for synthesizing targeted zeolites and modulating the atomic distributions of zeolites (see the figure).

All known zeolites are crystallized as metastable phases. Metastable materials are less energetically favorable than stable ones; therefore, to synthesize them, this energy penalty must be taken into account. To compensate for this penalty, OSDAs and/or hydrated inorganic cations (sometimes called inorganic structure-directing agents) are used to provide the extra energy needed to stabilize the metastable zeolite phases.

The metastability of zeolites is dependent on framework structures and chemi-

[1]Research and Services Division of Materials Data and Integrated System, National Institute for Materials Science, Ibaraki 305-0044, Japan. [2]Department of Chemical System Engineering, The University of Tokyo, Tokyo 113-8656, Japan. Email: chaikittisilp.watcharop@nims.go.jp; okubo@chemsys.t.u-tokyo.ac.jp

cal compositions (*5*) and is usually expressed as the relative framework energy, which is the energy difference with respect to the closest stable phase, for example, the α-quartz phase in the case of pure silica. This relative framework energy is often used alongside the stabilization energy, which is calculated by subtracting the energies of empty zeolite and OSDA molecules from the energy of OSDA–zeolite complexes, for evaluating the structure-directing ability of OSDAs.

To find the best OSDA for the synthesis of a specific zeolite, the size and shape of the OSDA must be commensurate with the porous cavities of the zeolite, so that the OSDA can be grown de novo within the interior spaces of the zeolite until the cavities are filled (*6*). Although this method can in theory produce the best OSDAs, some of them may be difficult or impossible in practice. An evolutionary algorithm that uses readily available chemicals and known reactions can be used to design synthesizable OSDAs (*7*). In addition to identifying OSDAs from an energetic standpoint, this approach identifies possible routes and reactants for the synthesis of OSDAs themselves, which provides a more comprehensive assessment of the practicality of the OSDAs. Furthermore, for zeolites to be commercially viable, the cost of OSDAs must also be integrated into the design process, for example, by using a multiobjective optimization algorithm (*8*).

Another limitation of the existing design methodology lies in the lack of transferability. By only trying to find the best OSDA for a specific zeolite, the approach may miss the fact that the same OSDA can also be the best OSDA for other zeolites. Schwalbe-Koda *et al*. propose an evaluating metric they call the templating energy, which includes both the structure directivity and competitivity for each given OSDA–zeolite pair. The directivity of an OSDA evaluates how close an OSDA is to being the best OSDA for a given zeolite from an energetic standpoint, whereas the competitivity of a zeolite represents how close a zeolite is to being the best zeolite for a given OSDA. As a result, the templating energy can simultaneously rank the best OSDA for a given zeolite and the best zeolite for a given OSDA.

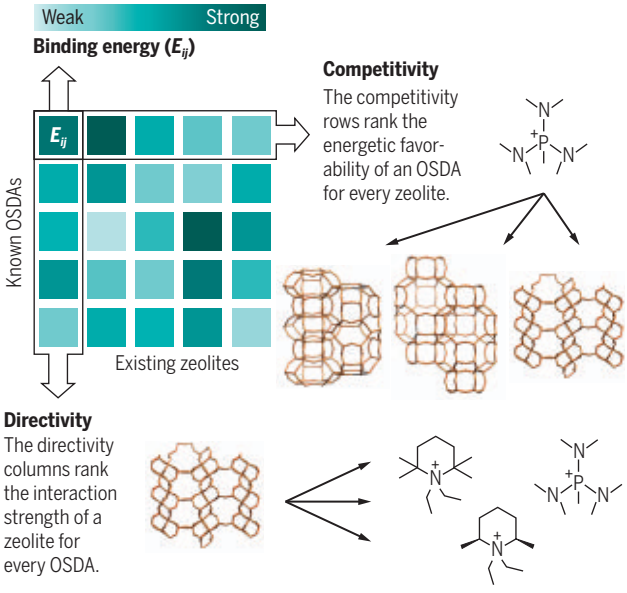
Schwalbe-Koda *et al*. demonstrate the applicability of this approach in designing the OSDAs for synthesizing zeolites with different structures, including the AEI structure, the CHA structure, and ones with a CHA/AEI intergrown structure, as well as modulating the distribution of aluminum in CHA zeolites. They collected data on the OSDAs previously used for zeolite synthesis from the literature using automated data extraction tools (*9*) and then calculated the binding energies of each OSDA for the stabilization of most of the known zeolite frameworks. Besides energetic scores, the size and shape of each OSDA are also considered as part of the assessment process. They used a principal components analysis to reduce the dimensionality of the OSDA down to two-dimensional space (*10*) and used a voxel-based method to estimate the OSDA volume. In addition to the templating energy and the geometric parameters, the authors also include the score describing the synthetic complexity (*11*) in their metrics for identifying the OSDAs suitable for synthesizing specific zeolites. A database of OSDA–zeolite pairs using the aforementioned evaluating scores was constructed and made publicly accessible by the authors, which will allow researchers to use it as a guide when selecting OSDAs.

Although the scope of the study is limited to known OSDAs and existing zeolites, it provides a systematic design protocol for zeolite synthesis and may be extended to consider hypothetical zeolites, although the computational cost will likely be higher. As is well known, small changes in synthesis conditions of zeolites can lead to huge differences in the crystalline phases of the products (*12*). The current study considered a comprehensive roster of OSDAs but was limited to previously reported synthesis conditions such as the temperature and the chemical compositions of starting reagents. It is of great importance to apply this computational workflow to unknown zeolites and develop a systematic approach for a priori design for all synthesis conditions, but this might prove to be very challenging. There may exist some correlations between structures and synthesis conditions of zeolites (*13*, *14*), and the ideal synthesis conditions for an unknown zeolite may be found by comparing to an existing zeolite through structure–synthesis similarity mapping (*15*).

Researchers have toiled for decades attempting to identify ways to synthesize zeolites on the basis of intuition and experience, combined with mechanistic studies of crystallization pathways and various techniques in computational chemistry. The data-driven approach proposed by Schwalbe-Koda *et al*. provides a path to formalize and formulate this intuition and experience. A workflow that integrates computations and experiments, in conjunction with data science, could realize the truly designed synthesis of zeolites. ■

## A pairing guide for OSDAs and zeolites

Binding energy, also known as stabilization energy, is a useful metric for evaluating the structure-directing ability of organic structure-directing agents (OSDAs) toward targeted zeolites. By using the binding energies for every known OSDA–zeolite pair, Schwalbe-Koda *et al*. quantify the directivity and competitivity energies for assessing the phase selectivity of zeolites.




GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

### REFERENCES AND NOTES


1. M. E. Davis, *Nature* **417**, 813 (2002).
2. C. Baerlocher, L. B. McCusker, Database of Zeolite Structures; www.iza-structure.org/databases.
3. Y. Li, J. Yu, *Chem. Rev.* **114**, 7268 (2014).
4. D. Schwalbe-Koda *et al*., *Science* **374**, 308 (2021).
5. K. Muraoka, W. Chaikittisilp, T. Okubo, *J. Am. Chem. Soc.* **138**, 6184 (2016).
6. D. W. Lewis, D. J. Willock, C. R. A. Catlow, J. M. Thomas, G. J. Hutchings, *Nature* **382**, 604 (1996).
7. R. Pophale, F. Daeyaert, M. W. Deem, *J. Mater. Chem. A* **1**, 6750 (2013).
8. K. Muraoka, W. Chaikittisilp, T. Okubo, *Chem. Sci.* **11**, 8214 (2020).
9. Z. Jensen *et al*., *ACS Cent. Sci.* **5**, 892 (2019).
10. Z. Jensen *et al*., *ACS Cent. Sci.* **7**, 858 (2021).
11. C. W. Coley, L. Rogers, W. H. Green, K. F. Jensen, *J. Chem. Inf. Model.* **58**, 252 (2018).
12. M. D. Oleksiak, J. D. Rimer, *Rev. Chem. Eng.* **30**, 1 (2014).
13. K. Itabashi, Y. Kamimura, K. Iyoki, A. Shimojima, T. Okubo, *J. Am. Chem. Soc.* **134**, 11542 (2012).
14. P. Guo *et al*., *Nature* **524**, 74 (2015).
15. K. Muraoka, Y. Sada, D. Miyazaki, W. Chaikittisilp, T. Okubo, *Nat. Commun.* **10**, 4459 (2019).


### ACKNOWLEDGMENTS

W.C. acknowledges the Japan Society for the Promotion of Science for their financial support through a Grant-in-Aid for Scientific Research (B) (KAKENHI: 20H02528).

10.1126/science.abm0089

ATMOSPHERIC SCIENCE

# Gulf Stream and Kuroshio synchronization

## The mid-latitude jet stream ties together the two major western boundary currents

*By* **Paola Cessi**

Extratropical cyclones (and anticyclones) in the Northern Hemisphere are preferentially located along storm tracks with centers of action at the exit regions of the major western boundary ocean currents. These currents are the Gulf Stream and the Kuroshio in the North Atlantic and North Pacific, respectively. As these currents strengthen or weaken, or migrate in latitude, the strength and location of storm tracks is affected. On page 341 of this issue, Kohyama *et al.* (*1*) use satellite observations, high-resolution coupled atmosphere-ocean models, and a simple conceptual model to demonstrate that the Gulf Stream and Kuroshio are synchronized through the atmospheric Atlantic and Pacific storm tracks.

These currents bring warm subtropical water northward, forming strong thermal fronts where they collide with southward-flowing boundary currents from the north: the Labrador Current for the North Atlantic and the Oyashio for the North Pacific. These thermal fronts extend eastward into the oceans as boundary currents leave the continental coast, providing fuel for the formation of mid-latitude atmospheric weather disturbances (see the figure) (*2*). The storms develop downstream and eventually decay as they are carried away from the thermal front by the prevailing westerlies. In turn, the surface westerlies are accelerated and sharpened by the storms (*3*).

The westerlies are also responsible for driving the oceanic western boundary currents that generate the oceanic thermal fronts in mid-latitudes. In this way, the ocean-atmosphere interaction between the western boundary currents, the atmospheric storms, and the associated surface westerlies form a self-sustaining system (*4*). Because western boundary currents take several years to adjust to changes in the surface westerlies, the coupled system oscillates irregularly with multiyear periodicity, dictated by the oceanic adjustment time (*5*).

**"The Atlantic and Pacific storm tracks are coupled aloft across the North American continent by the westerly jet..."**

Kohyama *et al.* demonstrate that the Gulf Stream and Kuroshio ocean currents are synchronized through storm tracks, which are anchored along the western boundary current frontal zones. The Atlantic and Pacific storm tracks are coupled aloft across the North American continent by the westerly jet, through mid-latitude teleconnections such as the Northern Hemisphere annular mode (*6*, *7*).

The Gulf Stream and Kuroshio exhibit different adjustment times to changes in the winds, about 3 and 5 years, respectively. This difference arises because the North Pacific is wider than the North Atlantic: Waves must travel the whole longitudinal span of each ocean in order for the western boundary current to adjust to changing winds. However, atmospheric coupling can synchronize their oscillations (*5*). The synchronization is analogous to that observed by Huygens for two pendulum wall clocks attached to a common support (*8*).

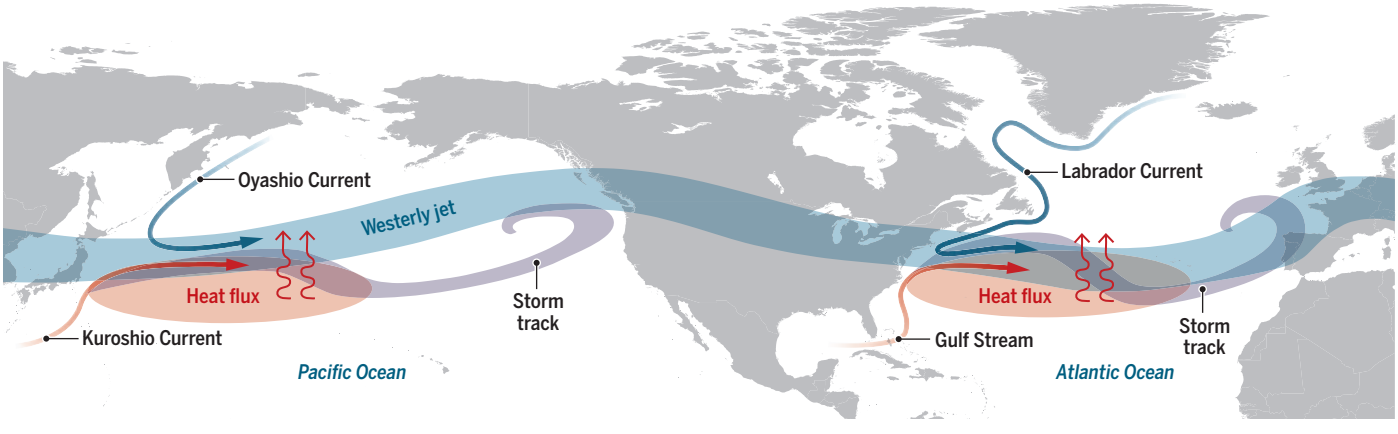
Because the atmosphere-ocean system is more complex than two coupled clocks, with motions occurring on many time scales, the synchronization of the Gulf Stream and Kuroshio emerges after filtering out the fast time scales associated with individual weather events. The interannual to decadal synchronization is characterized by meridional shifts of the boundary current fronts associated with sea surface temperature anomalies centered at the exit region of the Gulf Stream and Kuroshio.

When the sea surface temperature at the Gulf Stream and Kuroshio fronts increases (decreases), the mid-latitude jet migrates northward (southward). The effect of the

Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093, USA.
Email: pcessi@ucsd.edu

### Synchronization of ocean boundary currents

Warm (thick red arrows) and cold (thick blue arrows) western boundary currents collide to form thermal fronts with large upward heat fluxes (light red ellipses and red wavy arrows). The thermal fronts anchor the storm tracks in the lower atmosphere (light purple curly ribbons) and sharpen the westerly jet (light blue wavy ribbon) in the upper troposphere, coupling the two storm tracks. In turn, the westerly jet drives the oceanic western boundary currents.




GRAPHIC: K. FRANKLIN/*SCIENCE*

western boundary current on the overlying atmospheric track is only captured by models that resolve the sharp oceanic temperature front with length scales smaller than 100 km (*9*, *10*). Only in this case can the oceanic mid-latitude warming affect the atmosphere throughout its depth and effectively anchor the storm tracks and the westerly jet to the western boundary current thermal fronts. Thus, long-time, large-scale synchronization of the Gulf Stream and Kuroshio is only captured by coupled models that resolve fast-time, small-scale processes.

The interannual to decadal synchronization of the Gulf Stream and Kuroshio has important implications because the mid-latitude storm-track and jet activities modulate extreme temperature and precipitation over land (*11*). Low storm-track activity in summer leads to persistent, stationary anticyclones and heat waves. For example, the summer of 2018 was associated with large positive sea surface temperature anomalies just south of the Gulf Stream front, shifting the North Atlantic jet far to the north and leaving northern Europe, the United Kingdom, and Scandinavia under dry, hot conditions for months (*12*). The widespread concurrent heat anomaly in the North Pacific sector, leading to heat waves in Japan, South Korea, eastern China, and the western United States, was associated with a large positive sea surface temperature anomaly at the Kuroshio front (*13*).

The analysis of Kohyama *et al*. indicates that there is no substantial lag between the oceanic temperature anomalies at the Kuroshio and Gulf Stream fronts. Nevertheless, the synchronization of the western boundary current fronts, and associated storm-track modulation, enhances our ability to predict increases in the frequency of extreme events on interannual to decadal time scales. ■


### REFERENCES AND NOTES

1. T. Kohyama *et al*., *Science* **374**, 341 (2021).
2. K. A. Kelly *et al*., *J. Clim.* **23**, 5644 (2010).
3. H. Nakamura, T. Sampe, Y. Tanimoto, A. Shimpo, *Geophys. Monogr.* **147**, 329 (2004).
4. B. J. Hoskins, P. J. Valdes, *J. Atmos. Sci.* **47**, 1854 (1990).
5. B. Gallego, P. Cessi, *J. Clim.* **14**, 2815 (2001).
6. D. W. J. Thompson, J. M. Wallace, *J. Clim.* **13**, 1000 (2000).
7. B. J. Hoskins, K. I. Hodges, *J. Clim.* **32**, 1743 (2019).
8. C. Huygens, *Phil. Trans.* **4**, 937 (1669).
9. S. Minobe, A. Kuwano-Yoshida, N. Komori, S.-P. Xie, R. J. Small, *Nature* **452**, 206 (2008).
10. H. Nakamura, T. Sampe, A. Goto, W. Ohfuchi, S.-P. Xie, *Geophys. Res. Lett.* **35**, L15709 (2008).
11. J. Lehmann, D. Coumou, *Sci. Rep.* **5**, 17491 (2015).
12. M. McCarthy *et al*., *Weather* **74**, 390 (2019).
13. C.-C. Hong *et al*., *J. Clim.* **34**, 6319 (2021).



### ACKNOWLEDGMENTS

Funding was provided by the NSF (OCE-1634128) and by NASA (80NSSC20K0796).

10.1126/science.abl9133


DECISION SCIENCE

# Physicians' flawed heuristics in the delivery room

## Delivery complications can influence later decision-making


*By* **Meng Li**[1] *and* **Helen Colby**[2]


Clinical decisions require physicians to integrate a complex and changing set of stimuli to determine the best course of action for the patient, often under time pressure. Although physicians are highly trained experts, they are also human, and decades of research have shown that we frequently rely on heuristics or decision rules when making complex decisions. Although efficient and sometimes adaptive given our limited cognitive capacity, heuristics can also lead to suboptimal decisions (*1*). This understanding of bounded rationality has led to effective interventions in government policy (*2*), health behavior (*3*), environmental choice (*4*), and other domains but has not yet been widely accepted by medical professionals (*5*). On page 324 of this issue, Singh (*6*) demonstrates that physicians delivering babies appear to rely on a suboptimal heuristic, tending to switch to the alternative delivery modality (vaginal versus cesarean) for their next patient if their immediately prior delivery had complications, and this leads to negative patient outcomes.

The relative lack of acknowledgment of potential decision biases in clinical practice may be partially because of the difficulty studying, and thus limited evidence of, physicians' cognitive biases in real-world clinical settings. Indeed, physician decision research has mostly relied on hypothetical scenarios (*7*). Using experiments in the clinical setting would cause moral concerns because subjecting physicians to experimental conditions that may elicit suboptimal decisions would mean harming real patients. Using nonexperimental design means relying on analysis of existing medical records, and it is often difficult to definitely identify what the rational medical decision should be for a given patient on the basis of their records alone. Singh overcomes these challenges by focusing on the switching pattern in delivery modality decisions, using more than 86,000 patient health records, from both an urban and a suburban academic hospital, spanning January 2000 to September 2020.

Delivery decision-making is an especially astute context for studying clinical decision biases not only because it has a serious impact on both a mother and a newborn but also because it provides a circumscribed set of decision options—the delivery can only occur by a vaginal birth or a cesarean section. Singh finds that if a physician performed a vaginal delivery that had complications, the physician is more likely to perform the next delivery by cesarean than if the prior vaginal delivery had no complications—and vice versa for when it was an earlier cesarean that had complications. This switching behavior is specific to complications in the patient immediately before the current delivery, and is more pronounced with greater numbers of, and more severe, complications in the prior patient.

Delivery decisions made by physicians are complex and must take into account a wide variety of data, as well as the available resources and the preferences of the patient. In some instances, there may be a clear indication for vaginal or cesarean delivery. However, there are many less–clear-cut situations in which there is considerable physician leeway (*8*). For example, it is unclear how to optimally balance patient preferences with the speed of labor progression to determine whether to choose cesarean delivery for a patient whose vaginal delivery is progressing slowly. Although there is not a discoverable "correct" method of delivery for each patient, and perhaps exactly because of such ambiguity, delivery mode decisions can be influenced by factors that should not be part of the rational calculation; in this case, Singh found that it is influenced by what happened to the physician's immediately prior patient.

Singh's results are of more than just theoretical interest. Delivery modality switching leads to concrete, albeit small, negative patient outcomes for both mother and newborn. Switching after complications leads to an increased probability of death for the mother or baby (by 0.04 SD), a decreased probability of the patient being discharged home (by 0.03 SD), and an increased num-


[1]Department of Health and Behavioral Sciences, University of Colorado Denver, Denver, CO, USA. [2]Kelley School of Business, Indiana University, Indianapolis, IN, USA. Email: meng.li@ucdenver.edu

ber of outpatient visits after delivery (by 0.04 SD within 1 month after delivery) compared with switching after no complications. Singh's results are also impressively robust, showing that the switching effect stands after accounting for physician fatigue, patient volume, similarity between patients, death in prior delivery, time between the two deliveries, and scheduled cesareans. The effect also stands across 2560 robustness checks with different statistical specifications of the model.

This "win-stay-lose-shift" heuristic has been identified in other contexts as a learning strategy, but it only performs well in certain situations (*9*, *10*). In the medical context, this heuristic would be rational if the specifics of the prior patient matched the specifics of the current patient and thus provided a useful learning experience. In that case, if one patient's delivery went wrong, it can tell the physician that the same delivery plan may not work well with another patient with very similar characteristics and indications. However, two patients who happen to have consecutive deliveries by the same physician are not expected to be so highly similar, and, thus, using the outcome from the immediately prior patient to determine delivery modality for the next patient is a misapplication of an otherwise intuitive and perhaps useful heuristic.

Physicians' use of the "win-stay-lose-shift" heuristic does not mean that they lack medical expertise or training. Indeed, Singh finds that more experienced physicians use this decision rule more often, even though physicians who rely more on this rule have worse patient outcomes over time. The result demonstrates that the use of heuristics or simplified decision rules is a common human tendency even among smart, well-intentioned, highly trained doctors.

Three well-established cognitive biases stand out as potential causes of this maladaptive switching of delivery modes, whether individually or in concert: recency, affect heuristic, and confirmation bias. The most recent events are the most prominent in memory (*11*), thus complications that occur in the most recent delivery may exert an outsized influence on a physician's choice with the next patient. The finding that the switching behavior is only present based on the outcome of the immediately prior delivery—but not deliveries two, three, or four patients prior—supports this explanation. In addition, people rely on their affective response to gauge how large a risk is (*12*), and delivery complications can be highly emotional and thus distort a physician's estimate of such risk, prompting them to switch to a different delivery mode to avoid the increased perceived risk. A third potential mechanism is confirmation bias, or the tendency to seek out information that provides support for one's current opinion or position (*13*). Physicians who are uncomfortable with a modality because they recently had a negative experience with it may inadvertently seek and interpret evidence in a way that is consistent with such fears, which helps them feel comfortable that modality switching is appropriate for the current patient.

How can we help physicians overcome reliance on maladaptive heuristics or decision rules? The first step is to acknowledge the phenomenon without condemning physicians. Although understanding decision biases usually does not entirely ameliorate them (*14*), teaching doctors about heuristics may promote the acceptance of potential interventions. More research and clinical efforts need to then focus on designing and testing decision aids that are beneficial to patients and user-friendly to physicians in the labor and delivery context. In addition to making sure that the decision aid has a high degree of scientific accuracy in recommending the optimal treatment option, studies also need to examine whether physicians will accept and use such recommendations, because physicians may have understandable concerns about recommendations from a "black box" (*5*). It is also useful to explore non-algorithm-based decision aids such as decision trees or check lists, which are more transparent (*15*), although it remains to be seen if such simple decision aids are appliable in the complex labor and delivery environment. In addition, there is a slightly larger delivery-mode shifting effect for consecutive deliveries within 20 hours, and the effect evaporates for deliveries more than 72 hours apart. This suggests that the influence of a delivery with complications fades with time. Further research should explore interventions to help physicians clear their minds between deliveries in a way that is similar to the natural effect of time.

It is time to acknowledge the prevalence of heuristics and decision biases in clinical practice and to view these patterns as predictably human instead of blaming individual doctors. Only then can we start helping doctors improve clinical decisions and, as a result, the health of the public. ■

Complications during this cesarean delivery would make it more likely that the physician would pursue vaginal delivery with the next patient.

PHOTO: BSIP/EDUCATION IMAGES/UNIVERSAL IMAGES GROUP/GETTY IMAGES


## REFERENCES AND NOTES

1. A. Tversky, D. Kahneman, *Science* **185**, 1124 (1974).
2. D. Halpern, M. Sanders, *Behav. Sci. Policy* **2**, 52 (2016).
3. M. Li, G. B. Chapman, *Soc. Personal. Psychol. Compass* **7**, 187 (2013).
4. N. Dean, *Nat. Energy* **3**, 616 (2018).
5. P. Croskerry, *N. Engl. J. Med.* **368**, 2445 (2013).
6. M. Singh, *Science* **374**, 324 (2021).
7. J. S. Blumenthal-Barby, H. Krieger, *Med. Decis. Making* **35**, 539 (2015).
8. V. Berghella, A. D. Mackeen, E. R. Jauniaux, in *Gabbe's Obstetrics: Normal and Problem Pregnancies*, M. B. Landon *et al.*, Eds. (Elsevier, 2020).
9. M. Nowak, K. Sigmund, *Nature* **364**, 56 (1993).
10. F. A. Matsen, M. A. Nowak, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 18053 (2004).
11. A. D. Baddeley, G. Hitch, *Mem. Cognit.* **21**, 146 (1993).
12. P. Slovic, M. L. Finucane, E. Peters, D. G. MacGregor, *Eur. J. Oper. Res.* **177**, 1333 (2007).
13. R. S. Nickerson, *Rev. Gen. Psychol.* **2**, 175 (1998).
14. B. W. Smith, M. B. Slack, *Diagnosis* **2**, 117 (2015).
15. O. Wegwarth, W. Gaissmaier, G. Gigerenzer, *Med. Educ.* **43**, 721 (2009).



## ACKNOWLEDGMENTS

We thank G. Chapman for helpful comments and suggestions.

10.1126/science.abl5647

CELL BIOLOGY

# Resealing and rebuilding injured muscle

## Reconstructing an injured muscle involves nuclear and mitochondrial contributions

By **Elizabeth M. McNally**[1] *and* **Alexis R. Demonbreun**[1,2]

Skeletal muscle is organized into elongated and aligned myofibers that can rapidly shorten to elicit "on-demand" movement. Myofibers are individual cells, enclosed in a single plasma membrane with many nuclei per fiber. Within mouse skeletal muscle, there are between 5 and 10 nuclei for every 100 micrometers of myofiber length (*1*). These nuclei localize at the periphery of myofibers, and the myofiber cytoplasm is dominated by the near-crystalline array of contractile units (sarcomeres). Repetitive or excessive muscle contraction can disrupt, or injure, the plasma membrane of myofibers, triggering cellular repair processes. The large and multinucleated nature of myofibers presents distinct challenges for repair machinery, yet the large size of myofibers also enables real-time visualization of cellular repair. On page 355 of this issue, Roman *et al.* (*2*) show the importance of nuclear movement and local protein translation for repairing injured myofibers and uncover key components that regulate muscle reconstruction.

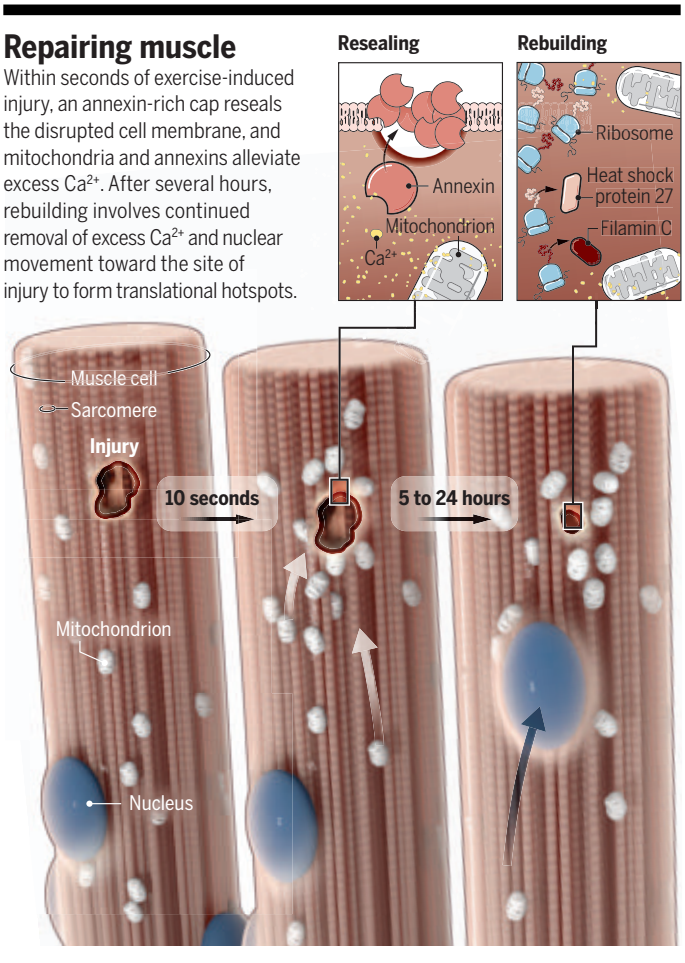
Severe muscle injury—for example, after trauma or surgery—destroys muscle architecture. This necessitates muscle regeneration, which is mediated by muscle-intrinsic stem cells, or satellite cells. Satellite cells contribute to muscle growth during periods of muscle development or after severe injury in which muscle architecture has been destroyed (*3*). However, with less-severe forms of muscle injury, cellular repair processes are independent of stem cell–mediated regeneration. In this mild injury paradigm, within seconds calcium accumulates within the myofiber adjacent to the site of membrane disruption (*4*, *5*). Resealing proteins are also immediately recruited to the injury site, including annexins, dysferlin, and mitsugumin-53 (MG53) (*5*). Annexins bind calcium, and multiple annexins (including annexins A1, A2, and A6) form a cap over the injured site; repair caps can be visualized within 10 s after membrane disruption (*6*). Furthermore, mitochondria near the injury site are required for this initial resealing to patch the cellular wound (*7*, *8*).

Roman *et al.* evaluated muscle repair after these initial phases of resealing, focusing on molecular events hours after muscle injury. They describe the formation of protein accumulations, called "scars," seen at 5 hours or even 24 hours after a single bout of acute exercise in mice and in biopsies taken from human muscle after exercise. The protein scars were enriched with filamin C and heat shock protein 27 (HSP27). Notably, animals that underwent exercise training before the acute exercise sessions had fewer scars, suggesting that exercise training may protect muscle from membrane breakdown.

The authors then carefully evaluated this later phase of cellular repair using a cultured cell model (myotubes). In cell culture, myotubes arise from the fusion of singly nucleated muscle precursor cells called myoblasts. Like muscle fibers, myotubes contain multiple nuclei per cell. Roman *et al.* discovered that myotube nuclei moved toward the injury site within 5 hours of injury and that nuclear movement was dependent on an intact cell division control protein 42 (CDC42) signaling cascade, a pathway that regulates cell motility. Inhibiting microtubule polymerization or dynein-mediated transport reduced the speed of nuclear movement. Repair of the sarcomeres at the site of injury was nearly complete 24 hours after injury. The authors described the formation adjacent to injury sites of translational hotspots, with messenger RNAs (mRNAs) that encode proteins found in scars, filamin C, and HSP27, overlapping with these translational hotspots.

In addition to being needed for resealing, mitochondria are also needed for rebuilding. During the rebuilding phase, the authors showed that mitochondria help sequester the excess calcium that accumulates at the injury site. The annexin-rich repair caps also remove calcium accumulation by extruding excess intracellular calcium into the extracellular space (*6*), and together, the data support that both annexins and mitochondria resolve calcium overload in early and late phases of resealing and rebuilding.

Cultured myotubes in a dish have much more myonuclear movement than does skeletal muscle in vivo (*9*). Cultured myotubes have increased cytoplasmic and nuclear fluidity be-




GRAPHIC: V. ALTOUNIAN/*SCIENCE*

[1]Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. [2]Department of Pharmacology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. Email: elizabeth.mcnally@northwestern.edu

cause they are not as full of sarcomeres, nor are they as rigidly held in place by extracellular scaffolds compared with myofibers in vivo. Consistent with this, Roman *et al.* found less nuclear movement in myofibers isolated from muscle compared with the nuclear movement of cultured myotubes in the dish, suggesting that the more fixed muscle environment restricts nuclear movement. Within the muscle, individual nuclei have defined nuclear domains in which mRNAs derived from a given nucleus are restricted in the subcellular areas in which protein translation takes place (*10*). Muscle injury may disrupt these nuclear domains.

Overall, these studies support a time course of myocellular repair that begins with an immediate resealing phase to urgently close the plasma membrane by means of the annexin repair cap. After this immediate resealing phase, reconstruction ensues, which requires calcium cleanup, new protein translation, and rebuilding of sarcomeres. Local mitochondria contribute to rebuilding through multiple mechanisms, including to help remove local excess calcium (see the figure). The intracellular network that supports myocellular rebuilding is microtubule dependent, whereas the resealing process requires actin (*2*, *4*–*6*). Resealing and repair in other cell types likely use similar pathways, although the need for sarcomere rebuilding is distinct to skeletal and cardiac muscle (*11*). Resealing proteins can be delivered to promote more rapid resolution of membrane tears, enhance resolution of associated inflammation, and enhance regeneration and rebuilding (*6*, *12*). These new findings on muscle rebuilding provide additional targets—such as HSP27, CDC42, and others—to enhance recovery after injury or with disease. ■


### REFERENCES AND NOTES

1. J. C. Bruusgaard, K. Liestøl, M. Ekmark, K. Kollstad, K. Gundersen, *J. Physiol.* **551**, 467 (2003).
2. W. Roman *et al.*, *Science* **374**, 355 (2021).
3. E. S. Folker, M. K. Baylies, *Front. Physiol.* **4**, 363 (2013).
4. J. R. McDade, A. Archambeau, D. E. Michele, *FASEB J.* **28**, 3660 (2014).
5. A. R. Demonbreun *et al.*, *J. Cell Biol.* **213**, 705 (2016).
6. A. R. Demonbreun *et al.*, *J. Clin. Invest.* **129**, 4657 (2019).
7. V. Debattisti *et al.*, *Cell Rep.* **29**, 1274 (2019).
8. A. Horn *et al.*, *Sci. Signal.* **10**, 495 (2017).
9. L. L. Englander, L. L. Rubin, *J. Cell Biol.* **104**, 87 (1987).
10. K. A. Murach, D. A. Englund, E. E. Dupont-Versteegden, J. J. McCarthy, C. A. Peterson, *Front. Physiol.* **9**, 635 (2018).
11. E. A. Scarborough *et al.*, *Nat. Commun.* **12**, 1547 (2021).
12. B. A. Whitson, T. Tan, N. Gong, H. Zhu, J. Ma, *Curr. Opin. Pharmacol.* **59**, 26 (2021).



### ACKNOWLEDGMENTS

The authors are supported by the National Institutes of Health (HL140398, NS047726, and HL061322). The authors are co-inventors on patent applications related to the use of recombinant annexin A6.


10.1126/science.abm2240

GRAPHIC: N. CARY/*SCIENCE*

CELL BIOLOGY

# A cell cycle regulator branches out

## A key $G_1$ kinase targets transcription to drive cell cycle commitment

*By* **Robert P. Fisher**

The commitment to cell division, at Start in yeast or the restriction (R) point in metazoans, is among the most strictly regulated events in the life of a cell. To promote this transition, a $G_1$-specific cyclin and cyclin-dependent kinase (CDK) pair—Cln3-Cdk1 in yeast or cyclin D–CDK4/6 in metazoans—induces the transcription of genes that encode cyclins needed at Start (Cln1 and Cln2) or the R point (cyclin E) (*1*). Functionally analogous transcription factors drive Start (SBF) and R-point passage (E2F), opposed by Whi5 and retinoblastoma (RB), respectively. During $G_1$, RB is sequentially phosphorylated and inactivated by cyclin D–CDK4/6 and cyclin E–CDK2, leading to stepwise activation of E2F. Whi5 is inactivated in a Cln3-Cdk1–dependent fashion as cells approach Start, seeming to complete an analogous circuit in yeast (*2*, *3*). On page 347 of this issue, Kõivomägi *et al.* (*4*) upend this model by revealing more direct communication between cell cycle and transcriptional machineries.

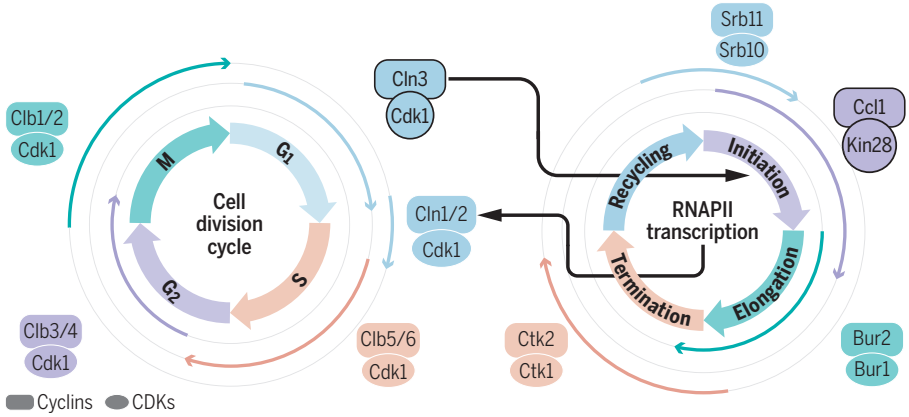

Department of Oncological Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
Email: robert.fisher@mssm.edu


The model of metazoan $G_1$ control has itself undergone recent revision, with the revelation that cyclin D–dependent kinases monophosphorylate RB on most available sites, rather than multiply phosphorylating select residues (*5*). How exclusive monophosphorylation is achieved, and the potential for functional specialization by different mono-phosphoRB isoforms, is beginning to be understood (*6*). A more complete picture of the budding yeast $G_1$ control network is also emerging, and perhaps diverging from the metazoan model.

Kõivomägi *et al.* now show that the distribution of phosphorylated Whi5 isoforms in $G_1$ is unaffected by Cln3 removal. By contrast, Whi5 hyperphosphorylation that coincides with Start is delayed, consistent with the Cln3 requirement to induce *CLN1* and *CLN2* expression and complete SBF activation. The electrophoretic mobilities of Whi5 phospho-isoforms argue against monophosphorylation and are more consistent with modification on multiple sites.

## Division and transcription cycles

Largely nonoverlapping sets of CDKs govern cycles of cell division and RNAPII transcription by phosphorylating key proteins in each stage. In yeast, the cyclin Cln3 activates Cdk1 in $G_1$ to drive expression of SBF target genes (*CLN1* and *CLN2*) needed for cell cycle commitment and S-phase entry. To do so, Cln3-Cdk1 takes on a function that is usually performed by the transcriptional CDK Ccl1-Kin28 (the TFIIH-associated kinase): phosphorylation of the RNAPII subunit Rpb1 CTD to promote the early steps of the transcription cycle.




CDK, cyclin-dependent kinase; CTD, carboxyl-terminal domain; RNAPII, RNA polymerase II; TFIIH, transcription factor IIH.

Further studies will be needed to uncover the importance of these phosphorylations and to identify the responsible kinase or kinases. Cln3 and Whi5 inputs can also be separated genetically, as indicated by additive advancement of Start when *whi5* is deleted and a *CLN3* gain-of-function mutation is expressed. In retrospect, the enzyme-substrate relationship inferred for Cln3-Cdk1 and Whi5 was suspect. Genetic data implicated $G_1$ cyclins generically, not Cln3 specifically, in Whi5 inactivation and nuclear exclusion (*2*, *3*), whereas biochemical evidence for Cln3-directed Whi5 phosphorylation consisted of a single result (*2*), now shown by Kõivomägi *et al.* to likely be an artifact. Whi5 is actually a poor substrate for Cln3-Cdk1 in vitro. Cdk1 activity is nevertheless required for Cln3 to execute its $G_1$ function, raising the following question: What is the relevant Cln3-Cdk1 substrate?

To answer this, Kõivomägi *et al.* screened candidate substrates and found a surprising top hit: the carboxy-terminal domain (CTD) of the RNA polymerase II subunit, Rpb1. This repetitive structure is the nexus of a conserved signaling network that comprises multiple CTD kinases and phosphatases and CTD-binding proteins, including RNA-processing and chromatin-modifying machineries, the ordered recruitment of which is largely determined by the CTD phosphorylation state. Indeed, the transcription cycle resembles the cell cycle in its reliance on multiple cyclin-dependent CTD kinases (*7*, *8*). Kõivomägi *et al.* provide compelling evidence that the Rpb1 CTD is a (or the) physiologically relevant substrate of Cln3-Cdk1. Cln3 is recruited specifically to SBF-bound promoter regions, where the capping enzyme Ceg1—a marker of phosphorylation at $Ser^5$ of the CTD repeat—is rapidly lost when Cdk1 is inhibited. In addition, forced recruitment (by tethering to SBF) of a known $Ser^5$ kinase, Ccl1-Kin28—the cyclin-CDK associated with transcription factor TFIIH—rescues normal transcription of SBF targets and timing of Start in cells lacking Cln3. Therefore, Cln3-Cdk1 performs its rate-limiting function by taking over a role—CTD phosphorylation to promote $G_1$ transcription—usually performed by a cyclin-CDK pair dedicated to transcriptional regulation (see the figure).

CTD phosphorylation by kinases other than transcriptional CDKs is not unprecedented. In mouse embryonic stem cells, the extracellular regulated kinases ERK1 and ERK2 substitute for TFIIH at developmental promoters (*9*). Similarly, gene-specific CTD kinase function has been ascribed to human dual specificity tyrosine-phosphorylation-regulated kinase 1A (DYRK1A) (*10*) and the yeast CK1 ortholog Hrr25 (*11*). There is also evidence, albeit rare, for cross-talk between the CDK subfamilies that govern cell division and transcription. For example, as Kõivomägi *et al.* note, in metazoans, the CDK-activating kinase (CAK) needed for cell cycle progression is CDK7, an ortholog of Kin28 (*8*).

Among questions for future investigation is whether using a cell cycle CDK for an essential step in transcription is distinctive to yeast $G_1$ control. Another is whether transcriptional CDKs ever step in to perform cell cycle functions. Naïvely, this seems more risky; transcriptional CDKs are generally active throughout the cell cycle and perhaps ill-suited for precise temporal regulation (*8*). Cell cycle–regulated transcription, however, might require specialized or specifically localized forms of enzymes that act constitutively in other contexts. For example, activating phosphorylation of human CDK7, specifically in the chromatin-bound population, is up-regulated upon entry into $G_1$ from $G_0$, concomitant with CDK7-dependent activation of CDK4 (*12*). This might also be a way to promote $G_1$-specific gene transcription, analogous to the role of Cln3-Cdk1.

It also will be important to determine the division of labor between Cdk1 and Kin28 during transcription of cell cycle genes: As Kõivomägi *et al.* observe, both Cln3-Cdk1 and Ccl1-Kin28 bind SBF-driven promoters, and previous studies indicate that Kin28 recruitment depends on Cdk1 at SBF-regulated genes (*13*). Whether the two CDK family members act in competition, collaboration, or temporal sequence on their common substrate, the Rpb1 CTD, will be an intriguing question for future investigation. Moreover, both transcriptional and $G_1$ cyclin-CDKs have emerged as potential drug targets in cancer (*8*, *14*); understanding how they collaborate might guide efforts to develop new combination therapies. ■

## REFERENCES AND NOTES


1. C. Bertoli, J. M. Skotheim, R. A. de Bruin, *Nat. Rev. Mol. Cell Biol.* **14**, 518 (2013).
2. M. Costanzo *et al.*, *Cell* **117**, 899 (2004).
3. R. A. de Bruin, W. H. McDonald, T. I. Kalashnikova, J. Yates III, C. Wittenberg, *Cell* **117**, 887 (2004).
4. M. Kõivomägi, M. P. Swaffer, J. J. Turner, G. Marinov, J. M. Skotheim, *Science* **374**, 347 (2021).
5. A. M. Narasimha *et al.*, *eLife* **3**, e02872 (2014).
6. I. Sanidas *et al.*, *Mol. Cell* **73**, 985 (2019).
7. S. Buratowski, *Mol. Cell* **36**, 541 (2009).
8. P. K. Parua, R. P. Fisher, *Nat. Chem. Biol.* **16**, 716 (2020).
9. W. W. Tee, S. S. Shen, O. Oksuz, V. Narendra, D. Reinberg, *Cell* **156**, 678 (2014).
10. C. Di Vona *et al.*, *Mol. Cell* **57**, 506 (2015).
11. C. M. Nemec *et al.*, *Nat. Chem. Biol.* **15**, 123 (2019).
12. M. M. Schachter *et al.*, *Mol. Cell* **50**, 250 (2013).
13. M. P. Cosma, S. Panizza, K. Nasmyth, *Mol. Cell* **7**, 1213 (2001).
14. J. M. Suski, M. Braun, V. Strmiska, P. Sicinski, *Cancer Cell* **39**, 759 (2021).


## ACKNOWLEDGMENTS

R.P.F. is supported by National Institutes of Health grant R35 GM127289.

10.1126/science.abm2010

DEVELOPMENT

# Not all stress is bad for your heart

## Shear stress forces instruct valve formation during cardiac development

*By* **Rajan Jain** *and* **Jonathan A. Epstein**

Within the first month of human gestation, the heart begins to contract, even though it is still undergoing morphogenesis, including the development of valves to promote proper blood flow. Thus, the heart is subjected to mechanical forces throughout development, but whether mechanical forces instruct morphogenesis has remained unknown. Although mechanoresponsive factors have been identified, it is challenging to directly interrogate the role of mechanical inputs during cardiogenesis and to decipher how such forces are transduced into molecular signals in vivo (*1*). Mechanoregulation of cardiac development is clinically relevant because diminished blood flow is thought to contribute to the pathogenesis of hypoplastic left heart syndrome, a congenital heart disease associated with valve abnormalities (*2*). On page 351 of this issue, Fukui *et al.* (*3*) show how mechanical forces are sensed by endocardial cells (the layer of endothelial cells lining the inside of the heart) to direct cardiac valve formation in zebrafish.

The early heart tube is composed of a single layer of beating myocardial cells lined by a single layer of endocardium. As the heart tube loops, valves form between the atria and the ventricles, and between the ventricles and the arteries exiting the heart. Until recently, the study of valve formation has focused on the signaling cascades that affect endocardial cells in the valve primordia, triggering those cells to undergo endothelial-mesenchymal transition, delamination, and migration into the underlying matrix between the myocardial

Department of Medicine, Department of Cell and Developmental Biology, Penn Cardiovascular Institute, Penn Institute for Regenerative Medicine, Penn Epigenetics Institute, Perelman School of Medicine, University of Pennsylvania, PA 19104, USA. Email: epsteinj@pennmedicine.upenn.edu; jainr@pennmedicine.upenn.edu

and endocardial layers, proliferation of the resulting mesenchyme, and maturation into elongated valve leaflets (*4*).

The process of valve formation is appreciated in developmental biology as an example of inductive interactions between adjacent tissues mediated by growth factors, cell-cell adhesion molecules, cell surface ligands and their receptors on neighboring cells, and intracellular signaling cascades that affect transcription and cell fate. Key mediators include those of the Notch, Wnt, vascular endothelial growth factor (VEGF), transforming growth factor–β (TGFβ), and epidermal growth factor (EGF) families, and nuclear factor of activated T cells (NFAT), Hippo, RAS, and calcium signaling cascades that modulate transcription factor activity, including Kruppel-like factors (KLFs) (*4*–*6*). Elucidation of the molecular effectors of inductive interactions between tissues has contributed to a deeper understanding of embryogenesis (*7*, *8*), and disruptions of specific developmental pathways have been identified as causative in dozens of congenital disorders.

More recently, biologists have appreciated that mechanical forces affecting the cell surface, or the nucleus, can modulate many of the same signaling cascades involved in inductive tissue interactions to affect gene expression and phenotype (*9*). For example, in the adult cardiovascular system, shear stress caused by regional blood flow contributes to atherosclerosis and valve calcification (*10*). Mechanosensitive Piezo and transient receptor potential (TRP) channels have been described on endothelial (and other) cells that transform mechanical forces into calcium signals or activate intracellular signaling pathways (*11*, *12*).

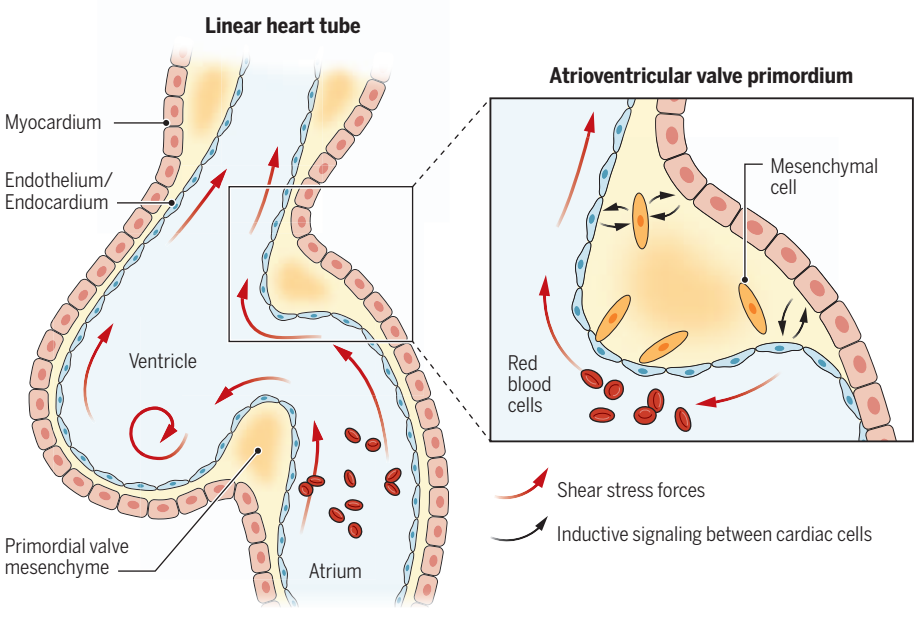
**"...mechanical forces are sensed by endocardial cells...to direct cardiac valve formation in zebrafish."**

Fukui *et al.* integrate the fields of mechanosensation with the classic understanding of valve morphogenesis using zebrafish as a model system and elegant micromanipulation techniques to modulate forces on specific endocardial cells within the embryonic heart. Using magnetic tweezers, they control the position of a small bead within the heart to modulate the amount of shear stress on endocardium. Zebrafish hearts can be arrested for short periods of time, which the authors leverage in their study. They show that shear stress produced by insertion of the bead is sufficient to induce ectopic valve-like structures, which express activated Nfat, Klfs, and Wnt9, which are markers of valve tissue. Ectopic shear stress results in $Ca^{2+}$ oscillations that activate Nfat signaling, which is dependent on adenosine triphosphate (ATP) receptors on the cell surface. Manipulation of ATP receptors and $Ca^{2+}$ transients causes aberrant Nfat activity and ectopic valve-like structures. Thus, local ATP concentrations are mechanosensitive triggers for downstream signaling. Notably, this pathway is independent of TRP and Piezo channels because loss-of-function mutations in these channels did not affect shear stress–induced $Ca^{2+}$ activation. The implication of these results is that shear stress resulting from regional blood flow differences in the early developing heart results in activation of specific signaling pathways that coordinate with inductive interactions between local tissues to specify the process of valve formation (see the figure).

Although Fukui *et al.* focused their studies on endocardial cells and $Ca^{2+}$-Nfat signaling, it will be exciting in the future to further define the complexity of how mechanosensing integrates with classic inductive signals, and how local shear stress may affect the development of other cardiac cell types. Moreover, it will be important to determine whether the specific pathways described in zebrafish can be extended to mouse and human development. Although the technologies used in zebrafish may be difficult to extend to mammalian models, other approaches may be possible. For example, three-dimensional organoid models (*13*) may facilitate understanding of how mechanosignaling affects aspects of human cardiac morphogenesis and disorders such as hypoplastic left ventricle syndrome. Similarly, development and incorporation of reporters of mechanosignaling will allow for the elucidation of the array of pathways that are likely affected by shear stress. ■

## Shear forces direct valve morphogenesis

Shear stress–induced mechanosignaling in the linear heart tube integrates with inductive signaling between the myocardium and endocardium to drive valve morphogenesis in zebrafish.




GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

### REFERENCES AND NOTES


1. C. J. Chan, C. P. Heisenberg, T. Hiiragi, *Curr. Biol.* **27**, R1024 (2017).
2. E. J. Hickey, C. A. Caldarone, B. W. McCrindle, *J. Am. Coll. Cardiol.* **59** (suppl.), S43 (2012).
3. H. Fukui *et al.*, *Science* **374**, 351 (2021).
4. E. J. Armstrong, J. Bischoff, *Circ. Res.* **95**, 459 (2004).
5. C. P. Chang *et al.*, *Cell* **118**, 649 (2004).
6. J. L. de la Pompa, J. A. Epstein, *Dev. Cell* **22**, 244 (2012).
7. N. Bazin-Lopez, L. E. Valdivia, S. W. Wilson, G. Gestri, *Curr. Opin. Genet. Dev.* **32**, 73 (2015).
8. S. M. Myllymäki, M. L. Mikkola, *Curr. Opin. Cell Biol.* **61**, 72 (2019).
9. A. L. Duchemin, H. Vignes, J. Vermot, R. Chow, *Curr. Opin. Genet. Dev.* **57**, 106 (2019).
10. V. Mehta *et al.*, *Sci. Adv.* **7**, eabg5060 (2021).
11. A. L. Duchemin, H. Vignes, J. Vermot, *eLife* **8**, e44706 (2019).
12. V. Shah, S. Patel, J. Shah, *Dev. Dyn.* dvdy.401 (2021).
13. P. Hofbauer *et al.*, *Cell* **184**, 3299 (2021).


### ACKNOWLEDGMENTS


The authors are supported by the NIH (R35HL140018 to J.A.E. and R01HL139783 to R.J.), NSF (15-48571 to R.J.), Burroughs Wellcome Fund (R.J.), and the American Heart Association–Allen Foundation (R.J.).

10.1126/science.abm1858

BOOKS *et al.*

Violet Lawson facilitates a controlled burn in Australia's Northern Territory in January 2020.

PHOTO: MATTHEW ABBOTT/*THE NEW YORK TIMES*

ECOLOGY

# The history and future of fire

## More prescribed burns and less fossil fuel combustion could help rebalance our relationship with fire

*By* **Mary Ellen Hannibal**

In *The Pyrocene*, Stephen J. Pyne takes a measured, historical, and ecological approach to fire. A revered fire historian and a prolific author on the subject, he argues that we ought to embrace and not shun combustion. For one thing, we need it. And for another thing, it is not going away.

Pyne structures his history of fire in three phases. "First-fire" is fire that arises from nature. For millennia, notes Pyne, evolving plant life was regularly burned by lightning-driven flames. "As soon as plants rooted on land, lightning set them ablaze. They've been burning ever since," he writes. Burning decomposes biomass, taking apart what photosynthesis has cohered. In doing so, it helps redistribute and stimulate the elements of life. Burned biomes regenerate and thrive.

"Second-fire," fires set and exacerbated by humans, appeared more than 2 million years ago, when early hominins captured flames from the environment and directed them with intention, argues Pyne. Fire was a tool, a weapon, and a complete game changer for the creatures who would eventually give rise to *Homo sapiens* and their cousins. Pyne credits one particular "pyrotechnology"—cooking—for changing the evolutionary trajectory of hominins. Processing meat and vegetation with fire essentially predigests it, allowing the human gut to complete the combustion process more quickly and efficiently. This led to better nutrition and increased the amount of time available to develop tools, resulting in the burgeoning dominance of our once unassuming ancestors.

Humanity's use of fire was initially bounded by ecological parameters such as seasonality and weather, which influenced the timing and extent of intentionally lit fires. Indigenous people burned landscapes in cycles to stimulate the growth of various plant species, not only to use the plants themselves but also to attract and hunt the creatures that arrived with the various successional stages of vegetation.

And then "third-fire"—fire "no longer bounded by such ecological limits as fuel, season, sun, or the rhythms of wetting and drying"—came along. During this period, humans figured out how to move combustion into machines and to direct flame with ever more increasing intensity. Alongside this revolution came an even more impactful upheaval in fuels. Humanity once focused on setting fire to aboveground products of photosynthesis—trees, shrubs, and grasses. With third-fire, we turned to belowground sources of yesterday's photosynthesis such as coal, oil, and natural gas. The result, argues Pyne, has empowered "humanity on a gargantuan scale." All of this combustion, now freed from the fetters of ecological seasonality, has pumped effluents into the atmosphere that have "unhinged the old climate" and are currently driving record warming and drought.

Third-fire is not only characterized by more combustion of underground fuel reserves but by too little burning of aboveground biomass. Indigenous burning practices have been curtailed nearly worldwide, and biomes that evolved alongside first- and second-fire are now often deprived of both.

Pyne lays out three resulting paradoxes. The first is that the more we try to remove fire from places that coevolved with it, "the more violently [it] will return." This reality is currently playing out in California and many other landscapes around the world. The second paradox is that as media attention highlights the woes of wildfire bearing down on human-built environments, "the amount of land actually burned overall is shrinking." In 2020, writes Pyne, 4.2 million acres of California burned, but in preindustrial times, it is likely that more than 10 million acres would have burned in a single year. Pyne's third paradox emphasizes the need for less of some types of fire and more of others. We must ratchet down fossil fuel burning, he argues, and we must burn living landscapes more widely and frequently.

**The Pyrocene: How We Created an Age of Fire, and What Happens Next**
*Stephen J. Pyne*
University of California Press, 2021. 192 pp.

One of the most intriguing arguments Pyne makes in this brief but highly impactful book is that we must reconsider fire from a cultural perspective. The Enlightenment enabled us to decouple the components of fire and to categorize them as chemistry and physics, he argues. When fire was incorporated into machines, its ubiquity and power became invisible, and we forgot about it as a fundamental part of our history. Framing our current age as the Pyrocene centers the role of fire in how humans are shaping the planet and "integrates its surroundings—geographic, historical, institutional, intellectual. It addresses the search for a usable future." ■

10.1126/science.abl9129

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

BIOMEDICINE

# Toward a holistic model of Alzheimer's

## An overemphasis on amyloid has stymied progress on the disease, alleges a neuroscientist

*By* **Tara Spires-Jones**

In his new book *How Not to Study a Disease: The Story of Alzheimer's*, neurobiologist Karl Herrup presents a scathing critique of the amyloid cascade hypothesis of Alzheimer's disease—the idea that the accumulation of amyloid beta in the brain kicks off a chain of events that causes dementia—and the effect this hypothesis has had on Alzheimer's research. Herrup presents an engaging history of the field that includes the origins and evolution of the amyloid hypothesis, before putting forward three "inflations" that, in his opinion, have led to it having an outsized influence on how we think about Alzheimer's.

The first incident of inflation described in the book occurred in 1910, when Alois Alzheimer and Emil Kraepelin first described the condition that would eventually become known as Alzheimer's disease in a textbook. The pair's observations, which linked amyloid plaques and tangles of tau proteins with dementia symptoms, were based on a single patient. Since that time, amyloid and small amounts of tau have also been detected in the brains of healthy elderly people, but Alzheimer and Kraepelin's initial inklings were not completely without merit: The presence of both pathologies in the characteristic late-stage "Alzheimer's" pattern is very rarely observed without accompanying dementia symptoms.

Throughout the book, Herrup's focus on dementia being "caused by amyloid" glosses over the very important distinction that a diagnosis of Alzheimer's disease includes requisite tau pathology in specific regional patterns in the brain. There are few contemporary scientists who would argue that amyloid alone is sufficient to cause dementia.

Herrup's second inflation hinges on the politics of how Alzheimer's research is funded in the United States via the National Institute on Aging (NIA). Here, he provides a very interesting description of the early years of the NIA, including the political maneuvering that sought to secure more funding from the aging politicians who control the agency's budget. This section is very well argued, if US-centric, but it implies that there has been too much funding for Alzheimer's research. When compared with other biomedical fields, dementia frequently lags far behind in funding. In the UK in 2012, for example, dementia research received about one-sixth the funding that cancer research did (*1*).

Charles Flagg took part in an early trial of the amyloid-clearing drug aducanumab.

PHOTO: BRIAN SNYDER/REUTERS/ALAMY STOCK PHOTO

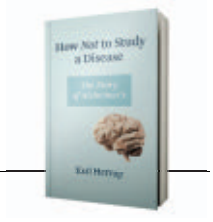
The final inflation, in Herrup's estimation, came from the 2011 revision of the diagnostic criteria for Alzheimer's, which included amyloid biomarkers and a defined preclinical phase of disease. Although Herrup rightly points out that the presence of plaques alone does not mean a person is on the path to Alzheimer's dementia, the revised criteria include biomarkers of neurodegeneration, tau, and functional changes. The presence of tau is also included in the preclinical definition. The inclusion of these other factors increases the likelihood that a diagnosis will truly reflect the presence of preclinical Alzheimer's disease rather than harmless age-related plaque accumulation.

Herrup proposes a well-supported holistic model of Alzheimer's pathogenesis that encourages research into the contributions of all cell types in the brain. Such a strategy has also recently been put forward by other investigators, including myself (*2*, *3*). However, he makes a less-convincing assertion that we should return to a definition of Alzheimer's disease solely based on dementia symptoms. Different patterns of brain atrophy and different pathological lesions can result in clinically similar dementias. Inherent differences in underlying pathology can be of vital importance in clinical settings, dictating which drugs are likely to work best, for example (*4*).

In June 2021, after this book was written, the first disease-modifying drug for Alzheimer's disease—aducanumab, which clears amyloid beta from the brain—was approved by the US Food and Drug Administration (*5*). Herrup discusses the lead-up to this approval, but I am sure that many more critiques will be coming soon. (If you think the amyloid cascade hypothesis is controversial, just wait until you learn about the approval of this drug.)

Despite the hiccups highlighted in this book, there has been remarkable progress in Alzheimer's research in the past several decades that has led to a better understanding of this inordinately complex disease. Although we may not have triumphed over it yet, the broadening of research avenues suggested in this book and embraced by many in the field is likely to result in life-changing treatments for dementias. ■

**How Not to Study a Disease: The Story of Alzheimer's**
*Karl Herrup*
MIT Press, 2021. 272 pp.



The reviewer is at the UK Dementia Research Institute and the Centre for Discovery Brain Sciences, The University of Edinburgh, Edinburgh EH8 9JZ, UK.
Email: tara.spires-jones@ed.ac.uk

### REFERENCES AND NOTES

1. R. Luengo-Fernandez, J. Leal, A. Gray, *BMJ Open* **5**, e006648 (2015).
2. B. De Strooper, E. Karran, *Cell* **164**, 603 (2016).
3. C. M. Henstridge, B. T. Hyman, T. L. Spires-Jones, *Nat. Rev. Neurosci.* **20**, 94 (2019).
4. I. Yunusa, A. Alsumali, A. E. Garba, Q. R. Regestein, T. Eguale, *JAMA Netw. Open* **2**, e190828 (2019).
5. A. Mullard, *Nature* **594**, 309 (2021).

10.1126/science.abl7597

As cities such as Shanghai develop, increasing numbers of glass buildings make bird collisions with windows more likely.

*Edited by* **Jennifer Sills**

## Bird-friendly buildings for China's cities

China's rapid urbanization has resulted in the proliferation of glass buildings, as has been the case elsewhere in the world. Bird collisions with glass buildings are now an important global factor in bird mortality (*1*). To prevent bird deaths in a country that hosts numerous vulnerable species, China should invest in research on bird-friendly building design and update building codes and regulations accordingly.

During the day, birds fail to perceive transparent windows as solid surfaces because the glass reflects images of the sky or local habitat (*2*). At night, artificial light can disorient and attract birds, leading to more collisions (*3*). In the United States, such collisions kill as many as 365 million to 988 million birds annually (*4*).

Three major migratory bird flyways (East-Asia Australasian, West Pacific, and Central Asian) cross China, covering almost the entire country (*5*). Tens of millions of migrating birds fly across China each year, with more than 400 bird species using the East-Asia Australasian flyway alone (*6*). The growth of tall buildings now poses a severe collision risk to birds. In 2020, 106 buildings exceeding 200 m in height were completed globally; more than half of them were in China (*7*), but little attention has been given to collision threats. Preliminary data from a citizen science project suggest that most collisions occur in rapidly urbanizing provinces such as Guangdong, Shanghai, and Zhejiang (*8*). Of the birds that were observed striking buildings, 78% died from their injuries (*8*).

China and other countries with substantial numbers of glass buildings should work to minimize bird collisions. Films or netting applied to the outside of windows are most effective (*9*). New building regulations should require glass that is less reflective, translucent, or opaque, and windows should be angled or have smaller panes (*10*). Green building standards already require energy efficiency and the use of recyclable building materials (*11*). Architectural green credentials should be expanded to include features that mitigate bird collisions with buildings.

**Hong Yang[1,2]*, Xianjin Huang[2], Julian R. Thompson[3], Roger J. Flower[3]**
[1]School of Geography and Ocean Science, Nanjing University, Nanjing 210023, China. [2]Department of Geography and Environmental Science, University of Reading, Reading RG6 6AB, UK. [3]Department of Geography, University College London, London WC1E 6BT, UK.
*Corresponding author.
Email: hongyanghy@gmail.com

### REFERENCES AND NOTES

1. S. R. Loss, T. Will, P. Marra, *Annu. Rev. Ecol. Evol. Syst.* **46**, 99 (2015).
2. C. S. Machtans, C. H. R. Wedeles, E. M. Bayne, *Avian Conserv. Ecol.* **8**, 6 (2013).
3. M. Rebke *et al.*, *Biol. Conserv.* **233**, 220 (2019).
4. US Fish & Wildlife Service, "Threats to birds: Migratory bird mortality—questions and answers" (2021); www.fws.gov/birds/bird-enthusiasts/threats-to-birds.php.
5. D. L. Yong *et al.*, *Front. Ecol. Evol.* 10.3389/fevo.2021.613172 (2021).
6. Y. J. Xu *et al.*, *Landsc. Ecol.* **34**, 243 (2019).
7. Council on Tall Buildings and Urban Habitat, "CTBUH year in review: Tall trends of 2020" (2021); www.skyscrapercenter.com/research/CTBUH_ResearchReport_2020YearInReview.pdf.
8. "123 bird-window collisions in three months" (2021); www.163.com/dy/article/GDTUQH420512RQI0.html [in Chinese].
9. N. Ocampo-Penuela *et al.*, *PeerJ* **4**, e1652 (2016).
10. D. Klem *et al.*, *Wilson Bull.* **116**, 69 (2004).
11. World Green Building Council, "What is green building?" (2021); www.worldgbc.org/what-green-building.

10.1126/science.abm3221

## Community stewardship of China's national parks

When striving to conserve natural habitats, ensuring that "wilderness" is completely free of human activities is not always the best option. Conservation goals can be achieved by seeking a balance between sustainable community development and nature conservation (*1*, *2*). This strategy has been successful in China, where ethnic minorities live in some of the country's most vulnerable protected areas.

The Qinghai-Tibet Plateau is in urgent need of environmental protection (*3*, *4*). In 2015, China began a pilot program in the region's Three-River-Source National Park to explore whether institutional innovation, especially economic assistance, can help the local community protect the environment (*5*, *6*). The 123,100-km$^2$ park is home to more than 70,000 herdsmen, many of whom lived in relative poverty (*7*). The residents, mainly Tibetan, are highly dependent on traditional husbandry. To benefit both the herdsmen and the environment, the park offered to employ one member of each family as a conservationist paid by government (*8*). The herders are paid to patrol the region to help prevent harmful activities such as illegal hunting and overgrazing.

Under the policy, more than 17,000 herdsmen have each earned about USD2930 annually as conservationists (*8*). This income, coupled with other opportunities to earn income such as crafting and employment in green industries (*9*), has improved the living standards of residents (*10*, *11*). Since the program began, the environmental quality of the Three-River-Source National Park has also improved:

PHOTO: GZ LAOBAI/SHUTTERSTOCK

Grassland coverage and grass yield have increased by more than 11% and 30%, respectively, compared with data from 10 years ago (*5*).

The success of this program in China should serve as a model for other countries that hope to conserve the environment and support the residents who live in vulnerable ecosystems. Balancing human economic development with environmental protection could also justify the designation of additional national parks.

**Yuyao Feng[1,2], Guowen Li[3], Jianping Li[4], Xiaolei Sun[1,2], Dengsheng Wu[1,2]***
[1]Institutes of Science and Development, Chinese Academy of Sciences, Haidian District, Beijing 100190, China. [2]School of Public Policy and Management, University of Chinese Academy of Sciences, Shijingshan District, Beijing 100049, China. [3]School of Management Science and Engineering, Central University of Finance and Economics, Haidian District, Beijing 100081, China. [4]School of Economics and Management, University of Chinese Academy of Sciences, Haidian District, Beijing 100190, China.
*Corresponding author. Email: wds@casipm.ac.cn

### REFERENCES AND NOTES

1. T. O'Riordan, S. Stoll-Kleemann, Eds., *Biodiversity, Sustainability and Human Communities: Protecting beyond the Protected* (Cambridge University Press, 2002).
2. US Environmental Protection Agency (EPA), "EPA's environmental justice collaborative problem-solving model" (Publication 300-R-06-002, EPA, 2008); www.epa.gov/sites/production/files/2015-02/documents/cps-manual-12-27-06.pdf).
3. Y. Zhao *et al*., *Sci. Adv*. **6**, eaay6193 (2020).
4. H. Xin, *Science* **321**, 633 (2008).
5. "China's national park system starts to yield fruits" (The State Council of the People's Republic of China, 2020); http://english.www.gov.cn/statecouncil/ministries/202008/25/content_WS5f447014c6d0f7257693af8d.html.
6. J. Li, W. Wang, J. C. Axmacher, Y. Zhang, Y. Zhu, *Science* **351**, 1160 (2016).
7. "China to complete building national park on 'roof of the world,'" *China Daily* (2020); http://global.chinadaily.com.cn/a/202009/10/WS5f599e39a310675eafc58c4a.html.
8. W. Yang, "Balancing growth and protection" (The State Council of the People's Republic of China, 2020); http://english.www.gov.cn/news/topnews/202009/28/content_WS5f71382cc6d0f7257693cc9c.html.
9. "Millions of herdsmen adapt to new jobs in new industries, which promotes the green development of the Three River Source region" (National Park Administration of China, 2019); www.forestry.gov.cn/main/5497/20190704/083703222939020.html [in Chinese].
10. "China's epic journey from poverty to prosperity," *China Daily* (2021); http://global.chinadaily.com.cn/a/202109/29/WS6153c41aa310cdd39bc6c469_11.html.
11. "China's Qinghai restores 574,000 hectares of desertified land," *China Daily* (2021); http://english.www.gov.cn/news/topnews/202107/16/content_WS60f1458ac6d0df57f98dd19a.html.

10.1126/science.abm2665

# Fences undermine biodiversity targets

Fences are often built to serve a specific purpose without regard for their effects on ecological processes, landscapes, and biodiversity. The Australian dingo fence (*1*) and the extensive Returning Grazing Lands to Grasslands project in China (*2*) were built to facilitate livestock grazing and restoration. European countries have erected fences to mark national borders and prevent refugee movement (*3*). China has begun building a fence along its southwestern border to curb the spread of COVID-19 (*4*). Although fences can successfully serve their narrow purposes to protect vulnerable areas and endangered species (*5*) or make political statements (*3*), they also have unintended consequences. These barriers increase habitat fragmentation, inadvertently trap and kill wildlife, and impede genetic exchange (*6*).

Fences can disrupt migratory routes. In South Africa, fences pose threats to populations of iconic migratory ungulates such as wildebeests (*7*). Habitat fragmentation, amplified by climate change, can have unexpected cascading effects on biodiversity and ecosystem services. In Botswana, fences prevented elephants from free movement. As a result, the elephants drank water from ponds in an area where warming had increased cyanotoxin blooms, triggering mass mortality (*8*). Unfortunately, even when fences are built in the name of conservation, we often focus on target species and overlook the broader impacts on other species and biodiversity (*9*).

Biodiversity is important for sustaining global food security, livelihoods, and economic development as well as for preventing future epidemics (*10*). Halting biodiversity loss is the cornerstone of Sustainable Development Goal 15 (Life on Land) of the United Nations' 2030 Agenda for Sustainable Development (*11*). In this light, the rapid proliferation of fences around the world should be carefully reconsidered and precautionary actions should be taken to mitigate risks. Reopening wildlife migration corridors across national boundaries by removing fences should be encouraged, and the necessity of erecting new fences (*4*), including the plans for a COVID-19 fence in China, should be assessed on a case-by-case basis. Where unavoidable, transient fencing (*6*) or permeable fencing could be a useful tool, but all kinds of fencing require a deeper understanding of the longer-term biodiversity implications.

**Jian Sun[1]*, Eryuan Liang[1], Isabel C. Barrio[2], Ji Chen[3,4,5], Jinniu Wang[6], Bojie Fu[7,8]**
[1]State Key Laboratory of Tibetan Plateau Earth System, Resources and Environment, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100101, China. [2]Faculty of Environmental and Forest Sciences, Agricultural University of Iceland, Reykjavík 112, Iceland. [3]Department of Agroecology, Aarhus University, 8830 Tjele, Denmark. [4]Aarhus University Centre for Circular Bioeconomy, Aarhus University, 8830 Tjele, Denmark. [5]CLIMATE Interdisciplinary Centre for Climate Change, Aarhus University, 4000 Roskilde, Denmark. [6]Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, China. [7]State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China. [8]State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China.
*Corresponding author.
Email: sunjian@itpcas.ac.cn

### REFERENCES AND NOTES

1. "NSW wild dog fence extension project" (2021); www.lls.nsw.gov.au/what-we-do/our-major-projects/nsw-wild-dog-fence-extension-project.
2. J. Hong, X. Xu, B. Pang, X. Ma, X. Wang,. *Land Degrad. Devel*. **32**, 535 (2021).
3. J. D. C. Linnell *et al*., *PLOS Biol*. **14**, e1002483 (2016).
4. "Chinese authorities take a leaf from the Trump playbook and 'build the wall' as part of Covid-19 curbs" (2021); www.scmp.com/news/china/science/article/3141126/chinese-authorities-take-leaf-trump-playbook-and-build-wall-part.
5. J. Ringma, M. D. Barnes, M. Bode, *Conserv. Sci. Pract*. **2**, e168 (2020).
6. J. Sun *et al*., *Sci. Bull*. **65**, 1405 (2020).
7. M. J. Kauffman *et al*., *Science* **372**, 566 (2021).
8. H. Wang *et al*., *Innovation* **2**, 100092 (2021).
9. D. Smith, R. King, B.L. Allen, *Biol. Rev*. **95**, 1590 (2020).
10. J. Tollefson, *Nature* **584**, 175 (2020).
11. United Nations, Department of Economic and Social Affairs, "Goal 15: Life on land" (2018); https://sdgs.un.org/goals/goal15.

10.1126/science.abm3642

## TECHNICAL COMMENT ABSTRACTS

### Comment on "Reintroduction of the archaic variant of *NOVA1* in cortical organoids alters neurodevelopment"

**Tomislav Maricic, Nelly Helmbrecht, Stephan Riesenberg, Dominik Macak, Philipp Kanis, Martin Lackner, Agrafena Daria Pugach-Matveeva, Svante Pääbo**

Trujillo *et al*. (Research Articles, 12 February 2021, eaax2537) conclude that the reintroduction of an ancestral amino acid substitution in the protein NOVA1 drastically alters the development of brain organoids. We show that cell lines used by the authors carry heterozygous deletions of the target DNA sequence, providing another plausible explanation for the effects observed.

Full text: dx.doi.org/10.1126/science.abi6060

### Response to Comment on "Reintroduction of the archaic variant of *NOVA1* in cortical organoids alters neurodevelopment"

**Roberto H. Herai, Ryan A. Szeto, Cleber A. Trujillo, Alysson R. Muotri**

Maricic *et al*. performed an undisclosed in silico–only whole-exome sequencing analysis of our data and found genomic alterations previously undetected in some clones. Some of the predicted alterations, if true, could change the original genotype of the clones. We failed to experimentally validate all but one of these genomic alterations, which did not affect our previous results or data interpretation.

Full text: dx.doi.org/10.1126/science.abi9881

HAPPY BIRTHDAY

**PERSPECTIVE**
Policing social unrest and collective violence p. 272

**POLICY FORUM**
The corporate shadow in democratic policing p. 274

**REVIEWS**
Assessing mass incarceration's effects on families p. 277
Exclusion and exploitation: The incarceration of Black Americans from slavery to the present p. 282
Toward an understanding of structural racism: Implications for criminal justice p. 286
The predatory dimensions of criminal justice p. 291

**RELATED ITEMS**
EDITORIAL p. 237 NEWS STORY p. 251 PODCAST

# AN OUTLIER OF INJUSTICE

*By* **Tage Rai*** *and* **Brad Wible**

For much of the 20th century, the incarceration rate in the US was relatively stable. But beginning in the early 1970s, several decades of "tough on crime" policies contributed to a dramatic rise in incarceration. Today, despite recent declines, the US incarceration rate remains a global outlier. This system of mass incarceration is particularly hostile to Black Americans, who have been imprisoned in stunningly disproportionate numbers. Amid burgeoning interest in scholarship on criminal justice, this special issue examines social science research on the state of mass incarceration in the US: its origin and expansion, its far-reaching effects on families and communities, and why the public tolerates and encourages it. Tracing the system's roots back to slavery, researchers examine the interplay between incarceration, labor demand, and racial domination in the labor market. As criminal justice infrastructure has grown more costly and vast, the system has extracted wealth from poor communities that it preys upon to fiscally survive. This ever-expanding web of incarceration entangles extraordinary numbers of people of all racial groups, with close to half of all Americans having a spouse or coparent, parent, sibling, or child that is or has been incarcerated. To support such a system, many Americans psychologically deny that structural racism is at the heart of criminal justice. Government responses to social justice protests often ignore root social causes and possible remedies and instead rely on policing. Also, law enforcement increasingly draws upon commercial technologies that challenge public oversight and democratic policing. Research on these topics is critical to reveal how we got here, as well as to inform and inspire change.

*Associate Editor, *Science* (June 2018 to August 2021); Current email: trai@ucsd.edu

ILLUSTRATION: ADRIAN BRANDON

**FROM THE ILLUSTRATOR:** In this apartment windows scene, three different families share the same building. Those on the top and bottom floors enjoy their time and freedom together, whereas the scene in the middle set of windows, blocked by jail bars, represents how the family with incarcerated loved ones is affected in its own form of isolation. Unlike the others, people in this family are not able to enjoy the luxury of the outdoors, which suggests the hardships that they endure. This illustration confronts freedom and the darkness that looms over families affected by the incarceration system. —*Adrian Brandon*

PERSPECTIVE

# Policing social unrest and collective violence

## Policies too often ignore root socioeconomic causes of "riots"

PHOTO: BETTMANN/GETTY IMAGES

*By* **Elizabeth Hinton**

The videos of police brutality in the United States that have been widely distributed across social media platforms in recent years gave rise to a mass movement for racial justice that has taken both nonviolent and violent forms. Government responses to this wave of protest point to discouraging linkages to the genesis of national crime control programs in the context of civil rights gains and the "Long, Hot Summers" of the 1960s. In the face of social science research indicating that addressing mass unemployment, failing public schools, and inhumane housing conditions offered an effective policy path to combat social unrest and collective violence, many policy-makers and officials consistently sought to manage the material consequences of socioeconomic problems (e.g., gun violence and drug abuse) with more police, more surveillance, and, eventually, more incarceration. But with growing awareness and recognition of the problem of systemic racism—grounded in a rich body of scientific and historical research that has illuminated the drivers of inequality—a host of social factors have converged to make the present an opportune moment to expect different results. The challenge is to draw from empirical evidence to enact a set of meaningful structural solutions beyond police.

Beginning with Harlem in 1964, federal policy-makers treated the so-called riots as expressions of mass criminality and reacted by simultaneously expanding and militarizing police forces in targeted low-income Black communities. Studies based on aggregate data, historical and economic changes, social relations, and legal structures—from the federal government's own National Advisory Commission on Civil Disorders (known popularly as the Kerner Commission) to the work of scholars such as Robert Fogelson, Morris Janowitz, and Louis Masotti—warned that escalating police force in the face of political and economic grievances was inadequate to solve the problem of racial inequality and consistently presented alternatives. Officials instead drew from the work of sociologist Daniel Patrick Moynihan, criminologist James Q. Wilson, sociologist Edward Banfield, and others who focused on personal and group behavior, and enlisted the police to manage the disorder. Far from deterring urban unrest, the violence—nearly always precipitated by aggressive encounters with police—only increased in frequency and intensity into the 1970s. The United States evolved into a mass incarceration society over the remainder of the 20th century and into the 21st, as gun violence and crime continued to pervade the same neighborhoods that policy-makers targeted with policing and surveillance measures. At best, these punitive policies came at a tremendous cost to taxpayers, with unsatisfactory results overall; at worst, their severe long-term impacts extended not only to the millions of people ensnared in the criminal legal system but also to their families and communities, with immeasurable humanitarian consequences.

From the killing of Michael Brown by local law enforcement in 2014 to George Floyd and Breonna Taylor in 2020, millions of people have taken to the streets demanding racial justice and a radical restructuring of law enforcement under the banner of a new and controversial slogan: "Defund the police." Although 93% of the demonstrations remained peaceful, in some cities, protesters engaged in violent encounters with officers, burned buildings, and looted stores (*1*). The lesson from the fires in Ferguson, Minneapolis, Louisville, and other cities is that violent protest will persist as long as crime control strategies remain at the center of domestic social programs in low-income communities of color.

The enduring impact of the violence of the 1960s and 1970s can be traced to the "War on Crime" that President Lyndon Johnson declared in March 1965, which began the targeting of Black neighborhoods for new policing practices. As the United States fought the Vietnam War, federal policy-makers built a pipeline to deliver surplus army weapons and technologies to local law enforcement to put down the collective violence, allocating $20.6 million to 359 separate programs that modernized police departments over an initial 3-year period (*2*). This marked an unprecedented investment in local law enforcement that grew exponentially over time. The federal allocation for local police forces increased 2900% during the Johnson administration, from zero dollars in 1964 to $10 million in 1965, $20.6 million in 1966, $63 million in 1968, $100 million in 1969, and $300 million in 1970 (*3*–*5*). This investment supported routine stop and frisk, the breaking up of community gatherings, the presence of armed, uniformed officers in the hallways of under-resourced public schools, and other practices that attacked people's dignity. Although such strategies helped suppress mass violence as a regular phenomenon, ironically, they also made further riots inevitable. These strategies remain at the center of domestic urban policies today.

In a televised address to the nation during the July 1967 unrest in Detroit, President Johnson himself acknowledged that "the only genuine, long-range solution for what happened lies in an attack—mounted at every level—upon the conditions that breed despair and violence" (*6*). Johnson's rhetoric indicated that he preferred social programs as the means of attack, but, in practice, he increasingly embraced law enforcement as a short-term solution to the collective violence. Conservatives had championed a tough police response to Black protest through the postwar period, but many liberals started to support Johnson's similar stance in the second half of the 1960s as authorities confronted unprecedented violence in the Watts section of Los Angeles (August 1965), Newark and Detroit (July 1967), and some 137 cities after the assassination of Martin Luther King Jr. (April 1968), alongside the rollout of civil rights legislation and new community action, job training, and remedial education programs.

Department of History, Department of African American Studies, and Yale Law School, Yale University, New Haven, CT, USA. Email: elizabeth.hinton@yale.edu

Protesters, seen through a paddy wagon window, sing and pray during a protest at the jail in Birmingham, Alabama, on 5 May 1963.

Johnson and other liberals never seriously asked whether the War on Crime, launched 1 year after the War on Poverty, may have exacerbated the violence. They believed that crime-war programs—of surveillance and militarized police in communities of color, as well as the increasing professionalization of officers through training and technology—provided the best strategies to handle the crisis moving forward. The Johnson administration's short-term solution to riots became the long-term reality.

Aggressive policing often tends to incite violence, especially when people are protesting the very thing to which they are then subjected. As such, urban unrest peaked only after Johnson's Omnibus Crime Control and Safe Streets Act went into effect in June 1968, which marked the extreme escalation of the War on Crime. Between the enactment of the Safe Streets Act and 1972, some 960 segregated Black communities across the United States witnessed 1949 separate uprisings (in contrast to the 329 total rebellions between 1964 and the end of May 1968) (*7*). In nearly every American city, small or large, where Black people lived in segregated, unequal conditions, residents confronted police with rocks and bottles, shot at them with rifles, plundered local stores, smashed the windows of businesses and institutions, and hurled firebombs.

Although this era of collective violence cannot be separated from the rise of militant civil rights protest and the Black Power movement during the second half of the 1960s, the assassination of King in April 1968 marked a turning point for the mainstream civil rights movement and its emphasis on nonviolence. A generation of young people had watched the direct-action strategies of the civil rights movement unfold with little concrete changes to their everyday living conditions amid the launch of the War on Crime and its new programs of social control. When it appeared that all other options that challenged exploitation, political exclusion, and criminalization had been exhausted, the response by this rising generation was not more nonviolence, but violence (*8*).

Black communities had long been subject to targeted surveillance, frequent encounters with police, mass arrests, illegal searches, and outright brutality, but after the Safe Streets Act, residents in segregated, low-income neighborhoods in big cities such as Baltimore, midsized cities such as Phoenix, and smaller cities such as Waterloo, Iowa, would be patrolled by police departments with veritable arsenals at their disposal. Riot-control training, military-grade weapons such as AR15 rifles and M4 carbines, steel helmets, 3-foot batons, masks, armored vehicles, tear gas—these and other techniques, weapons, and tools flowed into small cities and major metropolitan centers alike as a direct result of the legislation. With its additional $400 million outlay (about $3 billion today) for crime control, the 1968 law allowed cities to flood with police those urban areas that seemed prone to rioting (*9*). By 1970, federal policy-makers had allocated some $40 million worth (about $300 million today) of military-grade equipment for local law enforcement (*2*).

The decision to respond to collective violence—triggered by police but rooted in larger socioeconomic inequalities—with more police was not a forgone conclusion. For one, authorities failed to truly listen to the residents of the rebelling communities, who explained to law enforcement officials, politicians, reporters, and researchers how collective violence could be prevented in the future. What happened in Charleston, West Virginia, in August 1968 provides a telling example. Around 500 residents of the Black section of the city, known as the Triangle District, gathered on the Donnally Street playground, and several police cars "with frightening-looking riot guns," as the *Charleston Gazette* described them, arrived to monitor the situation (*10*). Residents were protesting police brutality and the broader system that had entrenched unequal conditions and anti-Black violence over generations. "We're tired of living in shacks," declared a young person. "When you've got four years of college and a white guy with a grade school education gets a better job than you, it makes you mad," a protestor exclaimed (*11*). On the basis of this testimony alone, officials could have legitimately concluded that law enforcement did not offer an adequate solution, yet as residents continued to air their grievances, two dozen police officers moved through the playground armed with shotguns to herd everyone out of the area. Protestors started throwing rocks, bottles, and cans, and the police, in turn, hauled a few of the young people away (*12*). Meanwhile, the community's larger calls for equality and greater inclusion in political and economic institutions went unanswered (*13*).

Even if officials were not prepared to take seriously the issues of police brutality, substandard housing, unemployment and underemployment, white supremacist violence, and racial discrimination as articulated by Black residents themselves, social science research backed up their demands. The Kerner Commission recognized the legitimate political roots of the collective violence and proposed a set of structural interventions to address fundamental inequalities. The commission had been established by executive order in July 1967 during the summer of rebellions in some 72 cities and was given 7 months to determine the origins of the violence and to recommend measures to contain future disturbances. During this time, the commission used a mixed-methods approach to conduct its investigative and analytical work. Members heard testimony from municipal, state, and federal officials, social workers, law enforcement authorities, academics, civil rights leaders, and "ghetto residents" over the course of 20 days of closed hearings; visited a number of cities that had suffered from serious incidents of collective violence; convened field research teams that interviewed more than 1200 people in 20 separate cities; conducted major surveys of

"Negro and white attitudes" with the support of the Ford Foundation; and analyzed census and local-level data on demographic trends.

In its final report of February 1968 (which went on to become a national mass-market bestseller), the commission warned federal policy-makers and the nation that absent a massive investment into poor Black communities outside of existing War on Poverty programs, Black rebellion and "white retaliation" would entrench racial inequality as a permanent feature of American life (*14*). Federal officials did not heed these warnings; Johnson himself refused to publicly comment on the Kerner Commission's report.

From the insights of the young people who confronted police in Charleston to the Kerner Commission, the socioeconomic roots of community violence were clearly identified, yet policy-makers still threw police at the problem. Over time, authorities have funneled billions of taxpayer dollars into the War on Crime, the War on Drugs, and the prison system. At the time of this writing, President Biden's fiscal year 2022 budget requests $388 million in funding to hire additional police officers, and members of Congress have proposed a budget resolution that would block federal funds to any municipality that defunds its police department, as well as a measure to fund the hiring of 100,000 officers. In effect, many national policy-makers responded to an unprecedented mass movement to de-escalate police force by attempting to do the very opposite.

Such an approach fails to reckon with the wisdom that King prophetically offered toward the end of his life: that only "social justice and progress are the absolute guarantors of riot prevention. There is no other answer. Constructive social change will bring certain tranquility; evasions will merely encourage turmoil" (*15*). Instead of contending with underlying causes and building policies around the needs of communities, the United States has built them around controlling communities and, at the same time, has erected the largest prison system in the world to warehouse those who exist at the margins.

In the face of the best available social science research, and at great fiscal costs, policy-makers at all levels continue to resist investing in alternative measures that could strengthen communities and improve safety overall. For instance, data indicate that children who participate in early childhood education programs are less likely to be arrested as teenagers. Yet funding for Head Start, the most visible remnant of the War on Poverty and a program that helps the nation's poorest children prepare for school between the ages of three and five, continues to shrink. Similarly, research has consistently shown that expanding job opportunities can reduce arrest rates and violence. The federal government spent about $2.7 billion on employment programs for low-income youth in 2020, while investing more than double that number, or $6.1 billion, on grants and assistance to local law enforcement, on top of the $115 billion state and local governments already spent on police that year. In addition to new investments in preschool programs and job-creation measures, mental health treatment and college scholarships would make for a safer society and a stronger democracy, as would establishing a justice system based on the principle of repair instead of retribution.

From the fires of the immediate post–civil rights period to our own time, the history of violent political rebellion underscores the need for public policies that will finally attend to the data and address the impact of structural racism and economic inequality instead of policing the consequences of these dynamics. This approach requires building a society where policing and incarceration are no longer the default responses to violence and other problems in low-income communities of color and broadening definitions of public safety to include access to adequate food, health care, housing, education, and employment for all. ■

## REFERENCES AND NOTES


1. Armed Conflict Location and Event Data Project (ACLED), Demonstrations and Political Violence in America: New Data for Summer 2020; https://acleddata.com/2020/09/03/demonstrations-political-violence-in-america-new-data-for-summer-2020/.
2. S. Schrader, *Badges without Borders: How Global Counterinsurgency Transformed American Policing* (Univ. of California Press, 2019).
3. "Granting Activity 1965–1966," p. 42, "Volume IV, Law Enforcement Assistance Part V, Law Enforcement Assistance—Narrative History [1 of 2]," Box 3, LBJ Library.
4. The Congress of the United States, Congressional Budget Office, "Federal Law Enforcement Assistance: Alternative Approaches" (US Government Printing Office, 1978); www.cbo.gov/sites/default/files/95th-congress-1977-1978/reports/78-cbo-032.pdf.
5. US Department of Justice, Office of Justice Programs, 50 Years of Building Solutions, Supporting Communities and Advancing Justice (2020); www.ojp.gov/ojp50/1968-beginning.
6. L. B. Johnson, Speech to the Nation on Civil Disorders, 27 July 1967; https://millercenter.org/the-presidency/presidential-speeches/july-27-1967-speech-nation-civil-disorders.
7. Radical Information Project, Uprisings/Disturbances Data; https://radicalinformationproject.weebly.com/uprisingsdisturbancesriots.html.
8. E. Hinton, *America on Fire: The Untold History of Police Violence and Black Rebellion Since the 1960s* (Liveright, 2021).
9. Omnibus Crime Control and Safe Streets Act of 1968, Pub. L. No. 90-351, 82 Stat. 197 (1968).
10. J. A. Haught, "Mood of the Triangle: The frustration begins to boil," *Charleston Gazette*, 8 August 1968, p. 29.
11. "Disturbance is repeated in Triangle," *Charleston Gazette*, 7 August 1968, p. 1.
12. T. Sawyer, "Policemen calm Triangle unrest," *Charleston Daily Mail*, 7 August 1968, p. 1.
13. K. W. Lee, "After two days of rumbling, calm returns with sign of pride," *Charleston Gazette*, 8 August 1968, p. 29.
14. National Advisory Commission on Civil Disorders, "Report of the National Advisory Commission on Civil Disorders" (US Government Printing Office, 1968).
15. M. L. King Jr., *Where Do We Go from Here: Chaos or Community?* (Beacon Press, 2010).


10.1126/science.abi9169

POLICY FORUM

# The corporate shadow in democratic policing

## Technology companies can elude accountability

*By* **Elizabeth E. Joh**

Facial recognition, body cameras, and other digital technologies are increasingly commonplace tools of police departments. These provide police with an increased ability to collect vast stores of information on our movements and habits, both online and in real life. These new forms of investigation have led to successes in ways that might not have been possible with traditional policing. But they also raise the specter of unchecked police surveillance. As a result, privacy and civil liberties groups have raised concerns about how much control we have over the most intimate details of our lives. But there is another, less obvious, challenge to democratic policing from these profound changes. Private companies, not public agencies, are typically responsible for the design, production, and sales of products like body cameras and facial recognition software. That the police neither design nor produce the tools that raise some of the most important questions about privacy and civil liberties today has profound consequences for how we think about democratic policing. Perhaps the central dilemma about this corporate shadow in policing is the tensions that exist between public accountability and private interests when police departments are customers.

Traditional policing assumes that investigators and patrol officers possess expertise based on their training and experience; their tools were once an afterthought for policy-makers. That model is both outdated and inaccurate. Policing technologies, whether software or hardware, can play central roles in criminal investigations. The rapid and widespread adoption of body-worn cameras by police in the past several years provides a good example. Although first proposed as a tool of police account-

School of Law, University of California, Davis, Davis, CA, USA. Email: eejoh@ucdavis.edu

ability after the nationwide protests of 2014 for increased police oversight and racial justice after Michael Brown's fatal shooting in Ferguson, Missouri, body cameras soon proved useful to police as evidence-gathering tools as well. The capabilities of these tools, however, are not static. Facial recognition software, for instance, can be applied to body cameras to turn them into tools of mass surveillance. This has raised new questions about how law enforcement agencies can adopt useful tools responsibly and with appropriate policies.

But of the more than 18,000 law enforcement agencies in the United States (*1*), those with in-house technological expertise are the exception rather than the rule. Most of the new forms of automation technology in policing are developed and owned by private companies like Axon, ShotSpotter, and TrueAllele (*2*). Law enforcement agencies are simply clients and customers. Indeed, the police are not even usually the only customers in the multi-billion-dollar commercial surveillance market.

In individual criminal cases, defendants often seek to find out how the police identified them as suspects. There may have been flaws in the identification process or sufficient doubts about the reliability of how the defendant was identified that suggest that the defendant's constitutional rights may have been violated. When it comes to new investigative methods used by police departments, local communities may be concerned that a proposed approach may have racially discriminatory impacts, may be flawed in design or in execution, or may simply be unjustified given its costs balanced against limited public resources. Journalists, researchers, and public interest organizations may also seek access to new technological means available to police to illuminate new forms of government power and to review scientific claims.

At the same time, companies that produce software or hardware for police clients have their own interests that may be at odds with these public values. Private companies are often wary of divulging too much about their products to gain competitive advantage over rivals. As a consequence, companies may decide to protect their intellectual property and market advantages by invoking trade secret privileges, demanding nondisclosure agreements with customers, and imposing other forms of property protections. These forms of commercial secrecy, common enough outside of the criminal justice system, pose challenges to basic police accountability.

## POLICING BY PRIVATE DESIGN

We might start with the very design of the technologies used by the police. Consider body cameras, now a common tool used by police officers around the United States. If a police department's patrol officers are equipped with body cameras, these are most likely Axon products. Formerly known as Taser International, the Axon corporation has designed not just a body camera for police officers but an entire digital platform: hardware in the form of body cameras, a digital platform that allows police to store the gigabits of video data, and software like automated transcription to apply to that information. Important questions in the design of the new digital ecosystems relied upon heavily by the police, including how and when body cameras begin to record and what types of software should be applied to body camera video, are often left to a company whose priorities are increasing market dominance and satisfying shareholders.

Whether Axon chooses to apply facial recognition to its body camera products is a private decision, made without substantial public input.

Police departments stand as one set of customers, not necessarily integral parts of the research or design of the products that document millions of hours of police-civilian interactions, form the basis of civilian complaints against the police, and support investigations for individual prosecutions. And in the absence of robust regulatory regimes, private companies also become de facto policy-makers.

In 2019, Axon announced that it would not use facial recognition technology in its body camera products (*3*), even as that technology is both capable of adoption in its products and being rapidly adopted in the private sector. Axon's decision is not binding on any other private company, nor even on itself. It can choose to reverse its decision at any time. And although the company stated that its decision was based on the advice of its own technology ethics board, the board's members themselves are subject to nondisclosure agreements about the details of Axon's products (*4*). Accordingly, whether the largest body camera company in the United States chooses to apply facial recognition to its products, and thus provide access to that technology to thousands of police officers around the country, is a resolutely private decision, made without substantial public input.

## PUBLIC ACCESS, PRIVATE DATA

Communities are sometimes in a worse position than the police when it comes to knowing more about privately developed and sold technologies used in policing. First, the public may have no idea that its local police department has adopted a new surveillance technology. In the early 2000s, when police departments began to use cell-site simulators—devices that mimic cell phone towers to identify nearby cell phones—the warrantless use of these devices was kept secret because of nondisclosure agreements. In its application for regulatory approval, the then-dominant manufacturer of these cell-site simulators, the Harris Corporation, asked for confidentiality from the Federal Communications Commission for its product, and those conditions were granted (*5*). Local police departments wanting to use cell-site simulators often found themselves bound by nondisclosure agreements enforced by either the Federal Bureau of Investigation or the company itself. As defense attorneys and journalists uncovered details of the technology, increasing calls for regulation ensued. Today, as a result of this advocacy, the US Department of Justice and a few states impose warrant requirements for the use of cell-site simulators (*6*).

In other cases, a local community may know that its police department has availed itself of a new surveillance technology but otherwise lacks the ability to examine its details because of the company's proprietary claims. Dozens of cities throughout the United States use ShotSpotter, a service that uses sensors and software to detect the sound and location of a gunshot to direct a law enforcement response. When a gun is fired, ShotSpotter's sensors pick up the sound and transmit the information to the company, where its proprietary software determines the approximate location of the gunshot (*7*). That information is sent back to the police department, which then sends its patrol officers to the location.

PHOTO: SEBASTIAN GOLLNOW/PICTURE ALLIANCE/GETTY IMAGES

This technology, in theory, could help the police with the persistent problem of gun violence. Automation of this sort augments the capabilities of the police without requiring more police staffing. But a community or researcher who wants to know more about ShotSpotter—its accuracy and its flaws—may find a dead end. ShotSpotter's contractual arrangements with its police customers provide them with results, and results only. The company claims ownership not just of its proprietary software but also of the data its technology generates (*7*). This means that conventional tools of disclosure, like state public records requests laws, have no purchase on any acoustic gunfire detection system because such systems remain within private hands (*8*).

This veil of secrecy has not gone completely unnoticed. A recent study counted at least 14 local governments that have passed municipal code changes that outline, to varying degrees, how and when their local police departments can purchase and engage in new surveillance technology practices (*9*). Perhaps the most comprehensive example can be found in Oakland, California. In Oakland, both the city council and a newly fashioned privacy advisory commission share responsibility for approving the acquisition or use of new surveillance technologies (*9*). This governance mechanism includes features like a requirement for a surveillance impact report and public hearings on the proposed technology acquisition. Although these are worthy efforts, 14 cities are just a miniscule percentage of all the communities where these new technologies are in use.

## DEFENDANT ACCESS TO ALGORITHMS

We can see a final example of the corporate influence in policing in individual criminal cases. DNA evidence has been used to identify suspects since the 1980s, but until recently, investigators have found it difficult to analyze biological samples that are either degraded, contain the DNA of several people, or are too small for conventional analysis. Thus, for instance, recovery of trace amounts of DNA from a crime scene were of little use in police investigations. The development of "probabilistic genotyping," however, has changed how crime labs evaluate this evidence. This type of genetic analysis uses software to calculate the likelihood of whether DNA within a sample came from a particular person.

One probabilistic genotyping program in particular, TrueAllele, has been used in more than 850 criminal cases over the past 20 years (*10*). Anyone who wishes to examine the algorithm responsible for these calculations, however, will run into the proprietary claims of Cybernetics, the company that developed and owns TrueAllele. The company claims that access to its algorithms would compromise its intellectual property. Thus, journalists, researchers, and even the government labs that rely on TrueAllele's results have been unable to examine the underlying algorithms on which the program is built.

Until recently, the secrecy invoked by Cybergenetics posed an obstacle even to criminal defendants who were identified by TrueAllele's algorithm but found themselves unable to gain access to the program to test its reliability or accuracy. That may be slowly changing. In a handful of recent state criminal cases, judges have ordered Cybergenetics to allow defendants access to TrueAllele's source code, in the same way that defendants can have access to other forensic tests that the government uses to identify individual suspects (*11*). These cases may reveal whether TrueAllele contains bugs or errors that raise questions about how particular defendants were identified. These cases may also eventually reveal more about this particular privately owned program, but they will take time and will unfold on a case-by-case basis.

## MITIGATING CORPORATE SECRECY IN POLICING

The growing reliance on artificial intelligence and automation in policing is here to stay. And although some larger police departments may develop their own in-house technological expertise, most police departments will continue to rely on the private development and ownership of many technological tools. (The adoption of in-house expertise for police technologies would likely introduce new oversight questions as well, rather than just address existing ones.) We can also expect the capabilities and use of these tools to increase, just as they have spread through our economy and society. But in a democracy, policing isn't just about efficiency. As the past year has shown us, democratic policing also requires trust and legitimacy. Protests held around the country after George Floyd's death were not calls for more technological tools. If anything, this national attention to policing reminds us that policing remains a social institution. These reform movements face hurdles when private companies shield the tools of policing from public scrutiny.

There is no single answer for how to move forward responsibly in policing. As the examples here show, each type of corporate secrecy poses its own specific set of challenges. Communities and individuals subjected to these policing technologies deserve transparency about how their police use these tools, whether there are potential flaws, and, of course, whether these tools are worth using at all. And although companies do have justifiable concerns about their intellectual property, invocations of proprietary information cannot become an all-purpose shield against public accountability when their customers are police departments.

The fissures in this wall of secrecy—whether through local government oversight, individual criminal litigation, or public pressure—are promising but piecemeal. Some states and local governments have moved forward in setting responsible limits on new police technologies. But states and the federal government can and should take much more regulatory action in setting limits on how much companies can demand secrecy for these policing tools. Relying on court cases is an unpredictable and slow method of change. Thoughtful legislation at the national level that considers the increasing incursions into our privacy rights while acknowledging that police also need tools is urgently needed. Such legislation should be future-proofed as well, that is, not fixated on a technological tool, narrowly construed, that may well seem antiquated in a few years time. Whatever the path forward, one thing is clear: We now lack the right balance between the use of innovative policing tools developed by the private sector but overseen by public accountability and transparency practices. ■

### REFERENCES AND NOTES


1. Bureau of Justice Statistics, "National sources of law enforcement employment data" (NCJ 249681, US Department of Justice, 2016).
2. R. Wexler, *Stanford Law Rev.* **70**, 1343 (2018).
3. R. Smith, "The future of face matching at Axon and AI ethics board report," 27 June 2019; www.axon.com/company/news/ai-ethics-board-report.
4. J. Schuppe, "Axon's police body cams could be getting an AI upgrade. Civil rights groups are worried," NBCnews.com, 26 April 2018; www.nbcnews.com/news/us-news/axon-s-police-body-cams-could-be-getting-ai-upgrade-n869071.
5. E. E. Joh, *N.Y.U. L. Rev. Online* **101**, 106 (2017).
6. US Department of Justice, "Justice Department announces enhanced policy for use of cell-site simulators," 3 September 2015; www.justice.gov/opa/pr/justice-department-announces-enhanced-policy-use-cell-site-simulators.
7. S. Gordon, "Milwaukee committed to ShotSpotter but outcomes, data remain elusive," Wisconsin Public Radio, 11 May 2016; www.wpr.org/milwaukee-committed-shotspotter-outcomes-data-remain-elusive.
8. A. Gatens, J. Reichert, "Police technology: Acoustic gunshot detection systems" (Report, Illinois Criminal Justice Information Authority, 2019).
9. M. Fidler, *Santa Clara High Technol. Law J.* **36**, 481 (2020).
10. L. Kirchner, "Powerful DNA software used in hundreds of criminal cases faces new scrutiny," *The Markup*, 9 March 2021; https://themarkup.org/news/2021/03/09/powerful-dna-software-used-in-hundreds-of-criminal-cases-faces-new-scrutiny.
11. J. Jouvenal, "A secret algorithm is transforming DNA evidence. This defendant could be the first to scrutinize it," *Washington Post*, 13 July 2021; www.washingtonpost.com/local/legal-issues/trueallele-software-dna-courts/2021/07/12/66d27c44-6c9d-11eb-9f80-3d7646ce1bc0_story.html.


10.1126/science.abi9150

REVIEW

# Assessing mass incarceration's effects on families

Hedwig Lee[1]* and Christopher Wildeman[2,3]*

In this Review, we assess how mass incarceration, a monumental American policy experiment, has affected families over the past five decades. We reach four conclusions. First, family member incarceration is now common for American families. Second, individuals who will eventually have a family member incarcerated are worse off than those who never will, even before the incarceration takes place. Third, family member incarceration has negative effects on families above and beyond these preexisting disadvantages. And finally, policy interventions that address the precursors to family member incarceration and seek to minimize family member incarceration would best enhance family well-being. If the goal is to help all American families thrive, then the importance of simultaneous changes in social and criminal justice policy cannot be overstated.

In a little more than a year, in 2023, mass incarceration reaches a major milestone: its 50th birthday (*1*). The degree to which mass incarceration has transformed the lives of American men—and especially African American men with little education living in poor neighborhoods—during this time can hardly be overstated. For African American men who did not finish high school and are approaching midlife, incarceration at some point in their lives is a modal outcome, with upward of 70% having been to prison (*2*). Incarceration is also consequential, and a large literature catalogs the myriad damages that men face after release as a result of doing time (*3*); and this is to say nothing of what they experience while they are confined to a cell. As such, most contemporary research on the direct effects of incarceration on individuals underestimates the toll that mass incarceration has taken in terms of human suffering.

In this Review, we explore how incarceration affects not the men for whom this event has become so common but their families. We see this shift in focus as important for four reasons. First, the family members of the incarcerated have rarely (if ever) been involved in the crimes that their incarcerated family members have committed and, as a result, are the collateral damage of the criminal justice system in a very real and tangible way. Second, many of them did not choose to have a criminally active or justice-involved family member, making the harms caused by the incarceration of a family member all the more problematic. Third, and as we have argued before, the indirect consequences of mass incarceration, experienced by family members, are likely more sizable than those for the men who experience incarceration (*4*, *5*). Finally, the interests of individuals who experience incarceration and their families may, simply put, be misaligned in some instances, a possibility that complicates policy in ways that merely looking at the average effects of family member incarceration on individuals and families may not allow us to see.

In considering these effects, we proceed in three phases. In the first, we present new data showing how common family member incarceration is, how much more common it is for recent cohorts, and how unequally distributed it is. In the second, we highlight how poorly, on average, families were faring even before experiencing the incarceration of a family member. We do so to make clear that what many of these families need even prior to (or, better still, instead of) having a family member incarcerated is expanded social services and economic supports. In the third stage, we present recent evidence regarding the effects of family member incarceration. Here, we highlight not only how family member incarceration appears to do more harm than good but also that there are some situations in which family member incarceration may have some benefits for family life. We conclude with a discussion of limitations of existing research, the need for better data in this area, and some preliminary policy suggestions.

## Crushingly common, shockingly unequally distributed

Until mass incarceration came into being in the early 1970s, prison and jail incarceration were so uncommon that while they were tragic outcomes for individuals and those tied to them, their broader social importance would have simply paled in comparison to other institutional contacts. This could hardly be less the case now. In Fig. 1, we use the Family History of Incarceration Survey (FamHIS), which we designed in collaboration with a number of our colleagues and have described elsewhere (*6*), to show the percentage of adults (by race or ethnicity) from three large birth cohorts (ages 18–39, 40–59, and 60+) who have ever experienced immediate family member (defined as biological, step, or foster parents, siblings, or children, or current romantic partners or anyone the respondent ever had a child with) incarceration generally and the incarceration of a parent, sibling, or child specifically. The confidence intervals for these estimates are quite large and often overlap; nonetheless, we see these data as appropriate for highlighting the level of family member incarceration before the onset of mass incarceration and how that level has changed as mass incarceration has grown to maturity.

Figure 1 highlights the penetration of American life by family member prison and jail incarceration in three ways. First, as the leftmost panel shows, immediate family member incarceration is common regardless of age, with nearly one-half of the youngest cohort (ages 18–39) (48.2%) and over two-fifths of the oldest cohort (ages 60+) (40.8%) having experienced this event. Second, although this may seem counterintuitive initially given that only younger cohorts would have been exposed to mass incarceration for most of their lives, as the "Parent" and "Child" panels indicate, the similar risk of family member incarceration for these different cohorts is largely driven by the fact that very few 18- to 39-year-olds have children who are old enough to have experienced incarceration, essentially yielding the risk of experiencing this type of family member incarceration nil for them. Finally, although immediate family member incarceration is ubiquitous across racial or ethnic groups and birth cohorts, this event is especially common for non-Hispanic Black Americans, with the lowest cohort risk hovering just under three-fifths (59.4%).

## Deep disadvantage often precedes family member incarceration

Stratification in the risk of family member incarceration across a host of facets of American inequality—race or ethnicity, level of educational attainment, and neighborhood context, to name just three—as well as additional factors such as a history of mental health problems, addiction disorders, and prior criminal justice contact, is pronounced (*2*, *7*–*10*). Figure 2 uses data from the Fragile Families and Child Wellbeing Study to show how families that will experience parental incarceration (paternal only, maternal only, or both) in the coming years differ from those that will not on a range of factors that reflect both existing disadvantages families are exposed to and existing disadvantages that can possibly be amplified by family member incarceration.

Figure 2 shows that, across a range of indicators, both mothers and fathers are disproportionately struggling prior to incarceration in ways that are likely consequential both for their immediate family unit and for their kin networks. These families are often in dire need, whether it is in ways that are tied solely to material resources and opportunities or

[1]Department of Sociology, Washington University, St. Louis, MO 63130, USA. [2]Department of Sociology, Duke University, Durham, NC 27708, USA. [3]ROCKWOOL Foundation Research Unit, 1472 Copenhagen, Denmark.
*Corresponding author. Email: hedwig.lee@wustl.edu (H.L.); christopher.wildeman@duke.edu (C.W.)

more broadly to the poor mental and behavioral health that so often springs from these disadvantages. And as the rich ethnographic literature in this area shows, moreover, the risks that families that eventually have contact with prisons and jails faced in advance extend far beyond those that we are able to highlight with the data we had in hand for Fig. 2 (*11*, *12*). As such, these are families that do often need some form of state intervention even before the incarceration of a loved one.

### (Usually) doing more harm than good

For mass incarceration to exacerbate inequality among families, it must be unequally distributed and do harm. If family member incarceration merely reflects existing disadvantage, it is unlikely to exacerbate inequaltity much, if at all. But this is not the case; across a range of indicators, we find that family member incarceration has negative effects on family well-being above and beyond existing disadvantages prior to incarceration. Because our Review must be brief, we focus on how the incarceration of men shapes family and community life. Debates about the effects of the incarceration of women on families and communities remain strident (*13*–*16*), and although rates of female incarceration are increasing, men continue to make up the largest segment of the incarcerated population (i.e., 93.2% of individuals in prison are men) (*17*).

Qualitative research highlights how difficult it is to exhaustively document how incarceration shapes the structure and function of families (*11*, *12*). Yet there are at least three things we can say about how incarceration shapes this broader family system. First, there are strong signs that incarceration affects family structure. Currently incarcerated men marry at extremely low rates, and rates of union dissolution among currently and formerly incarcerated men are markedly higher than those of other men (*18*). Second, there are strong indications that the incarceration of a father lowers paternal prosocial family engagement (*19*) (although this effect is obviously limited to men involved in family life before incarceration in some way), increases maternal harsh parenting (*20*), and increases maternal depression (*21*), all of which indicate decreased quality of family life. Third, there is simply no getting around the fact that paternal incarceration increases the economic hardship that families face (*22*, *23*) and that mothers are unable to offset these hardships solely through their own earnings (*24*). Put simply, despite the many nuances of family life (*11*, *12*), there are strong signs that men's incarceration, on average, disrupts men's involvement in families and leaves those families in greater financial need than they were before. Of course, there may be moderating circumstances—the effects may differ if the father was struggling with mental health or addiction issues or engaging in domestic violence—but the averages suggest that family life is more often impoverished than not as a result of incarceration.

These broader effects then reverberate to individual family members. There is probably no area of greater consensus in the field than regarding the effects of paternal incarceration on children's behavioral and mental health problems. Although no one dataset is ideally suited to consider these effects, and obstacles to causal inference remain, research using a range of datasets ties paternal incarceration to greater behavioral and mental health problems in children (*25*–*27*). Although there is evidence for both externalizing and internalizing behavioral problems, evidence regarding externalizing and physically aggressive behaviors tends to be the strongest; it also tends to suggest that male children are more affected in this domain than female children.

These behavioral and mental health indicators are tightly linked with school readiness, and, as a result, it is perhaps unsurprising that children with incarcerated fathers are less prepared to enter school than their peers (*28*). Yet children of incarcerated fathers face

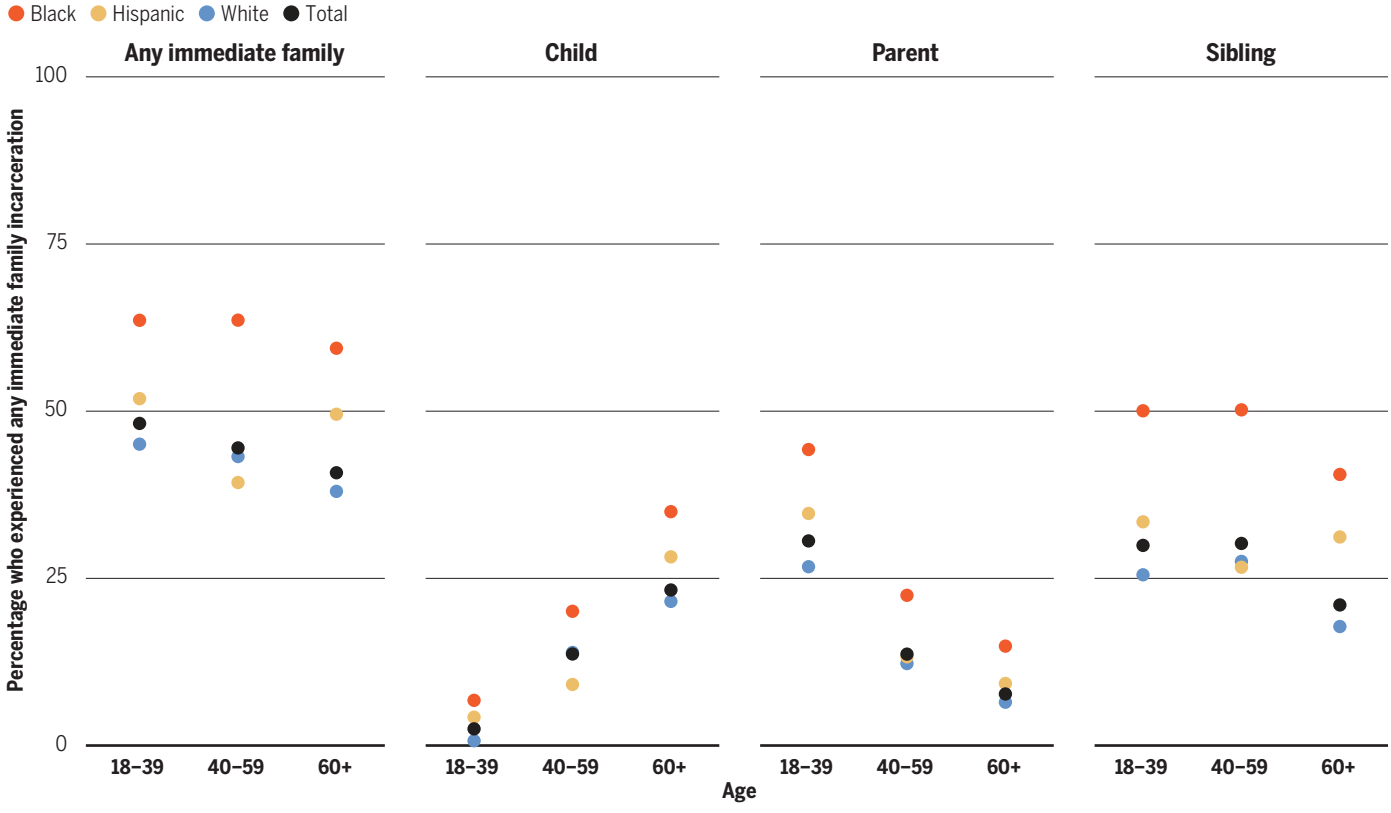



**Fig. 1. The prevalence of immediate family member incarceration in the US population.** Percentage of respondents who have ever had any immediate family members or a child, parent, or sibling, specifically, incarcerated by age (18–39, 40–59, and 60+ years old) and race or ethnicity (non-Hispanic Black, Hispanic, non-Hispanic white), per the 2018 Family History of Incarceration Survey (FamHIS). To download the FamHIS data, see https://ropercenter.cornell.edu/ipoll/study/31115615. For further information and the FamHIS data, see (*6*).

even more impediments to thriving in educational institutions after they set foot in the classroom. Having been incarcerated promotes withdrawal from educational institutions on the part of their parents (*29*). In conjunction with the stigma attached to paternal incarceration in the eyes of teachers (*30*), this leaves children of incarcerated fathers in triple jeopardy in schools—unprepared, with disengaged parents, and carrying the stigma of paternal incarceration. Given the strikes against them, it is no great surprise that children with incarcerated fathers are more likely to be held back in school (*31*), have lower cognitive functioning in middle childhood than their peers (*32*), and perform more poorly on a range of indicators of school success in high school (*33*).

While most disciplines have been intrigued by the broader social effects of mass incarceration on children that we just covered, criminologists have been laser-focused on the intergenerational transmission of crime and criminal justice contact (*34*). The associational evidence on the intergenerational transmission of physical aggression in childhood (*35*), delinquency and criminal behavior in adolescence and early adulthood (*36*, *37*), and criminal justice system contact is quite strong (*38*, *39*). Yet in this area, some of the measurement issues faced throughout the field are even more pressing. Current data generally measure either parental crime or parental criminal justice contact but rarely both (*34*), leaving it very much open for debate whether paternal incarceration is causally linked to children's criminal activity and criminal justice contact.

Yet men's incarceration does not affect just children. An emerging vein of research shows how harmful family member incarceration is to the women who are left behind to manage the fallout (*40*). Maybe the best example of these findings has to do with the effects of a son's incarceration on mothers. Qualitative research in poor communities has long documented the strain that a son's prison or jail incarceration places on mothers (*41*), with recent quantitative research also highlighting how a son's incarceration could increase the parenting burdens placed on grandparents (*42*). The results from two recent analyses are sobering: Mothers who had a child, almost always a son, incarcerated struggle mightily when it comes to a range of indicators of health, including but not limited to self-rated health, depression, and health limitations (*43*, *44*).

In Fig. 3, we highlight the core findings from one of these studies, which considers maternal physical limitations (*44*), a core predictor of mortality, indicator of general health, and factor shaping other domains of thriving (*45*). As Fig. 3 shows, even before experiencing a son's incarceration, women who would eventually experience a son's incarceration were in worse health than their peers, with 16% experiencing one or more physical limitations compared with only 12% of other mothers. By the end of the study, however, after some mothers had experienced a son's incarceration and others had not, this gap grew to a whopping 20% (46% compared with 26%). Within-person estimates, which provide a stronger basis for causal claims, indicate that having a son incarcerated increases mothers' risk of having a physical limitation fully five percentage points.

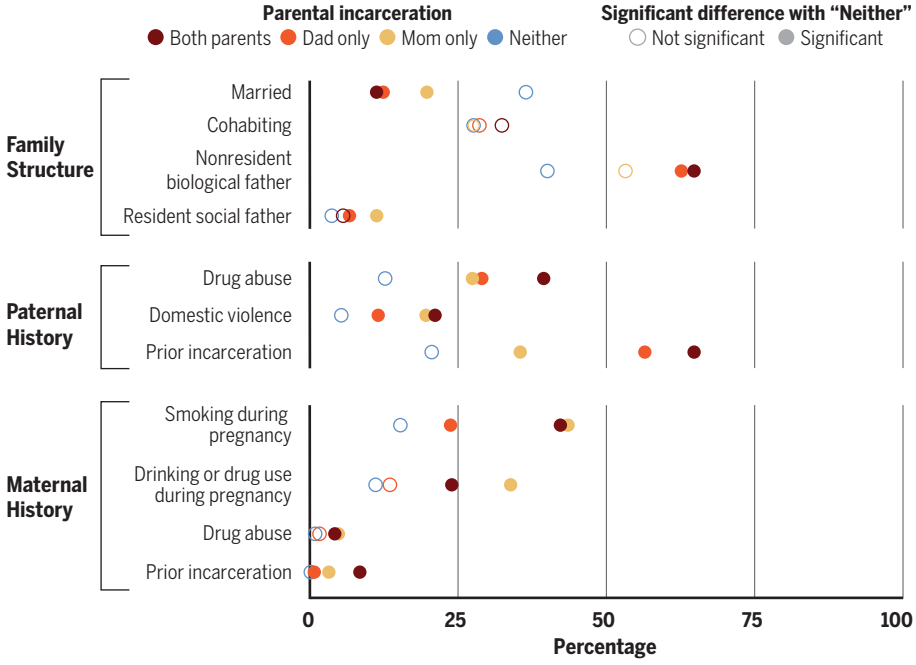



**Fig. 2. Differences between families that will eventually experience parental incarceration and those that will not.** Family structure and paternal or maternal mental and behavioral health and prior prison or jail incarceration among families with 1-year-old children who will not experience paternal or maternal incarceration, will experience only paternal incarceration, will experience only maternal incarceration, or will experience both paternal and maternal incarceration in the next two survey waves of the Fragile Families and Child Wellbeing Study. The figure is based on data from table 3 in (*16*).

Taken together, these studies suggest that family member incarceration may send generational ripples both down (to their children) and up (to their mothers). Because of the heavily unequally distributed risks of family member incarceration by race and class (*6*, *10*), these results indicate very real likely consequences of mass incarceration for family and health inequality.

Although negative effects of incarceration on families abound, there are also reasons to expect that some families benefit from incarceration—at least in the short-term—if the individual who experienced incarceration was struggling with addiction or mental health problems or endangering the safety of his family members through abuse or neglect (*8*). At least for some outcomes, the negative effects of paternal incarceration are concentrated in families where the father was not reported to engage in domestic violence (*35*). Ethnographic research on the perverse benefits of incarceration for families provides nuance to this flat statistical relationship by showing that in many cases, families that benefitted, or merely breathed a brief sigh of relief, when a family member was incarcerated do so not because the family member was actually incarcerated but because they were simply out of the picture briefly. As such, even when families might appear to benefit from incarceration, it is more appropriate to think not of incarceration as beneficial but of incarceration as beneficial in a society largely bereft of high-quality (in- and outpatient) services for addiction, mental health problems, and myriad other troubles.

Relatedly, we lack data to accurately assess the relationship between offense type and impacts on the family. In other words, the data we have leave us in the dark regarding whether the removal of individuals convicted of "low-level" offenses (e.g., drug possession) is detrimental to families or whether the removal of individuals convicted of "high-level" offenses (e.g., armed robbery) is beneficial to families (*46*). These are relationships that are often assumed but remain untested in the construction of policies aimed at creating alternative sentencing practices for caregivers.

Before moving on, it is worth thinking through how the spatial concentration of family member incarceration and its negative effects have the potential to ripple through entire neighborhoods. Because of historical and contemporary forms of structural racism, racially segregated neighborhoods with limited social and economic

resources and myriad environmental hazards and exposures are the same neighborhoods that are overpoliced and from which men of color are disproportionately imprisoned (*47*–*49*). If communities have a critical mass of families struggling because of family member incarceration while also not feeling protected by the same criminal legal system that is meant to protect them and/or not supported by the social service system that is meant to help them with their structurally induced stressors, how are those neighborhoods and the families that reside in them to function? Measuring the spillovers of mass incarceration to communities through families is immensely difficult, but it is hard to believe that such effects do not exist.

The COVID-19 pandemic provides a distinctive example of how family member incarceration affects communities, especially the most vulnerable ones. A growing body of research suggests that prisons and jails have served as vectors for COVID-19 spread. Conditions of confinement including, but not limited to, lack of proper ventilation, inability to properly social distance, and insufficient access to sanitation items lead to the rampant spread of COVID-19 among those in prison and jail (*50*). Individuals working or visiting prisons, often family members (prior to the halting of visitation), exposed to COVID-19 then spread this risk outside of prison to family and community members. And while decarceration has been used as a short-term policy to reduce COVID-19 spread in many areas, COVID-19 risk can still be spread to families that are unable to quarantine from family members returning from jail or prison owing to a combination of limited space at home and a lack of resources for alternative accommodations.

When decarceration has not been used to minimize the risk of COVID-19 infection, some of the other core strategies—greater use of solitary confinement, enhanced restriction of movement for incarcerated individuals to the point of resembling solitary confinement, and complete cessation of in-person visitation—show just how grim conditions of confinement must become in correctional facilities in order to minimize the spread of infections both between incarcerated persons and between incarcerated persons and their communities. Put more directly, for prisons and jails to not spread infections both internally and externally, they tend to become even more brutalizing institutions, with repercussive effects on incarcerated individuals and families alike.

## Concluding thoughts

In this Review, on the occasion of mass incarceration's 50th birthday, we have assessed how it has treated American families. Our assessment is that mass incarceration has largely negative consequences for families, which, because of how common family member incarceration is and how unequally distributed it is, has likely led to greater inequality between families in the United States.

Although we have tried to present a straightforward summary of the effects of incarceration on families, the reality is that this entire literature is grasping in the dark with the data available. None of the data currently used can provide a strong causal test that is also substantively informative. The few studies that purport to provide strong causal tests in the United States context raise at least as many questions as they answer, and it would be at best foolhardy to design any policies around them (*51*). Before we move on to our final conclusions, we will spend a moment addressing what we think we do and do not know about causal relationships in this field, as well as what our current perspectives might miss.

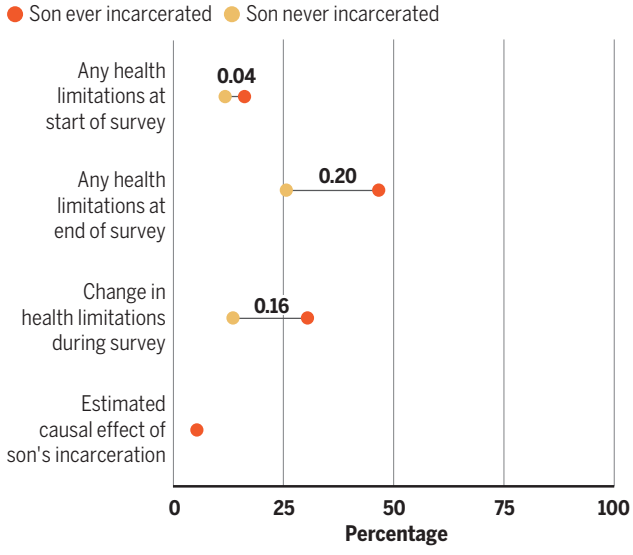



**Fig. 3. The effects of a son's incarceration on maternal health limitations.** Maternal health limitations, changes in health limitations by son's incarceration, and the causal effect of a son's incarceration on maternal health limitations, based on data from the 1979 National Longitudinal Survey of Youth. The figure uses data from tables 1 and 2 in (*44*).

Maybe most importantly, data limitations notwithstanding, there are areas of research where the causal effects of incarceration on individuals, families, and society are clear, and we need to start actively considering interventions, alternatives, and in some cases elimination of certain criminal justice policies and practices. Important examples include the inefficiency and costliness of long prison sentences, supermax facilities, and long-term solitary confinement, all of which cost the public great sums of money and either hold individuals in captivity for far longer than makes sense from a sociological or criminological perspective (in the case of long sentences) or returns them to the community, if ever, worse than when they left it (in the case of supermax facilities and long-term solitary confinement). The COVID-19 pandemic also laid bare the inhumanity of solitary confinement, when individuals who tested positive for COVID-19, many sick before and sicker still from the virus, were put in solitary confinement, a place where being in more than a few days leads to psychological trauma (*52*). So although our instincts as scientists suggest caution, there are some areas where causal effects are clear.

Yet unconvincing as the causal evidence is in some areas and convincing as it is in others, a hyperfocus on causal identification limits scientific advancement in two core ways. First, as we discussed earlier, a large body of ethnographic research has helped us to understand the many nuanced ways in which family member incarceration affects the well-being of family members. We must use insights derived from diverse methodological approaches to inform policies and practices that can protect and promote the health of those incarcerated and the loved ones and communities to which they are connected. In short, the methodological pluralism that breeds insight for both good science and good policy relegates the causal hammer to being just one of the many tools in our methodological belt. Second, both family member incarceration (the "treatment" in the studies we consider) and little to no positive intervention into family life or broader community supports (the "control" in virtually all studies we consider) are suboptimal. Incarcerating someone is costly, saves less in terms of crime reduction than it once did (*53*), and, at least according to much of the research we have reviewed, does harm to families. But doing nothing for an individual struggling with some combination of chronic unemployment, poor mental health, alcohol or drug addiction, or poor social relationships and interactions is also bad for families. As a result, we should think of establishing causal effects of incarceration not relative to no positive interventions but instead relative to interventions that simultaneously seek to enhance the quality of family life and to minimize the risks of further criminal activity.

In closing this Review, we aim to do three things. First and, from a research perspective,

most importantly, to plead for new data collection efforts focused exclusively, or at least primarily, on understanding how families come to experience incarceration, how they weather the time a family member is apart and reacclimating to family life, and how they fare as a result of this experience. In a perfect world, such a data collection would include both excellent measurement and a series of informative comparison groups such as families where a family member was found guilty of a crime and sentenced to something other than incarceration or where a family member never came to the attention of the criminal legal system but was given wraparound services to attend to struggles that could both imperil families and eventually lead to contact with the criminal legal system. Absent a major investment, we are, to be frank, at a loss in this area. We can consider a couple more outcomes or a couple of different types of criminal legal system contact, but we cannot provide much beyond that. This literature will very rapidly stall out absent investment.

Second, and on a more policy-focused note, it is worth considering, on the basis of what we gleaned in this Review, what family-friendly criminal justice interventions would look like. They would have at least three features: (i) They would focus more on diversion in combination with high-quality services in order to be as family-friendly as possible. By "diversion" we mean any interaction with an actor in the criminal legal system (e.g., a police officer, a prosecutor, a judge) that could have led to an arrest, jail incarceration, or prison incarceration but instead led to some lower-level sanction that did not involve confinement (e.g., a ticket, community service, treatment for addiction or other underlying mental health conditions, or some other approved program). Such policies have the benefit of not putting families in a situation where the choices are essentially either incarceration or chaos. (ii) When prison or jail incarceration is needed—and we fully acknowledge that some may bristle at the idea that it is ever needed—family-friendly policies would move beyond considering how to facilitate visitation and how to provide mentors for children to broad social interventions that treat the whole family. Research on the family lives of women with incarcerated partners, for instance, are replete with stories of how women struggle to juggle the number of demands on their time. Why not, instead of focusing on something women left to foot the criminal justice bill have not asked for, focus on something they have—time and money. High-quality childcare centers at free or no charge offer a mechanism through which the harms of incarceration for families could be lessened; more extensive economic supports, especially universal basic income, offer another mechanism for diminishing the effects of family member incarceration. (iii) There is broad consensus among criminologists that the time immediately before and after release is a sensitive one—that old habits, frustrations due to the difficulties of reentry, and frayed family ties may lead to recidivism. Moreover, economic stressors may be magnified when families already in economically precarious situations must now support returning family members who themselves struggle to find jobs and reintegrate. In some cases, rules and regulations may preclude family members with prior convictions from living in their family homes (*53*). In other cases, housing security is so tenuous that family members may be one flush away from eviction (*22*, *54*). Yet researchers and policy-makers generally do little to intervene on the behalf of families during this critical juncture, and truly family-friendly policies would seek to do just that.

Third, in a country that overincarcerates—and disproportionately so among poor Black, Latinx, and Native American communities—we must also seriously consider alternatives to incarceration, which we know was implicit in our first two points but bears mentioning on its own. We need state and county systems to partner with researchers to determine which types of changes (diversion, in-facility programs, postincarceration programs, etc.) best offset the costs of incarceration, with emphasis placed on costs borne by individuals, families, and communities. Moreover, alternative sentencing practices cannot be relegated to those convicted of nonviolent offenses. There is, to be as blunt as possible, no way to drastically shrink the imprisonment rate without cutting sentences for individuals convicted of violent crimes and convicted multiple times (*55*). Those sentences are simply so long that they keep our total imprisonment rate astronomically high, maintaining America's exceptionalism in mass incarceration. A focus on alternative sentencing practices solely for those who have committed a nonviolent offense ignores that there is little to no research suggesting that those who are convicted of different offenses vary in their importance to their families. The little research that exists suggests that that is hardly the case (*56*).

On the eve of its 50th birthday, let us do something revolutionary for mass incarceration: Let us invest in data that allows us to fully understand its effects on families and consider and implement both criminal justice and broader social policies and practices that prevent any future damage. Fifty years of harm is more than American families can be expected to endure.

## REFERENCES AND NOTES


1. D. Garland, *Punishm. Soc.* **3**, 5–7 (2001).
2. B. Pettit, B. Western, *Am. Sociol. Rev.* **69**, 151–169 (2004).
3. D. S. Kirk, S. Wakefield, *Annu. Rev. Criminol.* **1**, 171–194 (2018).
4. C. Wildeman, H. Lee, *Annu. Rev. Sociol.* **47**, 543–565 (2021).
5. S. Wakefield, C. J. Wildeman, *Children of the Prison Boom: Mass Incarceration and the Future of American Inequality* (Oxford Univ. Press, 2013).
6. P. K. Enns *et al.*, *Socius* **5**, 10.1177/2378023119829332 (2019).
7. R. J. Sampson, C. Loeffler, *Daedalus* **139**, 20–31 (2010).
8. R. J. Sampson, *Criminol. Public Policy* **10**, 819–828 (2011).
9. C. Wildeman, *Demography* **46**, 265–280 (2009).
10. H. Lee, T. McCormick, M. T. Hicken, C. Wildeman, *Du Bois Rev.* **12**, 269–282 (2015).
11. M. Comfort, *Doing Time Together: Love and Family in the Shadow of the Prison* (Univ. of Chicago Press, 2008).
12. M. Comfort, *Ann. Am. Acad. Pol. Soc. Sci.* **665**, 63–79 (2016).
13. J. Poehlmann-Tynan, K. Turney, *Child Dev. Perspect.* **15**, 3–11 (2021).
14. D. E. Roberts, *UCLA Law Rev.* **59**, 1474 (2012).
15. J. A. Siegel, *Disrupted Childhoods: Children of Women in Prison* (Rutgers Univ. Press, 2011).
16. C. Wildeman, K. Turney, *Demography* **51**, 1041–1068 (2014).
17. Federal Bureau of Prisons (BOP), BOP Statistics: Inmate Gender (2021); www.bop.gov/about/statistics/statistics_inmate_gender.jsp.
18. L. M. Lopoo, B. Western, *J. Marriage Fam.* **67**, 721–734 (2005).
19. K. Turney, C. Wildeman, *Am. Sociol. Rev.* **78**, 949–979 (2013).
20. K. Turney, *Soc. Forces* **92**, 1607–1636 (2014).
21. C. Wildeman, J. Schnittker, K. Turney, *Am. Sociol. Rev.* **77**, 216–243 (2012).
22. O. Schwartz-Soicher, A. Geller, I. Garfinkel, *Soc. Serv. Rev.* **85**, 447–473 (2011).
23. C. Wildeman, *Ann. Am. Acad. Pol. Soc. Sci.* **651**, 74–96 (2014).
24. A. Bruns, *J. Marriage Fam.* **79**, 1331–1352 (2017).
25. A. Geller, C. E. Cooper, I. Garfinkel, O. Schwartz-Soicher, R. B. Mincy, *Demography* **49**, 49–76 (2012).
26. S. Wakefield, C. Wildeman, *Criminol. Public Policy* **10**, 791–792 (2011).
27. K. Turney, *J. Health Soc. Behav.* **55**, 302–319 (2014).
28. A. R. Haskins, *Sociol. Sci.* **1**, 141–158 (2014).
29. A. R. Haskins, W. C. Jacobsen, *Am. Sociol. Rev.* **82**, 657–684 (2017).
30. C. J. Wildeman, A. R. Haskins, J. Poehlmann-Tynan, Eds., *When Parents Are Incarcerated: Interdisciplinary Research and Interventions to Support Children* (American Psychological Association, 2018).
31. K. Turney, A. R. Haskins, *Sociol. Educ.* **87**, 241–258 (2014).
32. A. R. Haskins, *Soc. Forces* **95**, 861–892 (2016).
33. J. Hagan, H. Foster, *Sociol. Educ.* **85**, 259–286 (2012).
34. C. Wildeman, *Annu. Rev. Criminol.* **3**, 217–244 (2020).
35. C. Wildeman, *Soc. Forces* **89**, 285–309 (2010).
36. J. Murray, D. P. Farrington, *J. Child Psychol. Psychiatry* **46**, 1269–1278 (2005).
37. M. E. Roettger, R. R. Swisher, *Criminology* **49**, 1109–1147 (2011).
38. J. Murray, C.-G. Janson, D. P. Farrington, *Crim. Justice Behav.* **34**, 133–149 (2007).
39. C. Wildeman, S. H. Andersen, *Criminology* **55**, 32–58 (2017).
40. H. Lee, C. Wildeman, E. A. Wang, N. Matusko, J. S. Jackson, *Am. J. Public Health* **104**, 421–427 (2014).
41. D. Braman, *Doing Time on the Outside: Incarceration and Family Life in Urban America* (Univ. of Michigan Press, 2004).
42. K. Turney, *Soc. Forces* **93**, 299–327 (2014c).
43. A. W. Goldman, *J. Health Soc. Behav.* **60**, 398–415 (2019).
44. C. Sirois, *Soc. Sci. Med.* **264**, 113264 (2020).
45. J. N. Houle, D. E. Keene, *J. Epidemiol. Community Health* **69**, 382–387 (2015).
46. S. Wakefield, H. Lee, C. Wildeman, *Ann. Am. Acad. Pol. Soc. Sci.* **665**, 8–21 (2016).
47. S. Herbert, K. Beckett, F. Stuart, *Law Soc. Inq.* **43**, 1491–1513 (2018).
48. B. I. Bryant, P. Mohai, Eds., *Race and the Incidence of Environmental Hazards: A Time for Discourse* (Westview Press, 1992).
49. M. Massoglia, G. Firebaugh, C. Warner, *Am. Sociol. Rev.* **78**, 142–165 (2013).
50. National Academies of Sciences, Engineering, and Medicine, *Decarcerating Correctional Facilities during COVID-19: Advancing Health, Equity, and Safety* (The National Academies Press, 2020).
51. S. Norris, M. Pecenco, J. Weaver, *Am. Econ. Rev.* **111**, 2926–2963 (2021).
52. C. Haney, *Annu. Rev. Criminol.* **1**, 285–310 (2018).
53. D. J. Harding, J. D. Morenoff, C. W. Herbert, *Ann. Am. Acad. Pol. Soc. Sci.* **647**, 214–236 (2013).
54. A. Geller, A. W. Franklin, *J. Marriage Fam.* **76**, 411–427 (2014).
55. J. Forman Jr., *Locking Up Our Own: Crime and Punishment in Black America* (Farrar, Straus and Giroux, ed. 1, 2017).
56. S. Wakefield, K. Powell, *Ann. Am. Acad. Pol. Soc. Sci.* **665**, 195–212 (2016).


## ACKNOWLEDGMENTS


We are especially grateful to C. Covington and S. Sernaker for research assistance and to the anonymous reviewer for helpful feedback. **Funding:** The authors have no funding to declare. **Competing interests:** The authors declare no competing interests.

10.1126/science.abj7777

REVIEW

# Exclusion and exploitation: The incarceration of Black Americans from slavery to the present

Christopher Muller*

Understanding long-run patterns in the incarceration of Black Americans requires integrating the study of racial inequality with the study of political economy. I offer a parsimonious framework describing how the Black incarceration rate has been affected by the dynamics of exploitation and exclusion over time and across space. This framework helps to explain otherwise puzzling facts, like why the Black incarceration rate was lower in the South than in the North for much of the 20th century, why it was lowest in the South's cotton belt, and why it began to tick upward when it did. It also enables us to better understand recent changes in racial and class inequality in incarceration in the United States.

In the early 1970s, the rate of incarceration in the United States began a sharp and unprecedented ascent. In 1970, the US imprisonment rate—excluding jails—stood at roughly 100 per 100,000 people. Forty years later, it had grown five times as large (Fig. 1). As the total imprisonment rate increased, the ratio of the Black and white imprisonment rates remained extraordinarily high. For most of the nearly four decades of prison growth, Black Americans were imprisoned at roughly six times the rate of white Americans (*1*).

These two features of US incarceration—its scale and its concentration among Black Americans—have inspired a large social scientific literature examining incarceration as a form of racial domination. This literature has shown that the study of incarceration in the United States must be situated in a broader history of racial inequality stretching back to slavery. But to understand long-run patterns in the Black incarceration rate from slavery to the present, we need to integrate an analysis of racial inequality with an analysis of political economy (*2*–*6*). Doing so enables us to explain otherwise puzzling facts, like why the Black incarceration rate was lower in the South than in the North for much of the 20th century (Fig. 2), why it was lowest in the South's cotton belt, and why it began to tick upward when it did (Fig. 3). It also helps us to make sense of why racial inequality in imprisonment has narrowed, why class inequality in imprisonment has widened, and why racial inequality exceeds class inequality in people's likelihood of having a family member imprisoned or living in a high-imprisonment neighborhood.

Department of Sociology, University of California, Berkeley, 496 Social Sciences Building, Berkeley, CA 94720, USA.
*Corresponding author. Email: cmuller@berkeley.edu

## Exclusion and exploitation

Consider the three most basic class relations governing the US political economy from the late 19th to the early 21st century: relations between employers, relations between workers, and relations between employers and workers (*7*). Relations between employers are typically defined by competition. In attempting to capture a bigger share of the market, employers are effectively trying to exclude their competitors. One way they do so is by devising ways of producing things using fewer or cheaper workers. Employers' need, under the competitive constraint, to produce more with fewer or cheaper workers is one of the principal reasons why their demand for workers rises and falls.

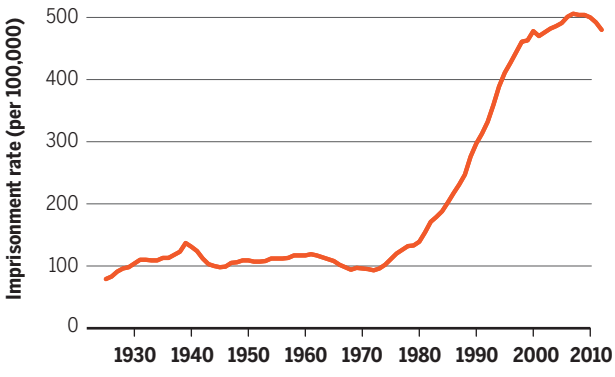



Fig. 1. United States imprisonment rate, 1925 to 2012.

In capitalist economies, workers also face each other as competitors. When jobs are scarce and class solidarity is low, workers hope to exclude other workers from competition. Workers compete as individuals, but throughout US history, they have often been led by racist ideologies to view other workers as members of distinct groups with interests opposed to their own (*8*). Competition over jobs and necessities like housing can entrench racist ideologies, just as racist ideologies can impede solidarity among workers (*9*). Moreover, when workers historically have organized movements that transcend these ideologies, such movements have often been violently repressed (*10*, *11*).

Relations between employers and workers, in contrast, are relations of exploitation: Employers derive direct economic benefits from the labor of workers. Relations of exploitation differ from relations of competition because they are defined by dependence (*12*). When employers' demand for workers is high, they want to exploit—not exclude—workers because they depend on workers' labor.

Black Americans disproportionately occupy subordinate positions within these class relations—a fact that cannot be understood without reference to slavery. Slavery in the United States was a system of exploitation in which slaveholders forcibly extracted labor from people they violently controlled, bought, and sold. Both those who owned enslaved people and those who did not justified this brutal practice by appealing to an ideological rationale holding that the enslaved belonged to a distinct group—a "race"—whose position as dependent agricultural laborers reflected their natural economic, social, and geographical "place" (*13*). Although slavery itself was abolished, its economic and ideological effects have endured. Slavery's regime of forced labor left the majority of Black people after emancipation with little or no wealth. This consigned them to the bottom of the labor market and made them especially vulnerable to fluctuations in the demand for labor—a situation reinforced by racial discrimination in employment.

The basic class relations I have described, together with Black Americans' position within them, bear upon the scale of incarceration in several ways. First, when welfare-state supports are weak, contractions in the demand for labor can leave workers who are expelled from the labor force without means of subsistence. These workers may turn to theft, illegal markets, or other criminalized forms of appropriation to survive. Second, unless employers can exploit the labor of people in prison, when their demand for labor is high, they typically will seek to prevent workers or potential workers from going to prison, thereby lowering the incarceration rate. In periods when employers could affect the incarceration rate, falling labor demand thus could have increased incarceration, particularly among Black Americans, who both occupied the bottom rungs of the labor market and faced especially harsh treatment from police and courts. Finally, competition between workers can exacerbate racist ideologies and lead workers to support punitive measures directed against their perceived competitors.

These claims, it should be stressed, are not transhistorical (*14*). For example, the relationship between labor demand, crime, and incarceration should be weaker when workers have ways to sustain themselves other than through work. When employers can use the labor of people in prison, they may try to increase, rather than decrease, the incarceration rate. And workers can resist—and historically have resisted—the sway of racist ideologies.

Through the mid-20th century, when the demand for Black labor was relatively high, the Black incarceration rate was strongly affected by whether people who exerted influence over incarceration sought to exclude or exploit Black workers. This demand collapsed in the second half of the 20th century, first with the mechanization of cotton harvesting, then with deindustrialization. Young Black men's labor force participation fell and crime rose amid a conservative reaction to both the second Great Migration and the Civil Rights Movement. The law-and-order politics that followed set the stage for an increasingly punitive turn in legislation and prosecution. Today, many of the economic changes that hit Black Americans first are hitting white Americans, whose rate of incarceration is rising.

### Slavery and convict leasing

During slavery, the South's Black incarceration rate was noteworthy primarily for how low it was (*15*). Scholars have long recognized that because slaveholders depended on enslaved people's labor, they had little reason to send enslaved people away to be punished through state prison systems (*3*). What has been less appreciated, however, is the extent to which this same dynamic carried over into the postbellum period.

After the Civil War, many southern states began punishing people by leasing them to private companies. People caught in this system, known as the convict lease system, labored in terrifying conditions and in many cases were literally worked to death. Owing to its brutality and trade in forced labor, the convict lease system has often been compared to slavery (*15*, *16*). Some have taken the argument further, claiming that convict leasing was a functional replacement for slavery (*17*). But two key differences between slavery and convict leasing are hard to square with this latter claim.

First, prior to the Civil War the vast majority of Black people in the South were enslaved. After the Civil War, in contrast, just a small fraction were confined in the convict lease system. In Georgia, for example, enslaved people made up nearly half of the state population in 1860, whereas people in the convict lease system made up less than a tenth of a percent 20 years later. Deprived of landownership and surrounded by laws and violence that restricted their mobility, most rural Black southerners labored instead as tenants, sharecroppers, or agricultural wage workers.

Second, slavery was a predominantly agricultural institution. Most people in the convict lease system, in contrast, were leased to industrial rather than agricultural operations (*18*). As the economic historian Gavin Wright has noted, agricultural workers sent to the convict lease system "were taken away from the area for a long stretch, not returned to the planter as a farm laborer" [(*19*), p. 459]. Thus, like slaveholders, planters had an interest in keeping workers or potential workers out of prison rather than in it.

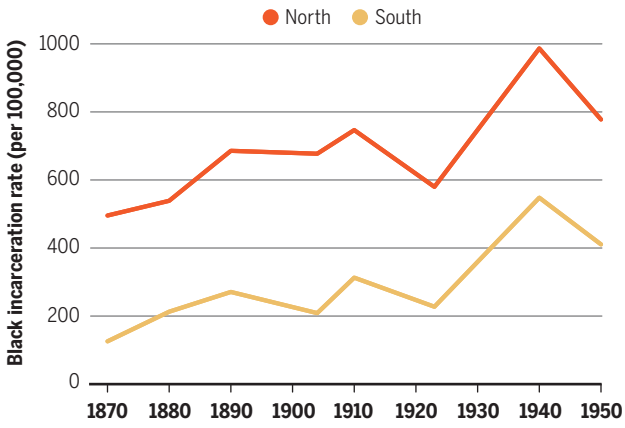



**Fig. 2. Black incarceration rate, 1870-1950.** Northern states include Connecticut, Illinois, Indiana, Maine, Massachusetts, Michigan, New Hampshire, New Jersey, New York, Ohio, Pennsylvania, Rhode Island, Vermont, and Wisconsin. Southern states include Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, Missouri, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, and Virginia. Incarceration data are from (*54*–*59*). The denominators used to construct the incarceration rates are from (*60*).

Some planters held onto accused workers by serving as character witnesses or using their influence over local courts to intervene in prosecutions. Others punished workers themselves, often using violence. But planters also used courts to procure a supply of forced labor. They gathered at courthouses and offered to pay the fines of "any defendants who seemed to be desirable workmen" [(*20*), p. 293]. Then they forced these defendants to work to pay off the debt (*16*, *21*). Defendants faced the impossible dilemma of choosing between the brutality of the convict lease system and the trap of peonage. Because defendants whose fines were paid did not enter state custody, this form of peonage—sometimes called the criminal surety system—diverted them from prison and consequently held the cotton belt's Black incarceration rate down.

Outside of the cotton belt, in contrast, peonage was less prevalent. There, Black defendants were more likely to be viewed as competitors to exclude than as workers to exploit. In the eyes of many white southerners, Black people who resisted dependent agricultural work by moving to cities or acquiring land had "rejected their 'place' in the agrarian social order" [(*18*), p. 71].

Studying the geographic distribution of convict leasing is challenging because census data record only the counties where people were confined, not the counties where they were convicted. However, it is possible to use archival records on where people in the convict lease system were convicted to calculate their risk of incarceration across space. The state of Georgia kept especially comprehensive records of incarceration that can be used for this purpose.

In 1880, Black men in Georgia were least likely to be imprisoned in the cotton belt and most likely to be imprisoned in urban counties or in counties where Black Georgians had accumulated an unusually large amount of land (*22*). This pattern held only for property crimes—those crimes white Georgians had the most discretion to enforce—and did not hold for imprisonment among white men. Notably, the Black incarceration rate for property crimes remained comparatively low in the cotton belt despite historical evidence suggesting that agricultural workers often "chose to commit criminal trespass or petty larceny, in order to dispose of what they saw as their share of the crop without the landlord's interference" (*23*–*25*). In short, Black men faced the lowest risk of imprisonment in counties where people who could influence the incarceration rate sought to exploit them and the highest risk of imprisonment in counties where such people sought to exclude them.

### When labor demand declines

The claim that planters' demand for labor held the cotton belt's Black incarceration rate down implies that declines in agricultural labor demand should have pushed the Black incarceration rate up. As noted above, the demand for Black agricultural workers in the South remained relatively high through the first half of the 20th century. But one notorious episode in the history of southern agriculture has allowed scholars to study the effects of a temporary decline in the demand for agricultural workers.

Beginning in the late 19th century, a beetle called the boll weevil spread northward and eastward from the base of Texas, reaching the East Coast early in the 20th century. Because boll weevils depend on cotton plants at all stages in their life cycle, they have the capacity to destroy

entire fields of cotton. Given the scope of the potential damage, the infestation drew the attention of the US Department of Agriculture, which tracked the weevil's movements in a series of maps it published in regular bulletins.

These maps have allowed scholars to study the boll weevil's effects as it moved, county by county, across the South. Economic historians and sociologists have shown that the infestation substantially reduced cotton yields and the extent of tenant farming (*26*, *27*). It also caused the demand for Black child labor to fall, leading Black children's rate of school enrollment to rise (*28*).

The resulting decline in the need for agricultural workers could have affected the Black incarceration rate in two ways. First, it could have pushed displaced agricultural workers to turn to crimes of survival. To the extent that such workers were caught, prosecuted, and sentenced to prison, this could have increased the incarceration rate. Second, the infestation could have caused planters to reduce the extent to which they paid defendants' fines or otherwise attempted to keep workers or potential workers out of prison. Although convict leasing had been abolished in Georgia in 1908, when the boll weevil arrived in 1915, defendants were still caught between peonage and imprisonment—now in the form of state-run chain gangs (*18*).

The boll weevil infestation was particularly consequential for Black southerners (*14*). Their concentration in cotton production made them especially susceptible to declines in cotton yields, and their low levels of wealth often prevented them from paying fines to avoid chain gangs and peonage (*20*, *21*). Moreover, historical research has shown that Black Americans were far more likely than white Americans to be punished by incarceration when the demand for their labor was low (*15*, *22*, *24*, *29*).

The boll weevil infestation in Georgia increased the Black prison admission rate for property crimes by more than a third (*14*). Its effect on Black prison admissions for homicide, in contrast, was negative and not statistically significant. Because there are no data on crime or peonage in early–20th-century Georgia, it is impossible to definitively determine how much of this effect was driven by increases in crime versus decreases in peonage. However, the infestation only increased prison admissions for crimes that could be punished with a fine—those crimes that allowed planters to entrap defendants in peonage. There is no evidence, in contrast, that it increased admissions for crimes that legally had to be punished with a prison sentence. This fact, together with qualitative evidence from historical sources, suggests that planters' practice of paying defendants' fines fell along with cotton yields. When planters' demand for agricultural workers declined, the Black incarceration rate increased.

## Migration

As the boll weevil completed its path across the southern United States, the first Great Migration of Black southerners to the North was getting underway. This migration had many causes, but two of the most important were war-induced shocks to immigration and the introduction of immigration restrictions in the 1910s and 1920s, both of which cut off the supply of industrial workers from Europe (*30*). With a shortage of laborers, particularly as northern workers were drafted and northern wartime production ramped up, labor agents headed south to recruit Black workers.

Black migrants typically were better off than those who stayed in the South, and they made great gains in income from moving (*30*). Moving north allowed migrants to escape the South's regime of racial terror and labor coercion and offered them social and political opportunities the South had closed off. But migrants also faced resentment and violence in the North as their numbers increased.

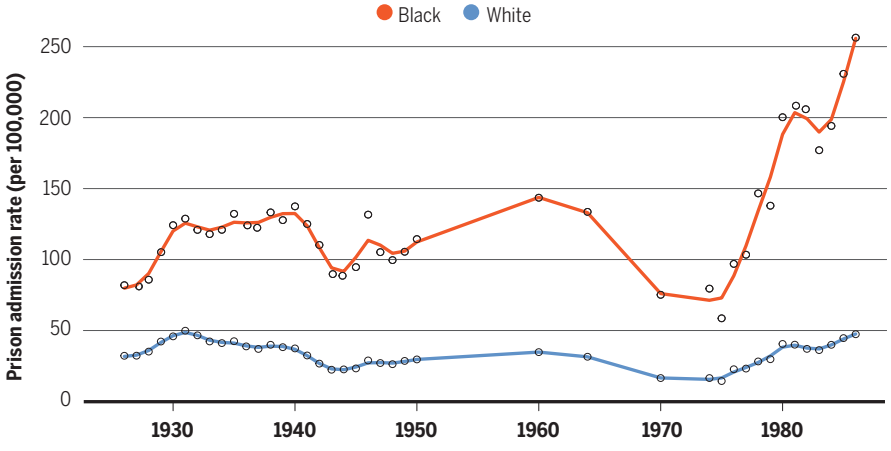



**Fig. 3. Black and white prison admissions, 1926 to 1986.** A smooth line summarizes the trend. Curves were fit to data using smooth-spline in R. Data are from (*61*).

The first Great Migration had two kinds of effects on the Black incarceration rate. First, migrants left a region where planters routinely tried to keep agricultural workers out of prison and in the fields. The absence of a similar practice in the North meant that the Black incarceration rate was higher in the North than in the South even before the migration began (*31*). Census microdata can be used to compare migrants' probabilities of incarceration to those of similar northerners and southerners who stayed in their respective regions. Moving north drastically increased migrants' likelihood of being incarcerated, even though migrants had higher average levels of education than those who remained in the South (*31*, *32*).

Second, the scale of the migration provoked a punitive reaction. Black migrants were forced into competition with northern workers, particularly European immigrants who were both concentrated at the bottom of the northern labor market and overrepresented on northern police forces. Many of these workers blamed migrants for their plight and greeted them with hostility rather than solidarity (*8*). The faster migrants moved into northern states, the sharper was the increase in the Black incarceration rate (*31*). Thus, the migration increased Black Americans' risk of incarceration not just because they left a region where the Black incarceration rate was held down by people trying to exploit them, but also because they entered a region where the Black incarceration rate was pushed up by people trying to exclude them.

Black Americans' earliest experiences with police, courts, and incarceration in both the North and the South engendered a profound sense of estrangement from these institutions (*33*, *34*). As early as 1901, W. E. B. Du Bois argued that the convict lease system had produced a marked decline in Black people's "faith in the integrity of courts and the fairness of juries" [(*15*), p. 742]. Persistent racial inequality in incarceration in both regions also gave rise to an ideological association of "blackness and criminality" that traced high rates of Black incarceration to an imagined group essence rather than to the dynamics of exclusion and exploitation described above (*29*, *35*). Recent work has demonstrated how the combination of legal estrangement and essentialist beliefs about criminality can generate self-reproducing cycles of face-to-face interpersonal violence and state violence (*36*).

## Agricultural mechanization

The causes of mass incarceration are too numerous and varied to summarize succinctly. But an underappreciated contributor to three of the most widely cited causes is the mechanization of cotton harvesting. Between 1950 and 1970, the percentage of US cotton harvested by machine rose from 5% to nearly 100%. Although mechanization had begun earlier in some parts of the

South and was itself, in part, a response to labor shortages, "with the successful breakthrough in mechanical cotton harvesting, the character of the labor market radically changed in the 1950s from 'shortage' to 'surplus'" [(*37*), p. 243]. Against the backdrop of a welfare state weakened by racial exclusion (*4*), this revolution in agricultural production had three important ramifications for the growth of the Black incarceration rate in the late 20th century.

First, mechanization accelerated the decline of young Black men employed in agriculture. In 1940, nearly a third of young Black men held agricultural occupations. By 1970, that figure had fallen below 3% (*38*). Historians have noted that the contraction in Black men's labor force participation "coincided with a stunning rise in their rates of incarceration" [(*39*), p. 82]. The effect of mechanization on crime and incarceration in the South has not been studied directly. But research on the final decades of the 20th century has shown that unemployment increased the property crime rate (*40*) and that decreases in employment among young Black men had a substantial effect on their prison admission rates (*41*). The imprisonment rates of southern cotton-producing states also rose earlier than those of states in other regions (*14*).

Second, the decline in agricultural work caused by mechanization strengthened the ongoing second Great Black Migration (*42*). Like the first, this migration was met with a backlash. Migrants arrived in northern and western cities just as urban jobs began to disappear (*37*, *43*); white residents left cities for suburbs (*42*, *44*); and police targeted poor, predominantly Black, neighborhoods (*44*, *45*). Recent research has shown that the reaction to the second Great Migration increased northern homicide rates, police spending, and Black incarceration rates (*46*).

Finally, some scholars have suggested that mechanization and the migration it inspired helped to create the conditions for the Civil Rights Movement (*30*, *47*). Conservative politicians responded to the urban uprisings and civil disobedience associated with this movement by connecting them to crime and appealing to voters with a promise to restore "law and order" (*44*, *48*). This message found a receptive audience, particularly, although not exclusively (*49*), among white voters invested in the social and economic exclusion that the Civil Rights Movement sought to combat. Crime policy became a more salient issue in electoral contests, and appearing "soft on crime" became a political liability.

Agricultural mechanization thus shifted the balance of efforts to exclude and exploit Black Americans. The demand for low-education Black workers fell as a conservative reaction to the second Great Migration and the Civil Rights movement took hold. This made it more likely that increases in crime—both real and imagined—would receive a punitive response. The prison admission rate shot up in the 1970s (Fig. 3), with most of the growth concentrated among Black men with low levels of education (*2*).

## Incarceration today

In the 21st century, the fall in the demand for low-education workers that first affected Black Americans began to affect white Americans too. This trend is most apparent in rising rates of mortality among white Americans with no college education, driven by deaths from suicide and alcohol and drug use (*50*). But less noticed has been the fact that the prison admission rate of white Americans without a college education rose sharply over the same period. This, along with falling prison admissions among Black Americans with no college education, substantially reduced racial inequality in incarceration for the first time in decades. Class inequality in incarceration, meanwhile, has exploded (*51*).

Despite these recent changes, however, racial and class inequality in incarceration remain tightly intertwined. Although the gap in the prison admission rates of Black and white Americans has narrowed, Black Americans still bear a vastly greater share of the harm of incarceration. This is because enduring structures of racial domination have made class boundaries among Black Americans more permeable than they are among white Americans. Persistent segregation has meant that middle-class Black Americans are much more likely than middle-class white Americans to live in or near poor neighborhoods (*52*). Because of historically low levels of wealth among Black families, middle-class Black Americans are also more likely than comparable white Americans to have poor family members (*53*). As a result, Black Americans with high levels of education and income are more likely than white Americans with low levels of education and income to experience the imprisonment of a family member or to live in a neighborhood with a high imprisonment rate (*51*).

## Conclusion

This is far from the first essay to propose that understanding incarceration in the United States requires integrating the study of racial inequality with the study of political economy [see, e.g., (*2*–*6*)]. But I have aimed to provide a parsimonious framework that describes precisely how the Black incarceration rate has been affected by the dynamics of exploitation and exclusion over time and across space. This framework is deliberately simple: There are changes in incarceration that it alone cannot explain and causes of incarceration that it leaves out. But it nonetheless helps us to make sense of patterns in the Black incarceration rate since slavery that otherwise would be difficult to understand.

Future research should evaluate the explanatory reach of this framework, as well as its limits. For example, the relationship between agricultural labor demand, peonage, and incarceration proposed here should be assessed using data from states other than Georgia. Scholars should also study in closer detail how mechanization and deindustrialization affected the Black incarceration rate. This work will necessitate gathering data from archival sources to fill in gaps in the administrative data currently available.

We are living in a moment of intense debate about how to keep people as free as possible from both state and interpersonal violence. But there is a growing recognition that this goal will be met by creating and expanding alternatives to the penal system. The patterns in the Black incarceration rate I have described were by no means inevitable: At key moments in the struggle for racial and economic justice, they could have been interrupted. Indeed, the framework I have sketched here suggests that striking at the roots of mass incarceration will require working simultaneously to eradicate systems of racial domination and to transform economic relations so that they are no longer predicated on exclusion and exploitation.

## REFERENCES AND NOTES


1. J. Travis, B. Western, S. Redburn, Eds., *The Growth of Incarceration in the United States: Exploring Causes and Consequences* (National Academies Press, 2014).
2. B. Western, *Punishment and Inequality in America* (Russell Sage, 2006).
3. R. W. Gilmore, *Golden Gulag: Prisons, Surplus, Crisis, and Opposition in Globalizing California* (Univ. of California Press, 2007).
4. L. Wacquant, *Punishing the Poor: The Neoliberal Government of Social Insecurity* (Duke Univ. Press, 2009).
5. M. Gottschalk, *Caught: The Prison State and the Lockdown of American Politics* (Princeton Univ. Press, 2015).
6. S. S. Smith, J. Simon, *RSF* **6**, 1–27 (2020).
7. R. Brenner, in *Marxist History-writing for the Twenty-first Century*, C. Wickham, Ed. (Oxford Univ. Press, 2007), pp. 49–111.
8. W. E. B. Du Bois, *Black Reconstruction in America* (Russell & Russell, 1935).
9. K.-Y. Taylor, *From #BlackLivesMatter to Black Liberation* (Haymarket Books, 2016).
10. R. D. G. Kelley, *Hammer and Hoe: Alabama Communists During the Great Depression* (Univ. of North Carolina Press, 1990).
11. A. Gourevitch, *From Slavery to the Cooperative Commonwealth: Labor and Republican Liberty in the Nineteenth Century* (Cambridge Univ. Press, 2015).
12. E. O. Wright, *Class Counts: Comparative Studies in Class Analysis* (Cambridge Univ. Press, 1997).
13. B. J. Fields, *New Left Rev.* **181**, 95–118 (1990).
14. C. Muller, D. Schrage, "The Political Economy of Incarceration in the Cotton South, 1910-1925," IRLE Working Paper no. 105-19, 2021.
15. W. E. B. Du Bois, *Mission. Rev. World* **14**, 737–745 (1901).
16. D. A. Blackmon, *Slavery by Another Name: The Re-enslavement of Black Americans from the Civil War to World War II* (Doubleday, 2008).
17. C. R. Adamson, *Soc. Probl.* **30**, 555–569 (1983).
18. A. Lichtenstein, *Twice The Work of Free Labor: The Political Economy of Convict Labor in the New South* (Verso, 1996).
19. G. Wright, *Georgia Hist. Q.* **81**, 452–464 (1997).
20. A. F. Raper, *Preface to Peasantry: A Tale of Two Black Belt Counties* (Univ. of North Carolina Press, 1936).
21. P. Daniel, *The Shadow of Slavery: Peonage in the South, 1901-1969* (Univ. of Illinois Press, 1972).
22. C. Muller, *Am. J. Sociol.* **124**, 367–405 (2018).
23. A. Lichtenstein, in *Slave Cultures and the Cultures of Slavery*, S. Palmié, Ed. (Univ. of Tennessee Press, 1995), pp. 176–186.
24. W. E. B. Du Bois, *Some Notes on Negro Crime Particularly in Georgia* (Atlanta Univ. Press, 1904).
25. G. D. Jaynes, *Branches Without Roots: Genesis of the Black Working Class in the American South, 1862-1882* (Oxford Univ. Press, 1986).

26. F. Lange, A. L. Olmstead, P. W. Rhode, *J. Econ. Hist.* **69**, 685–718 (2009).
27. D. Bloome, J. J. Feigenbaum, C. Muller, *Demography* **54**, 1029–1049 (2017).
28. R. B. Baker, *J. Econ. Hist.* **75**, 1128–1160 (2015).
29. K. G. Muhammad, *The Condemnation of Blackness: Race, Crime, and the Making of Urban America* (Harvard Univ. Press, 2010).
30. W. J. Collins, *Explor. Econ. Hist.* **80**, 101382 (2021).
31. C. Muller, *Am. J. Sociol.* **118**, 281–326 (2012).
32. K. Eriksson, *J. Econ. Behav. Organ.* **159**, 526–538 (2019).
33. M. C. Bell, *Yale Law J.* **126**, 2054–2150 (2017).
34. C. Muller, D. Schrage, *Ann. Am. Acad. Pol. Soc. Sci.* **651**, 139–158 (2014).
35. A. Y. Davis, in *The Angela Y. Davis Reader*, J. James, Ed. (Blackwell, 1998), pp. 74–95.
36. R. Sethi, in *Difference Without Domination: Pursuing Justice in Diverse Democracies*, D. Allen, R. Somanathan, Eds. (Univ. of Chicago Press, 2020), pp. 93–114 .
37. G. Wright, *Old South, New South: Revolutions in the Southern Economy Since the Civil War* (Basic Books, 1986).
38. C. A. Fitch, S. Ruggles, in *The Ties that Bind: Perspectives on Marriage and Cohabitation*, L. J. Waite, C. Bachrach, M. Hindin, E. Thomson, A. Thornton, Eds. (Aldine de Gruyter, 2000), pp. 59–88.
39. M. B. Katz, M. J. Stern, J. J. Fader, *J. Am. Hist.* **92**, 75–108 (2005).
40. S. Raphael, R. Winter-Ebmer, *J. Law Econ.* **44**, 259–283 (2001).
41. B. Western, M. Kleykamp, J. Rosenfeld, *Soc. Forces* **84**, 2291–2311 (2006).
42. L. P. Boustan, *Competition in the Promised Land: Black Migrants in Northern Cities and Labor Markets* (Princeton Univ. Press, 2017).
43. W. J. Wilson, *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy* (Univ. of Chicago Press, 1987).
44. L. Wacquant, *Daedalus* **139**, 74–90 (2010).
45. E. Hinton, *From the War on Poverty to the War on Crime: The Making of Mass Incarceration in America* (Harvard Univ. Press, 2016).
46. E. Derenoncourt, "Can You Move to Opportunity? Evidence from the Great Migration," Working Paper, Princeton University, 2021.
47. G. Wright, *Sharing the Prize: The Economics of the Civil Rights Revolution in the American South* (Harvard Univ. Press, 2013).
48. M. Alexander, *The New Jim Crow: Mass Incarceration in the Age of Colorblindness* (New Press, 2010).
49. J. Forman Jr., *Locking Up Our Own: Crime and Punishment in Black America* (Farrar, Straus and Giroux, 2017).
50. A. Case, A. Deaton, *Deaths of Despair and the Future of Capitalism* (Princeton Univ. Press, 2020).
51. C. Muller, A. F. Roehrkasse, "Racial and Class Inequality in U.S. Incarceration in the Early Twenty-First Century," IRLE Working Paper no. 109-20, 2021.
52. M. Pattillo, *Annu. Rev. Sociol.* **31**, 305–329 (2005).
53. C. M. Heflin, M. Pattillo, *Soc. Sci. Res.* **35**, 804–822 (2006).
54. U.S. Department of the Interior, Report on crime, pauperism, and benevolence in the United States at the eleventh census: 1890, part II (Government Printing Office, 1895), table 4.
55. US Department of Commerce and Labor, Special Reports: Prisoners and Juvenile Delinquents in Institutions: 1904 (Government Printing Office, 1907), Table 2.
56. US Department of Commerce, Prisoners and Juvenile Delinquents 1910 (Government Printing Office, 1914), Table 3.
57. US Department of Commerce, Prisoners 1923 (Government Printing Office, 1926), Table 139.
58. US Department of Commerce, Sixteenth Census of the United States: 1940, Population, Special Report on Institutional Population 14 Years Old and Over (Government Printing Office, 1943), Table 12.
59. US Department of Commerce, United States Census of Population: 1950, Special Reports: General Characteristics of Families (Government Printing Office, 1955), Table 36.
60. C. Gibson, K. Jung, "Historical Census Statistics on Population Totals by Race, 1790 to 1990, and by Hispanic Origin, 1970 to 1990, for the United States, Regions, Divisions, and States," Division Working Paper no. 56, US Bureau of the Census Population, 2002.
61. United States Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, and Inter-university Consortium for Political and Social Research, Race of Prisoners Admitted to State and Federal Institutions in the United States, 1926-1986, Inter-university Consortium for Political and Social Research [distributor], 1999-10-04.



**ACKNOWLEDGMENTS**

I thank E. Derenoncourt, A. Gourevitch, S. Naidu, A. Rana, R. Sampson, D. Schrage, and L. Wacquant for helpful comments.

10.1126/science.abj7781


REVIEW

# Toward an understanding of structural racism: Implications for criminal justice


**Julian M. Rucker[1]* and Jennifer A. Richeson[2]***



**Racial inequality is a foundational feature of the criminal justice system in the United States. Here we offer a psychological account for how Americans have come to tolerate a system that is so at odds with their professed egalitarian values. We argue that beliefs about the nature of racism—as being solely due to prejudiced individuals rather than structural factors that disadvantage marginalized racial groups—work to uphold racial stratification in the criminal justice system. Although acknowledging structural racism facilitates the perception of and willingness to reduce racial inequality in criminal justice outcomes, many Americans appear willfully ignorant of structural racism in society. We reflect on the role of psychological science in shaping popular understandings of racism and discuss how to contribute more meaningfully to its reduction.**


The criminal justice system in the US has been marred by the presence of profound racial inequality from its inception. Be it runaway slave patrols, which set the stage for contemporary policing; laws that sought to criminalize everyday activities when performed by formerly enslaved individuals after Reconstruction; or the bevy of discriminatory policies that led to an explosion in incarceration, disproportionately affecting Black and Latinx Americans (*1*), racial inequality has been produced and reproduced within the structures of the US criminal legal system. Yet, by and large, Americans not only tolerate these inequalities but also often support policies that exacerbate them. The persistence and broad acceptance of stark levels of racial inequality in the US criminal justice system are decidedly inconsistent with the oft-noted increase in Americans' embrace of racially egalitarian values during the past several decades (*2*). The purpose of the present Review is to shed light on this paradox through the lens of psychological science.

Although many psychological factors likely contribute to this gap between Americans' ostensible embrace of racially egalitarian values and their tolerance of a racially inequitable criminal justice system, the factor we consider here is how Americans think about the nature of racism itself. Do Americans think racism in contemporary society is primarily an interpersonal phenomenon that originates in prejudiced individuals or a structural phenomenon that originates in policies or laws that systematically disadvantage members of marginalized racial groups? We contend that Americans' tolerance of racial inequality in criminal justice outcomes is partly due to a profound, often willful ignorance of the role of structural racism in contemporary society.


[1]Department of Psychology and Neuroscience, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [2]Department of Psychology and Institution for Social and Policy Studies, Yale University, New Haven, CT 06520, USA.
***Corresponding author. Email: julian.rucker@unc.edu (J.M.R.); jennifer.richeson@yale.edu (J.A.R.)**


## The psychology of group-based hierarchy

It is imperative to ground this discussion in social dominance theory (SDT) (*3*), the prevailing psychological theory of societal hierarchy. SDT begins with the observation that group-based hierarchies are ubiquitous among human societies. It asserts that, typically, these hierarchies are created on the basis of age, sex, and at least one other dimension that is more arbitrary yet extremely meaningful in a given cultural context (e.g., race, ethnicity, religion). Irrespective of the dimension, members of dominant groups allocate a disproportionately large share of a society's valued resources (e.g., homes, jobs) to one another, at the expense of members of subordinated groups. In the case of race-based hierarchy in the US, for instance, members of the dominant racial group (e.g., white Americans) have repeatedly allocated positive societal goods (e.g., personal freedom, property rights, educational opportunities, access to health care) to themselves and negative societal goods (e.g., environmental toxins) to members of subordinated racial groups. This pattern of allocation, codified by law, reinforced by cultural norms and practices, and enforced by both official police as well as extralegal, vigilante groups, constitutes structural racism.

Once a hierarchy is established, social stratification is advanced, protected, and exacerbated by systems and processes that operate on multiple levels. At the individual level, for instance, both members of dominant groups and members of subordinated groups become motivated to justify, and thus maintain, the system. Dominant group members are motivated to maintain their higher social status, prestige, and preferential treatment. Members of subordinated groups are also often motivated to maintain the very hierarchical system that disadvantages them because of a greater psychological need to feel that the world is

predictable and controllable [see also system justification theory (*4*)].

Especially pertinent to this Review, SDT argues that criminal legal systems in most societies are thought to serve a hierarchy-maintaining function (*3*). Consistent with this perspective, members of subordinated groups throughout the world [e.g., Black Americans of lower socioeconomic status in the US (*1*), African and Asian individuals in the UK (*5*)] tend to be disproportionately targeted for surveillance and punishment. Moreover, SDT argues that some actors (e.g., prosecutors) embedded within the larger system typically work to maintain, if not enhance, the hierarchy, whereas others (e.g., prosecutors) typically work to attenuate it, eventually resulting in a hierarchy-preserving homeostasis. Ultimately, most members of the broader society come to behave in ways that support the system—consider, for instance, the readiness with which white Americans call the police in response to the unwelcome presence of racial minorities (*6*). Similarly, individuals living in disadvantaged neighborhoods, as a result of concentrated poverty and unemployment, may begin to engage in higher levels of criminal activity, ultimately exacerbating and seemingly justifying the increased police presence.

### Justifying racial hierarchy

According to SDT, at the psychological level, efforts to increase, maintain, or even decrease intergroup hierarchy are influenced by a constellation of values, attitudes, ideologies, and attributions known as legitimizing myths (or beliefs). Hierarchy-attenuating beliefs (e.g., "innocent until proven guilty") can serve as a challenge against societal inequality. By contrast, hierarchy-enhancing beliefs (e.g., negative racial stereotypes) are critical to preserving perceptions of societal fairness, despite the presence of group-based inequality. The endorsement and espousal of hierarchy-enhancing beliefs is not limited to members of dominant societal groups (*4*). For instance, periods of surprisingly widespread support for punitive criminal justice efforts among Black politicians and community members may be understood, at least in part, as a product of the adoption of hierarchy-enhancing beliefs [e.g., pathologizing Black criminality (*7*)]. Hence, hierarchy-enhancing beliefs can blunt efforts by members of racially subordinated groups to challenge the system despite its severe racially disparate outcomes. Notably, a legitimizing belief does not have to be even remotely true to be effective in maintaining the hierarchy; what matters is the extent to which enough people believe it to be true such that it justifies hierarchy-enhancing policies, practices, and outcomes.

For most of US history, when considering racial inequality in criminal justice, American elites (e.g., elected officials, education leaders, scientists) and institutions actively propagated the hierarchy-enhancing myth that Black Americans are inherently more "criminal" than others, be it due to biology or culture, thus justifying the group's unequal treatment and outcomes (*1*, *8*). Over the past several decades, however, belief in the fundamental, hierarchy-attenuating principle of racial equality began to take hold. For example, in the first major national surveys of racial attitudes in the US, conducted in the 1940s, 54% of the white Americans surveyed supported racially segregated public transportation and 54% thought that white Americans (versus Black Americans) should get preferential access to jobs. Only a few decades later (circa 1970s), support for equal access to public transportation and employment was so widespread among white Americans that these items were dropped from subsequent surveys (*2*). Indeed, this rise in egalitarian racial attitudes would seem to mark a shift toward support for a less hierarchical, racially stratified society.

Yet, during at least part of this same time period, the numbers of incarcerated racial minority group members, especially Black Americans, began to rise exponentially (Fig. 1) (*1*). The inmate population of jails and prisons in the US rose from ~300,000 around 1970 to roughly 2 million less than three decades later. Between 1983 and 2000, the zenith of this explosion in mass incarceration, the number of Black American prison inmates increased 27-fold (*1*). In the context of this new, racially egalitarian zeitgeist, the existence of a justice system whose outcomes are systematically influenced by racism would seem intolerable.

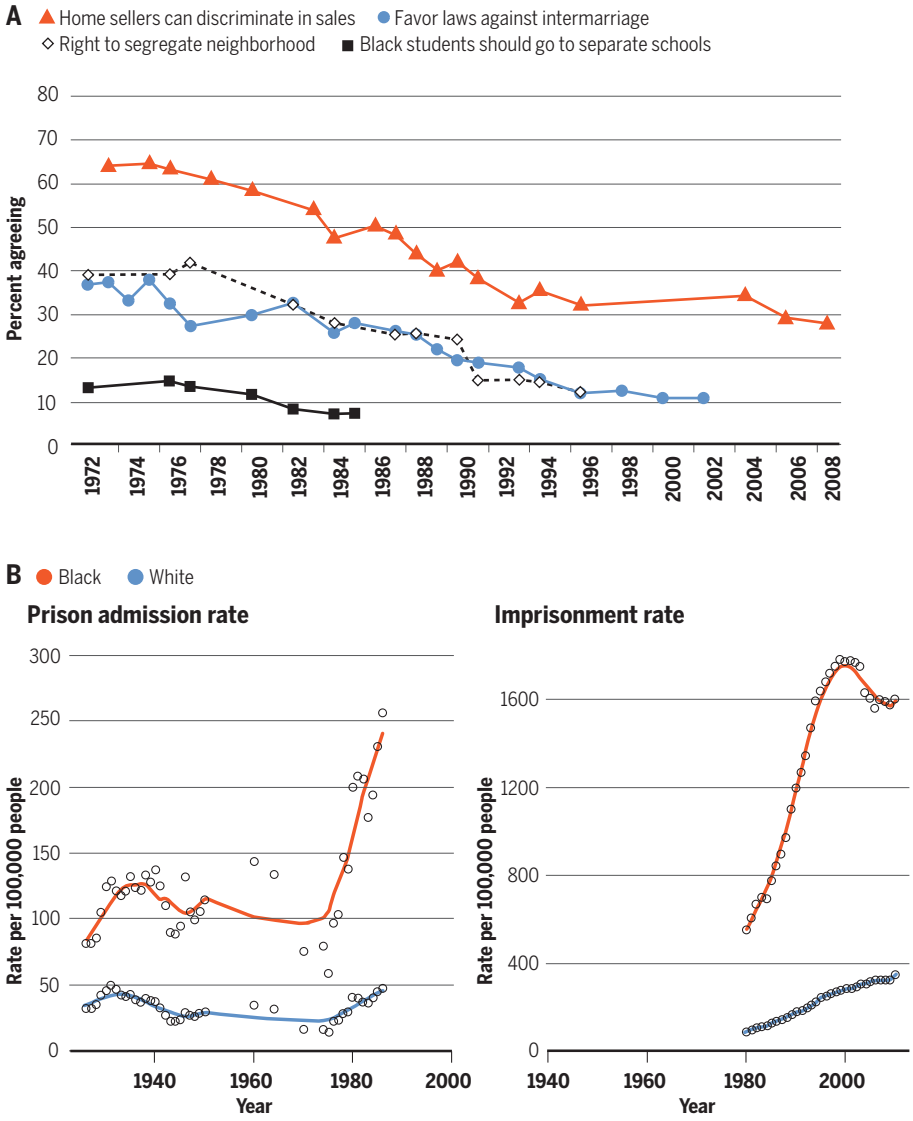



**Fig. 1. Trends in self-reported racial egalitarianism and societal incarceration rates.** (**A**) Declining self-reports of racial prejudice among white Americans from 1972 to 2008. [Republished with permission of Princeton University Press, from (*55*), permission conveyed through Copyright Clearance Center, Inc.] (**B**) Rising prison admission and imprisonment rates for white and Black Americans from 1926 to 2010. [Republished with permission of The National Academies Press, from (*56*), permission conveyed through Copyright Clearance Center, Inc.]

In the wake of this racially stratified criminal justice system, alternative explanations for systemic racial inequalities in carceral outcomes have emerged as a necessity.

We argue that Americans tolerate stark racial inequality in carceral outcomes partly because they fail to appreciate structural racism, focusing instead on the influence of interpersonal racism (e.g., focusing on "a few bad apples" among police officers). Because recognizing the role of structural racism in the criminal justice system would threaten its very existence, Americans are motivated to remain relatively ignorant of the structural racism that is endemic to criminal legal systems. Indeed, as recently as 2016, most Americans believed that racism in contemporary society is primarily an interpersonal problem rather than a structural one (*9*). The outsized endorsement of an interpersonal rather than structural view of racism is more prevalent among white Americans than among racial minority Americans (*10*). Notably, we contend that, as an implication of widespread endorsement of this legitimizing belief (e.g., the denial of structural forms of racism), Americans tend to be less responsive to evidence of structural racism than to evidence of interpersonal racism, both within the criminal justice system and in society more broadly. As a result, the racially stratified institutions in society are preserved. Fig. 2 depicts a conceptual framework to shed light on how hierarchical group rank (dominant versus subordinated) and motivation influence the acquisition of knowledge and understanding of structural racism that, in turn, shape the desire to maintain or dismantle racially inequitable systems and structures.

## Acquiring a structural understanding of racism

### *The role of information*

How do Americans learn about racism? For many, especially members of the dominant racial group, it is primarily through formal educational opportunities or the consumption of mass media, both of which have been found to limit, if not misrepresent, the very types of information that would foster an appreciation for structural racism (*11*–*13*). For others, especially members of racial minority groups, it is through formal education as well, but also partially through direct experience (*14*) coupled with the racial socialization efforts of family and community members (*15*–*17*). Although any of these informational pathways could lead to individuals' acquiring information about racism in general and structural racism in particular, there is reason to believe that these different routes to learning about racism contribute to the racial group differences in endorsement of an interpersonal or structural view of racism. Indeed, Nelson and colleagues (*18*) found that white participants were less likely than Black participants to acknowledge systemic forms of racism, in part because they had less knowledge about racism in US history than their Black counterparts. Similarly, Bonam and colleagues (*19*) found that an educational intervention designed to increase such historical knowledge (i.e., a brief lesson on racial "redlining" in US housing policy) subsequently increased the extent to which participants acknowledged structural forms of racism in contemporary society.

In other words, one reason why white Americans largely fail to recognize the role of structural racism in the criminal justice system—and in general—is the dearth of critical education, and the presence of miseducation, on the topic. This stands in contrast to the direct and vicarious experiences with racial discrimination in criminal justice that racial minority Americans have (*14*), in addition to exposure to historical accounts of mistreatment from parents, grandparents, and other members of their broader community (*15*). It is no wonder, then, that one component of civil rights activism in the 1960s, and again today, aims to increase awareness of racism in US society through education.

Such efforts, however, have always been met with institutional obstacles. For example, growing scrutiny has been placed on media outlets' overreliance on and underscrutinization of police accounts of crime reporting, especially in cases that may involve racial discrimination (*20*). Uncritical acceptance of these official police accounts, coupled with the exclusion (if not erasure) of perspectives of members of marginalized racial groups, leads to biased perceptions of these police incidents among the larger public. Consider, for instance, the demonstrably false official account of the murder of George Floyd submitted to media by the Minneapolis police department (*21*). Had there not been video footage of the event to contradict this misleading account, it is entirely likely that no one would have been held accountable, and the officer's deadly behavior would have been sanctioned.

Similarly, substantial political resistance, such as recent efforts to ban "critical race theory" from US school curricula (*22*), also represents a noteworthy obstacle to increasing a structural understanding of racism in criminal justice and beyond. Such efforts clearly illustrate the role that societal elites and institutions play in shaping not only how Americans learn about racism but whether they learn about it at all. The lack of education on structural racism in criminal justice is neither accidental nor unintentional (*23*). In line with SDT, these efforts likely reflect a motivation among the dominant group to maintain their place atop the status hierarchy by further propagating the denial of structural racism, a powerful legitimizing myth. Hence, it is important to consider the role of motivation as a barrier to the acquisition of a structural understanding of racism.

### *The role of motivation*

Denying structural forms of racism, past and present, serves the broad motivation to protect the group and self-image of white Americans [e.g., social identity threat (*24*)]. People are psychologically motivated to maintain a global, positive self-view and will strive to mitigate threats to their self-image (*25*, *26*). Learning about the racism that is foundational to the nation, especially as perpetrated by white Americans toward Indigenous peoples and racial minorities, can be experienced as a direct threat to the image of one's racial or national identity (*27*, *28*). Further, structural racism can also present a meritocratic threat to white Americans insofar as holding membership in a group that may be unfairly advantaged in society suggests that one's successes may not be solely attributable to one's own hard work, talent, and effort (*29*).

These self-protective motives largely work in opposition to the educational and informational processes that facilitate the acquisition of a structural understanding of societal racism. Consistent with this idea, Adams, Tormala, and O'Brien (*30*) found that affirming the self-concepts of white participants before they were exposed to potential incidents of racial discrimination toward Latinos led them to more readily acknowledge that incidents were discriminatory compared with white participants who had not been affirmed. Notably, a reanalysis of these findings (*31*) revealed that this effect was stronger for incidents of systemic and/or structural discrimination than for more interpersonal incidents, which suggests that the denial of structural racial discrimination, in particular, is a motivated process.

Although these motivations are more common among white Americans, as noted previously, members of racial minority groups can also perceive information about societal racism as threatening (*32*). Further, it is possible for individuals—including those within the criminal legal system—to be especially motivated to understand the systemic and structural bases of racial inequality in society. Just as some actors work to maintain and enhance hierarchical systems, according to SDT, others work to reduce group-based hierarchies (*33*). Such individuals tend to value egalitarianism and have low levels of social dominance orientation (SDO) (i.e., the preference for group-based hierarchy) (*3*). Recent research suggest that individuals with lower levels of SDO are more likely to notice evidence of inequality than their higher-SDO counterparts (*34*).

## Why does a structural understanding of racism matter?

We began this Review by noting the paradox between the rise in Americans' adoption of

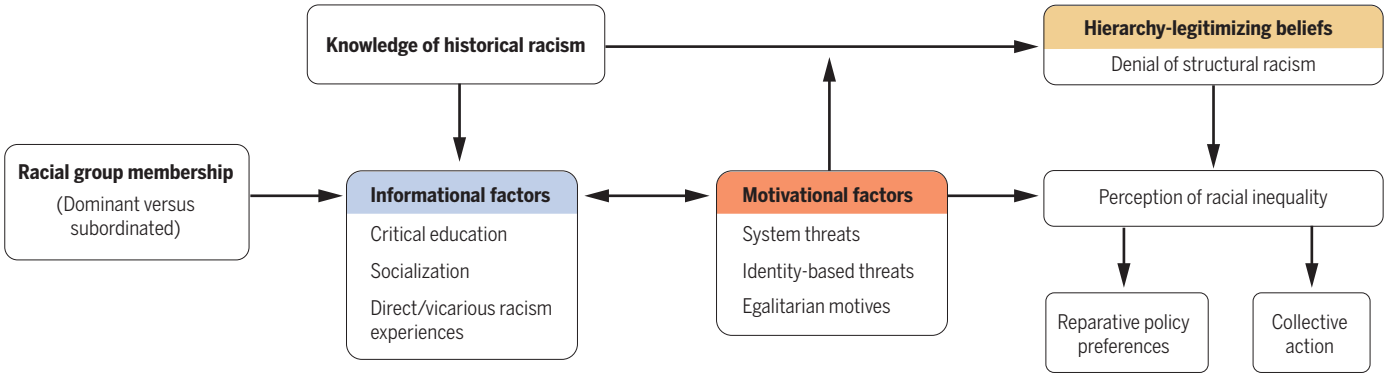



**Fig. 2. A framework of the antecedents and consequences of denial of structural racism.** Dominant versus subordinated racial group membership predicts differential exposure to racism (e.g., direct or vicarious experience, socialization, critical education), which informs individuals' overall level of knowledge about historical and contemporary racism. Group status also shapes individuals' motivations (e.g., self- and/or group esteem, system justification) to seek information about racism and acknowledge (or deny) the existence of structural racism, which in turn predicts willingness to perceive racial inequality and support for reparative policies or collective efforts designed to redress it.

racially egalitarian values and the coincident rise in racial disparities in carceral outcomes (*1*, *35*). One explanation, well-established in the social psychological literature, is the role of widespread implicit and/or automatic associations between young Black men and crime in shaping the views and behaviors of both the police and the public (*36*–*38*). We agree that these and other implicit racialized crime associations play an important role in the maintenance of systemic racial inequality in criminal justice. Indeed, it is structural racism and the inequities it produces that support and sustain these automatic associations between race and crime. In other words, implicit bias is itself a consequence of structural racism (*39*).

Additionally we argue that Americans' ignorance of the role of structural racism in contemporary society contributes to our willingness to tolerate such stark levels of racial inequity in criminal justice outcomes. Consider, for instance, New York City's "stop-and-frisk" program, which is ostensibly designed to get weapons off of the streets and allows the New York City Police Department (NYPD) to stop, question, and search individuals on the basis of "reasonable suspicion." Between 2004 to 2013, the NYPD stopped 4.8 million people, ~80% of whom were Black and Latino. Of those who were stopped during this period, both Black and Latino individuals were more likely to be frisked and less likely to be found with a weapon than white individuals (*40*). In other words, the program produced racially disparate criminal justice outcomes. We contend that it may only be through the acknowledgment of structural racism that these disparities can be accurately perceived and meaningfully addressed.

Consistent with this position, we found that individuals who tend to have a more structural (versus interpersonal) understanding of racism also tend to believe that racial minorities are disadvantaged in the criminal justice system relative to white Americans (*41*). Further, the tendency to endorse the idea that structural rather than interpersonal racism is a more important problem in contemporary society accounts for a substantial amount of the gap between white and Black Americans' perceptions that the criminal justice system is unfair. In other words, to the extent that white Americans understand structural racism, they also tend to perceive societal inequality in a way that is similar to Black Americans' views.

Because structural racism is characterized by policies, practices, and/or laws that have a disparate impact on members of particular racial or ethnic groups, evidence of racially disparate outcomes is the first indicator of its operation. Hence, racial disparities in a domain are likely to be met with extra scrutiny by people with a structural (but not necessarily those with an interpersonal) understanding of racism. Recent work suggests that absent an appreciation of structural racism, white Americans are likely to interpret evidence of racial disparities in incarceration through the lens of stereotypes about racial minority criminality (*36*–*38*, *42*), thereby justifying the inequality. Consistent with this notion, in a longitudinal study conducted shortly after Hurricane Katrina, O'Brien and colleagues (*43*) found that white undergraduate students in New Orleans who held a relatively interpersonal understanding of racism, compared with those who held a structural understanding, were less likely to perceive racism as a cause of the racially disparate relief efforts and outcomes.

However, white Americans with a relatively structural (versus interpersonal) understanding of racism—especially individuals with strong egalitarian attitudes—should be less likely to support punitive criminal justice policies after exposure to evidence of racial disparities in the criminal justice system (*41*). In a recent study, we examined this possibility directly and found that increasing awareness of structural racism (as opposed to increasing awareness of implicit racial biases) through an educational intervention led white Americans to express greater support for policies designed to reduce racial inequality (*10*). Similarly, Adams and colleagues (*11*) found that a multiday tutorial that expanded knowledge of the structural basis of racial discrimination increased support for antiracist policy initiatives (e.g., reparations, race-conscious admissions policies), compared with a similar tutorial that focused exclusively on interpersonal racial bias and with a no-treatment control. Together, this work suggests that a structural understanding of racism may be crucial to motivate efforts to address the stark and obstinate racial disparities in criminal justice, as well as in other societal domains.

In addition to the implications for members of groups atop the racial status hierarchy (e.g., white Americans), research suggests that a structural understanding of racism may have important implications for members of marginalized racial groups. This insight was central to the theorizing of Paulo Freire, who argued that education designed to foster a critical analysis of societal inequity is essential for members of subordinated groups to reclaim their full humanity and resist their oppression (*44*). More recent empirical work has demonstrated the benefits of developing an understanding of structural racism among marginalized youth, finding it to be associated with positive individual outcomes [e.g., improved mental health, greater educational achievement and engagement (*45*)] as well as increased civic engagement (*16*).

In sum, the literature surveyed here suggests that an appreciation of the structural nature of racism predicts not only the extent to which individuals perceive, or perhaps acknowledge, racial stratification in the criminal justice system but also their propensity to support reparative

social policies. Moreover, this work underscores the fact that for members of the dominant group (i.e., white Americans), strong egalitarian motives are insufficient to disrupt the normal psychological processes that result in the justification of racially inequitable criminal justice outcomes. Instead, egalitarian values need to be paired with an appreciation for structural racism. For members of marginalized racial groups, developing a structural analysis of racism may serve as an important buffer against internalizing negative stereotypes about one's group and, ultimately, engaging in behaviors that serve to justify the dominant group's position. Moreover, a structural understanding of racism may be especially vital to engendering collective action to change the system, for both members of advantaged and disadvantaged groups.

### Summary and self-reflection

In this Review, we sought to illustrate key social-psychological factors that shape the maintenance and justification of a racially unjust criminal justice system, despite large-scale support for racially egalitarian values. Psychological motives to substantiate the racial hierarchy and protect one's self-image work against opportunities to increase exposure to critical education on the structural underpinnings of contemporary racial inequality. In essence, ignorance and denial of structural racism protect against an indictment of the legitimacy of the criminal justice system. By contrast, acknowledgment of structural racism in society motivates efforts to reduce racially disparate outcomes. With this framework, it becomes clear that merely holding egalitarian attitudes is insufficient to reform and dismantle systems that reproduce racial inequality—a structural understanding of racism is integral to these objectives.

This work also has implications for psychological science as a field. We, as psychologists, should our role in encouraging, if not promulgating, an interpersonal understanding of racism at the expense of more structural accounts (*11*, *23*). Explicit (or implicit) animus toward or even stereotypes and beliefs about members of marginalized racial groups are troublesome and worthy of comprehensive inquiry, but it is essential that they be properly situated in a structural context (*39, 46*). Moreover, these attitudes and beliefs stem from our culture and our cultural products, be they the holidays we celebrate, the memorials we erect, the social policies we adopt, or the histories we teach. In other words, our collective cultural socialization shapes what we come to believe, consciously or unconsciously, about the racial inequalities we observe, or fail to observe, and our explanations for those inequities (*47*). To the extent that this socialization communicates that some groups are valued more than others or are more capable or suitable for certain tasks than are others, our implicit and explicit attitudes will reflect these contingencies (*48–50*).

Consequently, we join others (*23*, *51*) in calling for a psychological science of racism that carefully and regularly reconnects the processes of mind under examination to the sociocultural context(s) in which they arise. This "mind in context" approach is not novel. Social psychological research has long studied both interpersonal and structural (including cultural) forms of racism (*52–54*). In essence, we call for a return to this type of embedded scholarship. An embedded approach to understanding racism, in criminal justice and beyond, seeks to elucidate how individuals' everyday decisions, practices, and behaviors shape—and how they are shaped by—racially unjust societal structures with or without animus or intention. Without this embedded approach, the contributions of our field are likely to fall short of meaningfully reducing the pervasive racial inequality that exists in multiple domains of American life, including criminal justice. Instead, we are likely to contribute to a framing and (mis)understanding of racism that serves to uphold the societal status quo.

### REFERENCES AND NOTES


1. M. Alexander, *The New Jim Crow: Mass Incarceration in the Age of Colorblindness* (New Press, 2010).
2. L. D. Bobo, in *Racial Trends and Their Consequences*, N. J. Smelser, W. J. Wilson, F. Mitchell, Eds. (National Academies Press, 2001), vol. 1, pp. 264–301
3. J. Sidanius, F. Pratto, *Social Dominance: An Intergroup Theory of Social Hierarchy and Oppression* (Cambridge Univ. Press, 1999).
4. J. T. Jost, M. R. Banaji, B. A. Nosek, *Polit. Psychol.* **25**, 881–919 (2004).
5. D. Lammy, "The Lammy review: An independent review into the treatment of, and outcomes for, Black, Asian and minority ethnic individuals in the criminal justice system" (HMG, 2017); www.gov.uk/government/organisations/lammy-review.
6. V. M. Weaver, "Why white people keep calling the cops on black Americans" (Vox, 2018).
7. J. Forman Jr., *Locking Up Our Own: Crime and Punishment in Black America* (Farrar, Straus and Giroux, 2017).
8. K. G. Muhammad, *The Condemnation of Blackness: Race, Crime, and the Making of Modern Urban America* (Harvard Univ. Press, 2010).
9. Pew Research Center, "2016 Racial Attitudes in America III Survey [Data and Codebook]" (Pew, 2016); www.pewsocialtrends.org/2016/06/27/on-views-of-race-and-inequality-blacks-and-whites-are-worlds-apart/.
10. J. M. Rucker, "Thinking Structurally: The Antecedents and Consequences of a Structural Understanding of Discrimination," thesis, Yale University (2020).
11. G. Adams, V. Edkins, D. Lacka, K. M. Pickett, S. Cheryan, *Basic Appl. Soc. Psych.* **30**, 349–361 (2008).
12. B. D. Behnken, G. D. Smithers, *Racism in American Popular Media: From Aunt Jemima to the Frito Bandito* (Praeger, 2015).
13. G. E. Lopez, P. Gurin, B. A. Nagda, *Polit. Psychol.* **19**, 305–329 (1998).
14. J. Del Toro *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 8261–8268 (2019).
15. N. Anyiwo, J. Bañales, S. J. Rowley, D. C. Watkins, K. Richards-Schuster, *Child Dev. Perspect.* **12**, 165–170 (2018).
16. J. Bañales, C. Mathews, N. Hayat, N. Anyiwo, M. A. Diemer, *Cultur. Divers. Ethnic Minor. Psychol.* **26**, 176–188 (2020).
17. D. Hughes *et al.*, *Dev. Psychol.* **42**, 747–770 (2006).
18. J. C. Nelson, G. Adams, P. S. Salter, *Psychol. Sci.* **24**, 213–218 (2013).
19. C. Bonam, B. Nair Das, B. R. Coleman, P. S. Salter, *Soc. Psychol. Pers. Sci.* **10** (2018).
20. D. Shenkman, K. Slade, "Police reports shouldn't set the news agenda: A guide to avoiding systemic racism in reporting" (American Bar Association, 2021).
21. P. Bump, "How the first statement from Minneapolis police made George Floyd's murder seem like George Floyd's fault," *Washington Post* (2021).
22. M. Iati, "What is critical race theory, and why do Republicans want to ban it in schools?" *Washington Post* (2021).
23. P. S. Salter, G. Adams, M. J. Perez, *Curr. Dir. Psychol. Sci.* **27**, 150–155 (2018).
24. N. R. Branscombe, N. Ellemers, R. Spears, B. Doosje, in *Social Identity: Context, Commitment, Content*, N. Ellemers, R. Spears, B. Doosje, Eds. (Blackwell, 1999), pp. 35–58.
25. C. M. Steele, in *Advances in Experimental Social Psychology*, L. Berkowitz, Ed. (Academic Press, 1988), vol. 21, pp. 261–302.
26. H. Tajfel, J. C. Turner, in *Psychology of Intergroup Relation*, S. Worchel, W.G. Austin, Eds. (Hall Publishers, 1986), pp. 7–24.
27. A. M. Miron, N. R. Branscombe, M. T. Schmitt, *Gr. Pro. Inter. Rel.* **9**, 163–180 (2006).
28. K. N. Rotella, J. A. Richeson, *Soc. Psychol. Personal. Sci.* **4**, 730–737 (2013).
29. E. D. Knowles, B. S. Lowery, R. M. Chow, M. M. Unzueta, *Perspect. Psychol. Sci.* **9**, 594–609 (2014).
30. G. Adams, T. T. Tormala, L. T. O'Brien, *J. Exp. Soc. Psychol.* **42**, 616–626 (2006).
31. G. Adams, L. T. O'Brien, J. C. Nelson, *Ana. Soc. Iss. Pub. Pol.* **6**, 215–235 (2006).
32. J. Crocker, B. Major, C. M. Steele, in *The Handbook of Social Psychology*, D. T. Gilbert, S. T. Fiske, G. Lindzey, Eds. (McGraw-Hill 1998), pp. 504–553.
33. A. K. Ho, N. S. Kteily, *Curr. Opin. Psychol.* **35**, 108–113 (2020).
34. N. S. Kteily, J. Sheehy-Skeffington, A. K. Ho, *J. Pers. Soc. Psychol.* **112**, 136–159 (2017).
35. E. Hinton, *From the War on Poverty to the War on Crime: The Making of Mass Incarceration in America* (Harvard Univ. Press, 2016).
36. J. L. Eberhardt, P. G. Davies, V. J. Purdie-Vaughns, S. L. Johnson, *Psychol. Sci.* **17**, 383–386 (2006).
37. J. L. Eberhardt, P. A. Goff, V. J. Purdie, P. G. Davies, *J. Pers. Soc. Psychol.* **87**, 876–893 (2004).
38. P. A. Goff, J. L. Eberhardt, M. J. Williams, M. C. Jackson, *J. Pers. Soc. Psychol.* **94**, 292–306 (2008).
39. B. K. Payne, J. W. Hannay, *Trends Cogn. Sci.* 10.1016/j.tics.2021.08.001 (2021).
40. New York Civil Liberties Union, "Stop-and-Frisk in the de Blasio Era" (NYCLU, 2019); www.nyclu.org/sites/default/files/field_documents/20190314_nyclu_stopfrisk_singles.pdf.
41. J. M. Rucker, J. A. Richeson, in *Routledge International Handbook of Prejudice, Stereotyping, and Discrimination*, C. Tileaga, M. Augoustinos, K. Durrheim, Eds. (Routledge, 2021), pp. 13–25.
42. R. C. Hetey, J. L. Eberhardt, *Psychol. Sci.* **25**, 1949–1954 (2014).
43. L. T. O'Brien *et al.*, *Gr. Pro. Inter. Rel.* **12**, 431–444 (2009).
44. P. Freire, *Pedagogy of the Oppressed* (Seabury, 1970).
45. M. A. Diemer *et al.*, *Child Dev. Perspect.* **15**, 12–17 (2021).
46. J. L. Eberhardt, *Biased: Uncovering the Hidden Prejudice that Shapes What We See, Think, and Do* (Viking, 2019)
47. J. A. Richeson, "The mythology of racial progress," *The Atlantic* (2020).
48. N. Dasgupta, S. Asgari, *J. Exp. Soc. Psychol.* **40**, 642–658 (2004).
49. A. H. Eagly, W. Wood, in *Handbook of Theories of Social Psychology*, P. A. M. Van Lange, A. W. Kruglanski, E. T. Higgins, Eds. (Sage, 2012), pp. 458–476.
50. B. K. Payne, H. A. Vuletich, K. B. Lundberg, *Psychol. Inq.* **28**, 233–248 (2017).
51. C. D. Cogburn, *Milbank Q.* **97**, 736–761 (2019).
52. G. Allport, *The Nature of Prejudice* (Addison-Wesley, 1954).
53. J. M. Jones, *Prejudice and Racism* (McGraw-Hill, 1997)
54. J. A. Richeson, S. R. Sommers, *Annu. Rev. Psychol.* **67**, 439–463 (2016).
55. L. D. Bobo, C. Z. Charles, M. Krysan, A. D. Simmons, in *Social Trends in American Life: Findings from the General Social Survey since 1972*, P. V. Marsden, Ed. (Princeton Univ. Press, 2012), pp. 38–83.
56. National Research Council, *The Growth of Incarceration in the United States: Exploring Causes and Consequences*, J. Travis, B. Western, S. Redburn, Eds. (National Academies Press, 2014).


### ACKNOWLEDGMENTS

**Competing interests:** None declared.

10.1126/science.abj7779

REVIEW

# The predatory dimensions of criminal justice

Joshua Page[1]* and Joe Soss[1,2,3]*

Over the past 35 years, public and private actors have turned US criminal justice institutions into a vast network of revenue-generating operations. Today, practices such as fines, fees, forfeitures, prison charges, and bail premiums transfer billions of dollars from oppressed communities to governments and corporations. Guided by scholarship on racial capitalism, we argue that to understand how and why criminal justice operates as it does today, one must attend to its predatory dimensions. Analytically and politically, the concept of predation connects diverse forms of criminal legal takings to one another, to the extractive regimes of earlier eras, and to contemporary businesses that financially exploit subjugated communities. Analyses that focus on predatory relations encourage a reconsideration of some dominant understandings in the study of criminal justice today.

In 2016, Minnesota police killed Philando Castile in his car. Over the previous 14 years, officers had stopped Castile 49 times, issuing 82 citations for minor infractions that cost more than $7000 (*1*). At a rally for her deceased son, Valerie Castile spoke plainly:

> I told my son once before he had got murdered, "These people ain't even looking at you like a man, they looking at you as revenue…Because every time they stop you, they are going to give you a ticket, they are going to tow your car, so that ain't nothing but money." (*2*)

Castile's experiences of roadside injustice are consistent with a long history of racially authoritarian policing in the United States (*3*). Yet they also reflect pronounced change. Over the past 35 years, public and private actors have turned criminal justice institutions into a vast network of revenue-generating operations. Today, practices such as fines, fees, forfeiture procedures, prison charges, and bail premiums extract billions of dollars, disproportionately from race-class subjugated (RCS) communities (*4*), and deliver them to governments and corporations.

Guided by scholarship on racial capitalism, we argue that to understand how and why criminal justice operates as it does today, one must attend to its predatory dimensions. Indeed, analyses that foreground predation encourage a reconsideration of some dominant understandings in the study of policing and punishment.

## Predation and criminal justice

Throughout American history, criminal justice has played a critical role in racialized regimes of resource extraction (*5*). Slave patrols enforced chattel slavery, and military-policing operations facilitated Native dispossession. After the Civil War, prison industries flourished in the North, disproportionately extracting labor from ethnic groups deemed racially inferior. In the South, white elites used chain gangs and convict leasing agreements to rebuild the region. Landowners and merchants relied on criminal codes to trap Black southerners in ruthlessly exploitative debt, credit, and labor arrangements.

Predation connects these histories to the present. The concept specifies relations and practices that (i) are based on a subordinated group's oppression and marginalization and (ii) leverage the group's vulnerabilities and needs to pursue projects of expropriation, extreme exploitation, and/or dispossession. The concept of predation draws diverse criminal legal operations into a common frame of analysis that begins with dominant-subordinate relations and centers questions of power and wealth. As we use the term, "predatory" does not refer to individuals or their diverse motives; it refers to a kind of social relation and its associated practices.

Leading accounts of criminal justice and social inequality focus on dynamics of exclusion, emphasizing, for example, how people marked as "felons" encounter barriers to employment, housing, education, and public assistance (*6*). Such analyses tell half of a larger story. To see the rest, one must analyze a process that scholars such as Taylor theorize as "predatory inclusion" (*7*). Exclusionary practices create group-specific needs and vulnerabilities that can be exploited for profit. The unbanked, for instance, become ripe targets for payday lenders that offer access to quick money at usurious prices. In this view, deprivations are more than just hardships endured at the bottom of the social order; they also supply the basis for predatory projects that enrich advantaged beneficiaries at the top.

In the criminal justice field, practices that divide and exclude—for example, through criminalization, imprisonment, and the policing of social and residential boundaries—underwrite projects of predatory inclusion. The repressive effects of policing and punishment find their counterpart in practices that produce and position targets for resource takings.

By foregrounding such dynamics, the concept of predation provides distinctive terms for analyzing historical continuity and change. Now as in the past, criminal justice practices facilitate extractive projects that underwrite state development and capital accumulation. Nevertheless, several features distinguish the present era from earlier periods.

The dramatic growth of policing and punishment since the 1970s greatly expanded the scale of predatory operations. At the same time, the center of gravity in criminal legal predation shifted from labor to finance. Indeed, the present era is distinguished less by its "criminalization of poverty" or its racialized use of "justice" for predatory purposes (both of which are long-standing) than by the financialization of criminal justice itself. Consistent with developments that financialized the broader political economy (*8*), predatory criminal justice practices pivoted toward tools that charge prices, create debts, and pursue collections.

Prison labor exploits the convicted individual at a particular (carceral) site. New methods of financial takings allowed extractive practices to spread throughout criminal legal institutions and turned justice-involved people into entry points for practices that drain resources from larger social networks. Financialization also expanded predation temporally: By imposing debts, legal authorities lay claim to future resources, enhancing their ability to take what poor communities, by definition, lack.

From this perspective, the criminal legal system can be seen as part of a broader class of predatory enterprises, such as payday lenders, subprime mortgage loans, and high-interest credit card traps. All these operations disproportionately target RCS communities for revenue projects based on steep financial charges, perpetual debt traps, and aggressive collection efforts. Predatory businesses turn to criminal courts to enforce payments (*9*), while fees imposed by these same courts drive many legal debtors into the arms of predatory lenders (*10*).

## Predation and racial capitalism

Studies of racial capitalism explore the coevolving relationship between liberal capitalism and racial domination, as well as their intersections with other axes of oppression (*5*). Scholarship in this tradition clarifies how capitalist projects of exploitation and dispossession have structured shifting constructions of race and racial subjugation. Conversely, it illuminates how the relations of inequality that make capital accumulation possible—for example, owner-worker, creditor-debtor, breadwinner-caretaker, enslaver-enslaved, colonizer-dispossessed—have been organized by differences produced through racism, sexism, and other "isms" beyond class (*11*).

Race and class are long-standing concerns in the study of US policing and punishment. Yet conventional research rarely engages canonical

[1]Department of Sociology, University of Minnesota, Minneapolis, MN, USA. [2]Hubert H. Humphrey School of Public Affairs, University of Minnesota, Minneapolis, MN, USA. [3]Department of Political Science, University of Minnesota, Minneapolis, MN, USA.
*Corresponding author. Email: page@umn.edu (J.P.); jbsoss@umn.edu (J.S.)

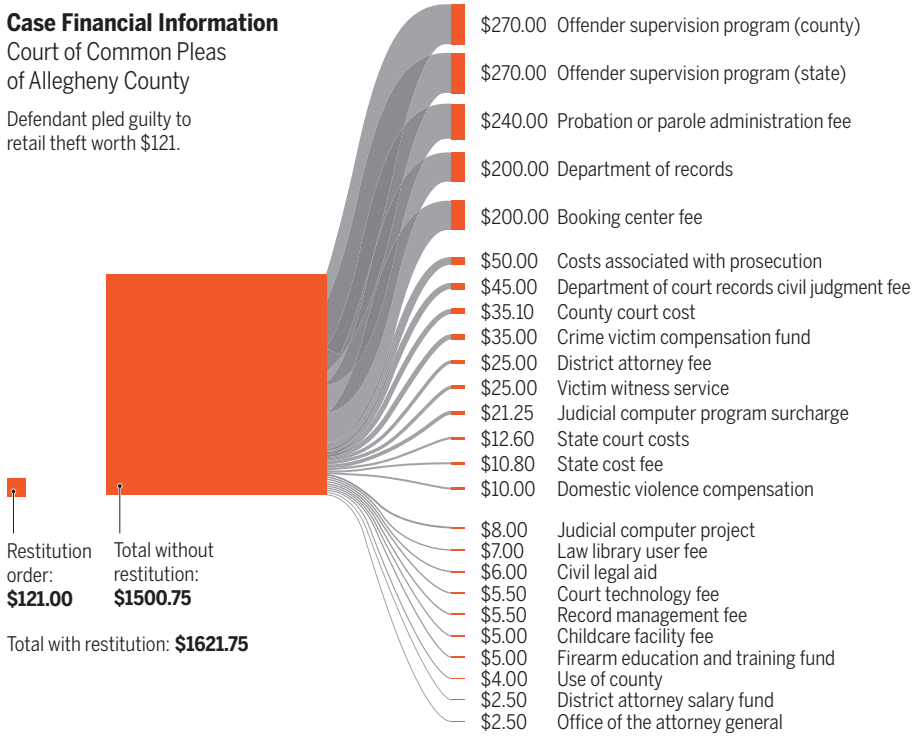



**Fig. 1. Revenue production has become a key responsibility of criminal courts.** Courts routinely impose "surcharges" on convicted persons, often on top of fines. For example, in Allegany County, Pennsylvania, a person who pled guilty to theft of retail goods worth $121 was sentenced to pay $121 in restitution plus an additional $1500.75 in fees, mostly unrelated to the crime. Data are from (*38*).

studies of racial capitalism, such as the works of Walter Rodney, Ida B. Wells, Manning Marable, Barbara Fields, Cedric Robinson, and W. E. B. Du Bois (*5*). Race-and-class–based predation has been a central preoccupation of this literature, with scholars such as Du Bois and Marable emphasizing the historically specific ways that criminal codes and practices have supported racialized modes of capital accumulation.

Against this backdrop, predation offers a generative concept for efforts to understand mass policing and incarceration today. Scholars have long argued that criminal justice institutions serve as tools for controlling subjugated groups that are deemed threatening, disruptive, and unproductive. In the era of mass incarceration, leading accounts frame this social control as a matter of containing and managing "*a surplus population* devoid of market utility" [(*12*), p. 105]. Such analyses illuminate social control but do not connect it to dynamics of predatory inclusion. Racialized practices of penal control do more than just contain and manage; they turn people who might otherwise be "devoid of market utility" into sources of revenue.

To understand criminal justice predation today, one must clarify how and why it differs from earlier formations. Studies of racial capitalism emphasize how the prevailing terms of race and political economy shift over time, often in ways that change the interface of predation and criminal justice. The political struggles of the Civil Rights era, for example, erected barriers to building predatory schemes on explicitly racial foundations. In subsequent decades, the expansion of policing and punishment produced new revenue needs alongside new capacities for resource extraction. In these same decades, privatization and financialization turned governance itself into a collection of sites for capital investment and profit.

The new modes of predation that emerged on this landscape in the 1980s and '90s received little attention in studies that focused on the causes and consequences of mass policing and punishment. In leading accounts, this era of punitive racial politics accelerated bipartisan initiatives that began earlier, in the 1960s and '70s (*13*, *14*). Recent research on "monetary sanctions" integrates financial takings into this framework by defining fines and fees as supplemental forms of punishment produced by the same forces that expanded policing and imprisonment (*15*). By contrast, we distinguish predation from punishment and treat their relationship as an empirical question.

There should be no confusion: Predatory revenue-seeking did not drive the turn toward mass policing and incarceration. Rather, criminal justice growth served as one of several major developments that, together, enabled the rise of new predatory operations.

In the 1980s, state and local governments faced severe fiscal pressures, stemming from several sources: tax cuts for corporations and the rich, reductions in federal funding supports, costly responsibilities shifted downward from the federal government, and spiraling criminal justice expenditures. Desperate to supplement their budgets, public officials began to reimagine criminal legal operations as revenue generators and levy a host of new financial charges on the "consumers" of policing, adjudication, and punishment.

The path forward was smoothed by a racialized politics of law and order that demanded ever-tougher efforts to make criminals pay. RCS communities, already the most heavily policed and punished, became the social and spatial epicenters for new extractive operations that radiated outward to a larger number of (mostly disadvantaged) Americans (*16*).

In this same period, criminal justice became a far more attractive site for corporate investors (*17*). The rapid growth of incarceration left public authorities scrambling to address unconstitutional levels of jail and prison overcrowding. Entrepreneurs developed new business models that, they claimed, would help state agencies overcome their cost and management problems. Corporate interests spent millions on political action to remove barriers to entry and install for-profit operations throughout the field. Increasingly, governments outsourced costly responsibilities and, in many cases, generated revenues through profit-sharing agreements.

### Predation in practice

Outside RCS communities, the steady expansion of financial takings drew so little public attention that many Americans were shocked when a 2015 US Department of Justice investigation revealed that police and courts in Ferguson, Missouri, were acting as agents of race-targeted plunder (*18*). City officials budgeted for "sizeable increases" in fines and fees each year, exhorted police and courts to produce larger cash flows, and closely monitored changes in revenue delivered [(*18*), p. 2]. Fines and fees (overwhelmingly imposed on Black residents) provided the city's second largest source of local revenue, creating dedicated funding for a wide range of public services, an $8 million fire hall, a $3.7 million renovation to the police station, and an 8% raise for all municipal employees [(*19*), p. 932].

The beneficiaries of this regime included most residents outside the city's most policed and punished communities. In this respect, Ferguson was no anomaly. Across the nation, predatory criminal justice revenues subsidize public services, pay for middle-class jobs in businesses and governments, and lessen the tax burdens of advantaged citizens. They fund corporate profits, lavish CEO salaries, and lucrative returns for Wall Street banks and investors. Criminal justice predation, in this sense, shapes life conditions not only in the

communities it targets but also among more privileged Americans who are scarcely aware of its existence.

Throughout the United States, police stop and ticket drivers, bicyclists, and pedestrians, most intensively in RCS communities. In 2016, for example, the city of Chicago brought in $264 million from auto-related charges alone, accounting for about 7% of the city's $3.6 billion operating budget (*20*). Policing also generates substantial revenue through asset forfeiture, a procedure for seizing cash and goods based solely on an allegation of illicit origins. Justified as a method to bring down drug kingpins and terrorists, asset forfeiture is mostly used to pursue large numbers of small takings from poor and working-class people. Between 2000 and 2019, federal and state governments seized assets worth a total of at least $68.8 billion [(*21*), p. 5]. Shockingly, and for the first time, the overall value of asset forfeitures in 2014 surpassed total losses from burglaries nationwide (*22*).

Revenue production has also become a key responsibility of criminal courts. In a landmark study of Washington state, Harris documented the proliferation of court fees and fines since the 1990s (*23*). Today, most states allow criminal courts to bill indigent defendants for legal representation, and some even charge fees for exercising the right to a jury trial. Courts routinely impose a costly array of "surcharges" on convicted persons, often on top of fines, which have grown steeper and more numerous. Michigan, for example, generates roughly $290 million from surcharges per year, making up about a quarter of all funding for the state's trial courts [(*24*), p. 7]. The costs of court fees for individuals can be staggering. A case from Allegany County, Pennsylvania, provides a striking example: A person who stole retail goods worth $121 was sentenced to pay $121 in restitution plus an additional $1500.75 in fees, mostly unrelated to the crime (see Fig. 1).

Courts also funnel defendants into the bail bond industry. Between 1990 and 2009, the frequency and size of monetary bail conditions set by judges in large counties rose substantially (*25*). Required bail amounts are routinely set beyond the financial reach of defendants. To secure pretrial release, defendants often must turn to bail bond companies and pay nonrefundable premiums (typically 10% of the bail). The full amount of money siphoned from poor communities by the bail industry is unknown because, unlike government agencies, private businesses are not obligated to disclose financial records. In 2012, however, the big insurance companies that underwrite bail contracts secured about $13.5 billion in bonds (*26*). Local "mom-and-pop" storefronts are the visible face of commercial bail, but corporate underwriters, owned in many cases by larger conglomerates and private equity firms, reap the lion's share of profits.

In jails and prisons, the accused and convicted are targeted by a further complex of revenue-generating practices (*27*). Today, almost every state charges imprisoned people "pay-to-stay fees" for costs such as room and board, medical and dental care, clothing, and toiletries. Governments also generate funds by selling corporations access to incarcerated populations. The companies exploit imprisoned people as a cheap source of labor and as a captive market for overpriced goods and services. In return for granting access, public agencies typically get a slice of the profits.

In this manner, telecom companies such as Securus and Global Tel-Link (GTL) have turned phone calls, video visitation, e-tablets, and other connections to the outside world into a $1.6 billion industry [(*17*), p. 73]. Carceral banking companies charge families and friends to deposit funds that imprisoned people need to buy basic necessities from commissaries, make phone calls, pay for medical services, and so forth. JPay, a leader in the field, brings in nearly $144 million in annual revenue [(*27*), p. 61]. In recent years, private equity firms have bought out some of the most prominent businesses in the sector (including JPay, Securus, and GTL), delivering the profits from carceral predation to a wider array of investors, including Wall Street banks and public pension funds.

Meanwhile, the corporations that manage private jails, prisons, and immigrant detention facilities, such as CoreCivic and GEO Group, have turned government contracts and extractive carceral practices into a $5 billion industry [(*17*), p. 30]. These publicly traded companies serve as investment vehicles for shareholders, as well as the banks that finance them.

Similar developments have transformed community supervision. From 1990 to 2014, the number of states charging people for their own probation and parole rose from 26 to 44 (*28*). Some states outsource probation supervision to for-profit companies, and most hire businesses to provide technologies and services. As of 2014, every state except Hawaii charged people on supervision for their own electronic monitoring (*28*).

As police and courts pursue payments on legal debts, they also act as enforcers for debts generated outside the field itself, such as child-support payments that public welfare agencies claim as "reimbursements" and money owed to private landlords, payday lenders, health care providers, credit card companies, and the like. When people fall behind on such payments, judges routinely declare them "willfully delinquent" and require them to serve time in jail (*8*).

### Consequences for communities

By beginning with dominant-subordinate group relations, the concept of predation decenters the policed, accused, or convicted individual. Criminal justice predation operates at the community level, organized by various axes of social inequality. Imprisoned individuals, for example, function as entry points for what Katzenstein and Waller describe as "an invisible system of revenue and taxation that exploits the ties of family dependency" [(*29*), p. 644)].

In RCS communities, such takings operate as a substantial source of wealth depletion and hardship (*10*). Fearing harsh reprisals for nonpayment, legal debtors and their kin drain what little resources they have. They forgo meals, medical care, and rent payments. They sell their blood and turn to predatory lenders. They set aside life goals and accept low-end jobs on whatever terms they can get. In a study of supervision fees, one probationer explained:

> I do without sometimes. I pick and choose what I eat. When I've worked in fast food, I'll take home what we didn't use… Don't buy shoes. Don't buy extra clothes. I need two hearing aids. I've been taking the money I have saved up for hearing aids to pay for this [probation]. [(*30*), p. 2]

The gendered organization of care in the United States combines with the race-and-class targeting of criminal justice to make the financial burdens of predation distinctively heavy for women from RCS communities (*31*). In ethnographic research conducted as a bail bond agent, one of us (Page) observed women disproportionately paying costly bail premiums and accepting onerous cosigning duties because they felt they had no choice. A Native American woman whose partner was charged with statutory rape, for example, felt forced to bail out her partner "for the kids." The defendant's mother was furious with her son but helped arrange his bail because, as she put it, the man was her "baby" [(*31*), pp. 161–162].

On behalf of loved ones, women (as mothers, wives, girlfriends, daughters, aunts, sisters, or friends) drain their savings, work extra hours, and skip utility, rent, and mortgage payments. A formerly incarcerated person told researchers:

> Everything that was put into bailing me out was everything my mother had in savings, and she borrowed some money from my grandparents. She was back to working paycheck to paycheck. Eventually, about a year and a half after being locked up, my mother had to give up the house she loved and move back to an apartment. [(*32*), p. 14]

Costly financial assistance of this sort can sow frustration and discord in relationships that are vital for personal and social well-being, weakening commitments to mutual aid as well as networks of emotional and material support.

As it deepens poverty and hardship, predation also reconstructs citizenship. Criminal legal authorities engage some Americans as equal citizens, deserving of protection, service, and justice, while mining others for dollars and subjecting them to ongoing, debt-based

social control. In so doing, they can teach potent lessons about racial hierarchy, civic standing, law, and the state (*3*, *33*). The meaning and experience of citizenship becomes tethered to an insistent question: What does it mean to be treated—by one's own government and in the name of law and justice—as a target of plunder?

Legal debtors become "perpetual subjects of the criminal justice system who at any time can be called to answer for their nonpayment and may even be incarcerated" [(*23*), p. 73]. Failures to pay can lead to revocations of probation or parole, driver's license suspensions, exclusions from public programs, and even the stripping of voting rights (*16*). Legal debtors live in fear of the consequences that might result from an ordinary encounter with law enforcement or a missed court date.

The ongoing threats of life disruption and state violence can be terrifying. As a respondent in one study put it: "[Legal debt] is overwhelming, [it] causes anxiety. I go to therapy because you are always scared they will be knocking at your door. I have started using [drugs] because of the anxiety" [(*34*), p. 19]. This haunting effect of state power is likely to be especially fraught for Black Americans, who have endured racist and authoritarian state violence for centuries and may reasonably fear that ordinary interactions with police will turn deadly (*3*).

### Political action and social science

Throughout history, people have fought back against predation. Today is no different. The past two decades have produced prison labor strikes, protests led by RCS communities, lobbying campaigns, and other efforts to reform or abolish predatory criminal legal practices. Activists have led divestment campaigns, and public interest lawyers have sued governments and corporations for revenue-driven rights violations.

Across a fragmented field of issue- and jurisdiction-specific campaigns, political action has produced a modest, but growing, collection of victories (*35*). In several states and localities, phone-justice campaigns have reduced or eliminated jail and prison telecom charges. A small number of states and counties have pared back fines and fees for juveniles and adults. Advocates have pushed some jurisdictions to curb debt-based driver's license suspensions and jailing practices. Bail reformers have pressured several states and localities into imposing limits on pretrial profiteering. Left and liberal organizations have sometimes found common ground with libertarians in the fight against asset forfeiture, producing successful legal and legislative challenges. Public shaming campaigns have led some large banks to cut ties with for-profit prison corporations.

Scholars have worked closely with many of these efforts, learning from affected communities and activists, producing research to inform advocates and policy-makers, developing plans for alternative institutional arrangements, and using their voices to publicly support policy change. Over the past decade, experts on monetary sanctions have made especially notable contributions to social science and policy reform (*15*). Research on privatization has illuminated the complex world of prison profiteering, aiding and, in some cases, motivating reform efforts (*17*, *27*). Legal scholars have analyzed the constitutional and statutory bases of government fees and seizures, clarifying how they might effectively be challenged in courts (*36*).

In this context, the concept of predation offers a unifying frame, capable of integrating disparate streams of scholarship and political action. Analytically and politically, it connects diverse criminal legal practices to one another, to the extractive regimes of earlier eras, and to the exploitative lending and credit businesses that pervade RCS communities today. Throughout, it foregrounds the fundamental questions of social structure prioritized by scholars of racial capitalism. The idea of predatory governance, resonant with connotations of group-based injustice, offers a powerful language for political mobilization and coalition building (*37*). For social scientists, it provides a generative basis for critical scholarship that connects empirical research to normative inquiry and action-oriented analyses of power and social change.

In the study of criminal legal predation, a number of pressing questions set directions for future research. At present, we lack even rough national estimates of the resources stripped from communities, their effects on poverty and inequality, and their contributions to specific sectors of state development and capital accumulation. Little is known about how organizations (e.g., courts, police agencies, and prisons) adapt their structures, cultures, and routines as they pursue predatory projects. How, if at all, have predatory policies altered the political landscape of criminal justice as they have drawn new interests to the field, created incentives to defend and expand revenue-generating operations, and, at times, mobilized popular resistance? How does pocketbook policing affect the amount and patterning of racialized police violence?

Going forward, scholarship on predation must do far more to ground its pursuit of knowledge in the diverse experiences, perspectives, and priorities of RCS communities. Because predation is a relationship, understandings of it are distorted by analyses that treat dominant groups and institutions as the only effective doers in social life, eclipsing the knowledge and agency of subjugated groups. Close attention to both sides of this relationship is essential for meeting the overarching challenge that confronts us today: to develop a historically specific and actionable understanding of criminal justice predation as a constitutive element of racial capitalism and liberal political economy. A serious reckoning with the predatory dimensions of criminal justice is long overdue—not only for scholars but also for all who hope to build a more just, humane, and democratic society.

### REFERENCES AND NOTES


1. B. Stahl, "Philando Castile was caught up in a cycle of traffic stops, fines," *Minneapolis Star-Tribune*, 17 July 2016.
2. News on Purpose, Remembrance of Philando Castile Rally, Facebook video, 6 July 2020.
3. V. M. Weaver, G. Prowse, *Science* **369**, 1176–1178 (2020).
4. J. Soss, V. Weaver, *Annu. Rev. Polit. Sci.* **20**, 565–591 (2017).
5. D. Jenkins, J. Leroy, *Histories of Racial Capitalism* (Columbia Univ. Press, 2021).
6. B. Western, *Punishment and Inequality in America* (Russell Sage Foundation, 2007).
7. K.-Y. Taylor, *Race for Profit: How Banks and the Real Estate Industry Undermined Black Homeownership* (Univ. of North Carolina Press, 2019).
8. G. R. Krippner, *Capitalizing on Crisis: The Political Origins of the Rise of Finance* (Harvard Univ. Press, 2011).
9. American Civil Liberties Union (ACLU), "A pound of flesh: The criminalization of private debt" (Report, ACLU, 2018).
10. Alabama Appleseed, "Under pressure: How fines and fees hurt people, undermine public safety, and drive Alabama's racial wealth divide" (Report, Alabama Appleseed, 2018).
11. J. Melamed, *Crit. Ethn. Stud.* **1**, 76 (2015).
12. L. Wacquant, *Punishm. Soc.* **3**, 95–133 (2001).
13. E. Hinton, *From the War on Poverty to the War on Crime: The Making of Mass Incarceration in America* (Harvard Univ. Press, 2017).
14. N. Murakawa, *The First Civil Right: How Liberals Built Prison American* (Oxford Univ. Press, 2014).
15. K. D. Martin, B. L. Sykes, S. Shannon, F. Edwards, A. Harris, *Annu. Rev. Criminol.* **1**, 471–495 (2018).
16. M. Feola, *Theory Event* **23**, 877–901 (2020).
17. L. Eisen, *Inside Private Prisons: An American Dilemma in the Age of Mass Incarceration* (Columbia Univ. Press, 2017).
18. T. Shaw, *The Ferguson Report: Department of Justice Investigation of the Ferguson Police Department* (The New Press, 2015).
19. K. Henricks, D. C. Harvey, *Soc. Forum* **32**, 930–951 (2017).
20. M. Sanchez, S. Kambhampati, "How does Chicago make $200 million a year on parking tickets? By bankrupting thousands of drivers," *Mother Jones*, 27 February 2018.
21. Institute for Justice, "Policing for profit: The abuse of civil asset forfeiture" (Report, Institute for Justice, December 2020).
22. C. Ingraham, "Law enforcement took more stuff from people than burglars did last year," *Washington Post*, 23 November 2015.
23. A. Harris, *A Pound of Flesh: Monetary Sanctions as Punishment for the Poor* (Russel Sage Foundation, 2016).
24. Trial Court Funding Commission, "Trial Court Funding Commission final report" (Report, State of Michigan, Trial Court Funding Commission, 2019).
25. Council of Economic Advisors, "Fines, fees, and bail" (Issue brief, Council of Economic Advisors, 2015).
26. S. Bauer, "Inside the wild, shadowy, and highly lucrative bail industry," *Mother Jones*, May/June 2014.
27. Worth Rises, *The Prison Industry: How it Started. How it Works. How it Harms.* (Worth Rises, 2020).
28. J. Shapiro, "As court fees rise, the poor are paying the price," National Public Radio, All Things Considered, 19 March 2014.
29. M. Katzenstein, M. Waller, *Perspect. Polit.* **13**, 638–656 (2015).
30. E. L. Ruhland, J. P. Robey, R. P. Corbett Jr., K. R. Reitz, "Exploring supervision fees in four probation jurisdictions in Texas" (Report, Robina Institute of Criminal Law and Criminal Justice, 2017).
31. J. Page, V. Piehowski, J. Soss, *RSF* **5**, 150 (2019).
32. S. deVuono-powell, C. Schweidler, A. Walters, A. Zohrabi, "Who pays? The true cost of incarceration on families" (Report, Ella Baker Center for Human Rights, Forward Together, Research Action Design, 2015).
33. M. C. Bell, *Yale Law Rev.* **126**, 2054 (2017).
34. F. Cook, "The burden of criminal justice debt in Alabama" (Report, UAB TASC, 2014).
35. Fines and Fees Justice Center, The Clearinghouse (2021); https://finesandfeesjusticecenter.org/clearinghouse/.
36. B. A. Colgan, *UCLA Law Rev.* **65**, 2–77 (2018).
37. H. L. Walker, *Mobilized by Injustice: Criminal Justice Contact, Political Participation, and Race* (Oxford Univ. Press, 2020).
38. A. B. Stark, G. Walsh, "Clearing a path to a new beginning: A guide to discharging criminal justice debt in bankruptcy" (Report, National Consumer Law Center, 2020).


### ACKNOWLEDGMENTS

We thank S. Piston for insightful feedback on an earlier draft of this essay.
**Competing interests:** The authors declare no competing interests.

10.1126/science.abj7782

# Your Legacy to Science

AN ESTATE GIFT TO THE
AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

**Since 1848, our founding year,** the American Association for the Advancement of Science (AAAS) has been deeply committed to advancing science, engineering and innovation around the world for the benefit of all people.

By making AAAS a beneficiary of your will, trust, retirement plan or life insurance policy, you become a member of our 1848 Society, joining Thomas Edison, Alexander Graham Bell and the many distinguished individuals whose vision led to the creation of AAAS and our world-renowned journal, *Science*, so many years ago.

Unlike many of its peers, *Science* is not for-profit. Your estate gift would provide long-term financial stability and durable annual income that will support operations and competitive innovation for years to come. **This support is vital.**

***"As a teacher and instructor, I bear responsibility for the younger generations. If you have extra resources, concentrate them on organizations, like AAAS, that are doing work for all."***

**—Prof. Elisabeth Ervin-Blankenheim, 1848 Society member**

If you intend to include AAAS in your estate plans, provide this information to your lawyer or financial adviser:

**Legal Name:** American Association for the Advancement of Science
**Federal Tax ID Number:** 53-0196568
**Address:** 1200 New York Avenue, NW, Washington, DC 20005

If you would like more information on making an estate gift to AAAS, cut out and return the form below or send an email to philanthropy@aaas.org. Additional details are also available online at **www.aaas.org/1848Society**.

1848 SOCIETY AAAS

*cut here*

## Yes, I would like more information about joining the AAAS 1848 Society.

**PLEASE CONTACT ME AT:**

Name: _____

Address: _____

City: _____ State: _____ Zip code: _____ Country: _____

Email: _____ Phone: _____

**RETURN THIS FORM TO:**
AAAS Office of Philanthropy and Strategic Partnerships • 1200 New York Avenue, NW • Washington, DC 20005 USA

# RESEARCH

## IN *SCIENCE* JOURNALS

*Edited by* **Michael Funk**

**Dexterous hand movements, such as those required for climbing, are directed in part by circuits in the cuneate nucleus of the mouse brainstem.**

NEUROSCIENCE

### Sensory signals to fine-tune and movement

Skilled hand movements are modulated not only by signals from the motor system but also by sensory feedback. However, little is known about the circuits that regulate these feedback signals and how such regulation might influence movement. Conner *et al*. combined molecular, electrophysiological, and behavioral approaches in the mouse to identify and characterize inhibitory circuits in the brainstem cuneate nucleus. These circuits can enhance or suppress the transmission of tactile information, thereby affecting behaviors that require dexterous movement of the hand. In addition, there are descending cortical inputs that can inhibit or excite the cuneate nucleus. These findings indicate a new anatomical and functional circuit architecture for the adjustment of tactile feedback. —PRS *Science*, abh1123, this issue p. 316

CATALYSIS

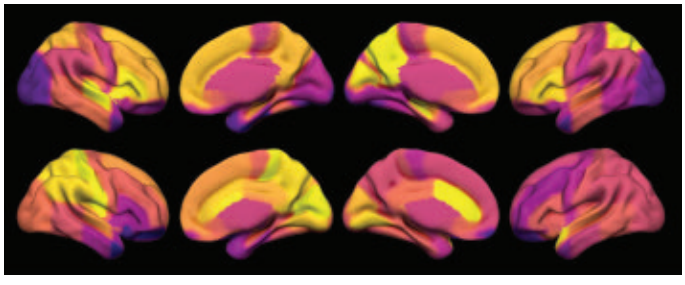
### Which phosphines squeeze together?

Phosphine ligands coordinated to palladium and nickel are essential tools for assembling the backbones of pharmaceutical compounds. For decades, descriptors that characterize spatial bulk have helped to guide phosphine optimization. However, these descriptors tend to apply to ideal geometries of a single ligand. Newman-Stonebraker *et al*. introduce a descriptor that considers how the ligand conformation might change in a crowded environment. Specifically, they found that the minimum percentage buried volume accurately predicts when one or two of a particular ligand will coordinate to a metal center, frequently a key determinant of successful catalysis. —JSY *Science*, abj4213, this issue p. 301

NEURODEGENERATION

### Influential microglia

Neurodegenerative diseases are associated with alterations in brain energy metabolism. Recent developments in imaging techniques allow for the visualization of glucose metabolism alterations using 2-deoxy-2-[$^{18}$F]fluoro-D-glucose positron emission tomography (FDG-PET). Xiang *et al*. investigated the role of microglia in determining the FDG-PET signal. The authors found that glucose uptake is greater in microglia compared with astrocytes and neurons, and that microglia state determined the alterations in FDG-PET signal seen in mouse models of neurodegenerative diseases. In patients, microglia activity correlated with FDG-PET, suggesting the translational value of the preclinical data. The results suggest that microglia activity should be taken into consideration when performing brain FDG-PET. —MM *Sci. Transl. Med.* **13**, eabe5640 (2021).



**Positron emission tomography reveals a correlation between microglia activity and glucose uptake in patients with neurodegenerative diseases.**

PHYSICS

### Getting more from Moire

Moire superlattices of twisted two-dimensional materials have emerged as a new paradigm for exploring strongly correlated electron systems. Once fabricated, however, the electronic correlations and the subsequent behavior of any device

CREDITS (TOP TO BOTTOM) INGO ARNDT/MINDEN PICTURES; XIANG *ET AL.*

structure tend to be fixed. Schwartz *et al*. show that further sandwiching of the bilayer heterostructures between graphene electrical gates can provide a route to controlling the electronic properties in each layer separately and the electronic interaction strengths overall. The fine control of the electronic interactions, in terms of a tunable Feshbach resonance, provides a platform to explore many-body physics and quantum simulation. —ISO

*Science*, abj3831, this issue p. 336

EXOPLANETS

## Compositions of rocky exoplanets

The interior compositions of small rocky exoplanets cannot be observed directly but are expected to relate to the composition of the host star. Adibekyan *et al*. analyzed a sample of rocky exoplanets, inferring the planets' iron fractions by combining their masses and radii with an interior structure model. The iron fractions of the host stars were calculated from stellar elemental abundances. The two iron fractions, that of the planets and that of the stars, correlated with each other, but the slope was steeper than 1, indicating that planet formation processes modify the compositions of rocky planets. —KTS

*Science*, abg8794, this issue p. 330

MUSCLE REPAIR

## Muscle repair without stem cells

Skeletal muscle is a mechanical organ that endures cellular damage after contraction. Lesions caused by external injury can be repaired by muscle stem cells, which fuse with injured cells or create entirely new myofibers. Roman *et al*. describe a cell-autonomous repair process that is independent of muscle stem cells (see the Perspective by McNally and Demonbreun). After localized damage, myonuclei migrate to injury sites and locally deliver messenger RNA for cellular reconstruction. This myofiber self-repair represents a model for understanding the restoration of muscle architecture in health and disease. —BAP

*Science*, abe5620, this issue p. 355; see also abm2240, p. 262

PSYCHOLOGY

## Viewing victims as morally virtuous

Victims of wrongdoing are often perceived to have elevated moral character compared with otherwise similar nonvictims. Drawing on results from 17 experiments in which study participants recruited online or from a university campus pool rated the moral character of a person in a hypothetical story, Jordan and Kouchaki found that this virtuous victim effect exists, not because of anything the victims themselves have done, but rather because of their experience of mistreatment. The authors also found that seeing victims as especially moral can motivate the punishment of perpetrators and the helping of other victims, which they call the "justice restoration hypothesis." This research has implications for how victims make decisions about disclosing their narratives and for policy and practice responses. —KT

*Sci. Adv.* 10.1126/sciadv.abg5902 (2021).

HEALTH ECONOMICS

## Decisions about delivery

Heuristics, or simplified decision rules, have been found to influence decision-making in several settings. Singh found that when patients experience a complication during birth either by vaginal or cesarean delivery, their doctors are more likely to switch to the opposite delivery mode for their next patient's birth (see the Perspective by Li and Colby). There is also evidence that this heuristic may lead to worse patient outcomes. —TSR

*Science*, abc9818, this issue p. 324; see also abl5647, p. 260

# IN OTHER JOURNALS

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

Pain killers, such as these oxycodone tablets, are linked to opioid-related overdose mortality among U.S. armed forces veterans caught up in the legal system.

SUBSTANCE ABUSE

## The injustice of overdose

Opioids have been used extensively for pain control but are associated with the collateral risk of dependence, addiction, or even death. Studies have shown that individuals released from incarceration are five- to 10-fold more likely than the general population to overdose on opioids. Finlay *et al.* examined the records of >5 million people attending the U.S. Veterans Health Administration during 2012. Of these, 32,284 had been involved with the legal system in the same year and, of these, 295 died from opioid-related overdoses in the ensuing 5 years. Population comparisons showed that when armed forces veterans were involved with the legal system, they had a higher risk of opioid-related overdose mortality. These data indicate that more effort should be put into substance disorder treatment and other mental health and life management support for veterans. —BAP

*Am. J. Prev. Med.* 10.1016/j.amepre.2021.06.014 (2021).

EPIGENOMICS

## Mapping disease similarities

Epigenetic effects can underlie human diseases and may explain how multiple diseases, including those with Mendelian inheritance, map to different genes yet exhibit phenotypic similarity. In mice, Luperchio *et al.* looked at how phenotypic disease commonalities with different underlying

PHOTO: GIPHOTOSTOCK/SCIENCE SOURCE

CARBON CYCLE

## Going nowhere

Carbon dioxide fertilization and climate warming have made Europe greener during this century but has that translated into an increase in ecosystem carbon uptake? Liu *et al.* report that simultaneous increases in ecosystem respiration negated the impact of the additional photosynthesis, and that heatwaves caused stronger reductions in carbon uptake than carbon loss. Increases in the frequency and severity of heatwaves that are expected to accompany climate warming will affect the European carbon cycle more in the future, beyond just the effects of rising average temperatures. —HJS

*Geophys. Res. Lett.* 10.1029/2021GL094870 (2021).

**The recent terrestrial greening of Europe has not led to an increase in net ecosystem carbon uptake.**

genetic causes may result from shared downstream epigenetic changes. Examination of chromatin accessibility in B cells of mutant mouse models for the similar human genetic diseases Kabuki syndrome 1 and 2 and Rubinstein-Taybi syndrome type 1 identified similar changes in chromatin accessibility patterns in 420 gene promoters. Transcriptome analysis showed an overlapping correlation between changes in promoter accessibility and gene expression. —LMZ

*eLife* **10**, e65884 (2021).

HEMATOLOGY

## Undivided attention on mitochondria

Myelodysplastic syndromes are malignant blood disorders associated with ineffective hematopoiesis and clonal expansion of abnormal bone marrow cells. To examine the underlying mechanisms responsible for ineffective blood cell development in these conditions, Aoyagi *et al*. studied mouse models with hematopoietic stem cells bearing some of the mutations seen in human patients. By characterizing the gene expression patterns and metabolic activity of the mutant cells, the authors detected mitochondrial fragmentation and examined the mechanism tying it to defective hematopoiesis. They also showed a protective effect of a compound inhibiting a key mitochondrial division protein, suggesting a potential therapeutic approach. —YN

*Cancer Discov.* 10.1158/2159-8290.CD-21-0032 (2021).

RNA INTERFERENCE

## Chopping up viruses

Plants and invertebrates use a gene-silencing mechanism called RNA interference (RNAi) to defend themselves against viral infection. When a virus invades a cell, its RNA replicates into double-stranded RNA, which is detected by the cell as invasive and is specifically cleaved by a molecular scissor called Argonaute-2. Although mammalian cells contain all of the genes that make up the RNAi arsenal, whether RNAi plays a meaningful role in mammalian antiviral immunity has been controversial. In self-defense, viruses produce proteins called viral suppressors of RNAi (VSRs). Fang *et al*. designed peptides that target the VSR of enterovirus A71 (EV-A71), which causes hand-foot-and-mouth disease in millions of children annually. The peptides liberated the natural antiviral RNAi to inhibit infection in both cells and mice. Therefore, RNAi is functional as an antiviral immunity in mammals. —DJ

*Immunity* 10.1016/j.immuni.2021.08.027 (2021).

C–H BOND ACTIVATION

## When PT edges out ET

Enzymes and chemical catalysts break strong carbon–hydrogen (C–H) bonds through multiple distinct combinations of electron transfer (ET) and proton transfer (PT). Barman *et al.* investigated the subtleties of how a manganese-oxido complex reacts with hydrocarbons of varying acidity and C–H bond strength. They concluded that, in this case, the reaction proceeds through an asynchronous mechanism, with PT wielding primary influence over the rate but ET still contributing to the rate-limiting transition state. The authors' semi-empirical relation weighting these contributions could help to elucidate similar mechanisms in other systems. —JSY

*Proc. Natl. Acad. Sci. U.S.A.* **118**, e2108648118 (2021).

NUCLEAR PHYSICS

## Doubly magical tin

In an atomic nucleus, a closed shell of either neutrons or protons promotes stability. These special proton and neutron numbers are called magic numbers, and when both the proton and the neutron number in a nucleus are magic, the nucleus is called doubly magic. The doubly magic nuclide of tin, $^{100}$Sn, has 50 protons and as many neutrons but hovers at the limit of stability. Its mass is deduced from those of neighboring nuclides of indium and its beta-decay energy. Mougeot *et al.* measured the masses of $^{99–101}$In, which, when combined with the beta-decay results from two previous studies, resulted in a substantial difference between the corresponding extracted masses of $^{100}$Sn. The researchers' numerical studies aim to resolve the discrepancy. —JS

*Nat. Phys.* 10.1038/s41567-021-01326-9 (2021).

PHOTO: NANCY CARELS/MINDEN PICTURES

## ALSO IN *SCIENCE* JOURNALS

*Edited by* **Michael Funk**

IMMUNOLOGY

### Mitochondria drive CTLs' killer instinct

Cytotoxic T lymphocytes (CTLs) can terminate both virally infected cells and cancer cells by secreting cytolytic proteins such as perforin and granzyme B. CTLs are particularly effective because they can sequentially kill multiple targets in a process called serial killing. Lisci *et al*. have identified mitochondria as important regulators of CTL killing. Mice lacking the deubiquitinase USP30 have CTLs acutely depleted of mitochondria, and these cells have reduced killing ability but normal motility, signaling, and secretion. Surprisingly, the mitochondria's metabolic functions were not required for this process. Rather, mitochondrial translation proved indispensable for CTL cytolytic protein synthesis and sustained CTL killing. —STS

*Science*, abe9977, this issue p. 299

NEUROSCIENCE

### Glia control white matter information flow

The myelinated axons of the white matter mediate rapid information transmission between the brain's processing nodes. Axonal excitability and conduction speed are key determinants of neural circuit function but it is largely unknown how they are regulated. Lezmy *et al*. examined the effects of adenosine on the functional properties of myelinated axons. They observed activity-dependent calcium rise in astrocyte processes, calcium-triggered ATP release from astrocytes, activation of adenosine receptors in the axon initial segment and the nodes of Ranvier, increased cyclic AMP concentration, activation of HCN2 channels, and depolarization of axons by 5 to 10 millivolts. This molecular pathway enabled astrocytes to increase the excitability of pyramidal neurons and to profoundly decrease the axonal conduction velocity. Astrocytes thus control signaling speed in myelinated axons. —PRS

*Science*, abh2858, this issue p. 300

ZEOLITE SYNTHESIS

### Selectivity control in zeolite synthesis

Zeolites are widely used in many industrial applications, but despite decades of research, their synthesis still relies on trial-and-error approaches. Complex nucleation mechanisms and topological diversity lead to strong phase competition, complicating the issue of rational design of zeolite synthesis. Using atomistic simulations, literature mining, human-computer interaction, synthesis, and characterization, Schwalbe-Koda *et al*. developed a computational strategy that enables a priori control of phase selectivity in zeolite synthesis (see the Perspective by Chaikittisilp and Okubo). This approach uses several metrics for designing organic structure–directing agents to crystallize target zeolites with controlled phase competition and intergrowth. These results may have profound implications for the materials science community if this method is shown to be successful in the synthesis of practically useful uncommon zeolites. —YS

*Science*, abh3350, this issue p. 308; see also abm0089, p. 257

TROPHIC CASCADES

### The importance of disturbance

Work in sea otters over the past few decades has transformed our understanding of the importance of specific species, or keystones, as drivers of community structure and stability. Foster *et al*. took the next step and tested whether otter foraging might influence genetic diversity in an eelgrass ecosystem (see the Perspective by Roman). The authors found that eelgrass genetic diversity was significantly higher where otters were present and that the impact was related to time: Longer otter presence was associated with higher genetic diversity. These results illustrate how the actions of a predator can affect the diversity of a producer in a tropic system. —SNV

*Science*, abf2343, this issue p. 333; see also abm2257, p. 256

OCEAN CIRCULATION

### Together from afar

The Gulf Stream and the Kuroshio are the large surface boundary currents in the western Atlantic and Pacific Oceans, respectively. They transport heat from the tropics to the extratropics and influence sea surface temperature, weather, and climate throughout the Northern Hemisphere. Kohyama *et al*. combined satellite observations with model experiments to show that these two currents are synchronized on decadal time scales by the meridional migration of the atmospheric jet stream (see the Perspective by Cessi). This coupling can help to explain climate events such as the abnormally hot summer of 2018 in the Northern Hemispheric extratropics. —HJS

*Science*, abh3295, this issue p. 333; see also abl9133, p. 259

DEVELOPMENTAL BIOLOGY

### Making cardiac valves via mechanical forces

Cardiac valves form in response to mechanical forces generated by the beating heart. Fukui *et al*. studied how patterning signals are generated in response to these forces (see the Perspective by Jain and Epstein). They show that two mechanotransduction pathways act in parallel to instruct cardiac valve progenitors: a well-established transient receptor potential mechanosensation pathway and an extracellular ATP-dependent purinergic receptor pathway that triggers $Ca^{2+}$ oscillations and results in nuclear translocation of the protein nuclear factor of activated T cells 1. These two synergistic mechanotransduction pathways generate positional information and control valve formation. The use of multiple pathways may be a general solution to increase the robustness and precision of responses to mechanical forces in biological systems. —BAP

*Science*, abc6229, this issue p. 351; see also abm1858, p. 264

T CELLS

### Cholesterol metabolite quiets T cells

Interleukin-27 (IL-27) is a cytokine with known immunoregulatory effects, as shown by prior studies of IL-27–deficient mice, but the mechanisms contributing to IL-27's suppressive effects are still poorly understood. Takahashi *et al.* identified the gene encoding cholesterol 25-hydroxylase (Ch25h), the enzyme that converts cholesterol to 25-hydroxycholesterol (25OHC), as being strongly and selectively induced in T cells stimulated with IL-27 versus other combinations of cytokines. Extracellular 25OHC selectively impaired in vitro growth of activated T cells through impairment of cholesterol biosynthesis while sparing quiescent T cells. In autoreactive T cells capable of inducing dermatitis after adoptive transfer, genetic absence of Ch25h or the IL-27 receptor resulted in more severe skin disease. These findings identify IL-27–induced 25OHC as an immunomodulatory cholesterol metabolite with therapeutic potential in autoimmune and inflammatory diseases. —IRW

*Sci. Immunol.* **6**, eabb6444 (2021).

CARDIAC PHYSIOLOGY

## Pressured to become inflammatory

Cardiac injury or disease can lead to pressure overload, inflammation, and heart failure. Abe *et al*. sought to identify how inflammatory monocytes and macrophages are recruited to the heart under pressure overload. The transcription factor nuclear factor-κB (NF-κB) was activated in a subset of cardiac fibroblasts in mice subjected to pressure overload, resulting in increased expression of the messenger RNA encoding CCL2, a chemokine that recruits inflammatory immune cells. Deleting the NF-κB activator IKKβ in this subset of cardiac fibroblasts reduced the cardiac dysfunction caused by pressure overload in mice. —WW

*Sci. Signal.* **14**, eabe4932 (2021).

CELL CYCLE

## Control by RNA polymerase II

Evidence indicates that yeast cells initiate DNA synthesis and transition from the $G_1$ to the S phase of the cell cycle when cyclin 3 accumulates and causes phosphorylation of Whi5, a functional equivalent of the mammalian Rb (retinoblastoma) protein. Kõivomägi *et al*. present evidence for a different cyclin-dependent kinase target (see the Perspective by Fisher). They found that the cyclin 3–cyclin-dependent kinase (Cdk) 1 complex in yeast promoted phosphorylation of RNA polymerase II and thus increased transcription at genes that control entry into the cell cycle. Cdks that regulate the cell cycle can thus act by similar mechanisms as so-called "transcriptional Cdks," which are known to act as transcriptional regulators but not to function in control of cell division. —LBR

*Science*, aba5186, this issue p. 347; see also abm2010, p. 264

RESEARCH ARTICLE SUMMARY

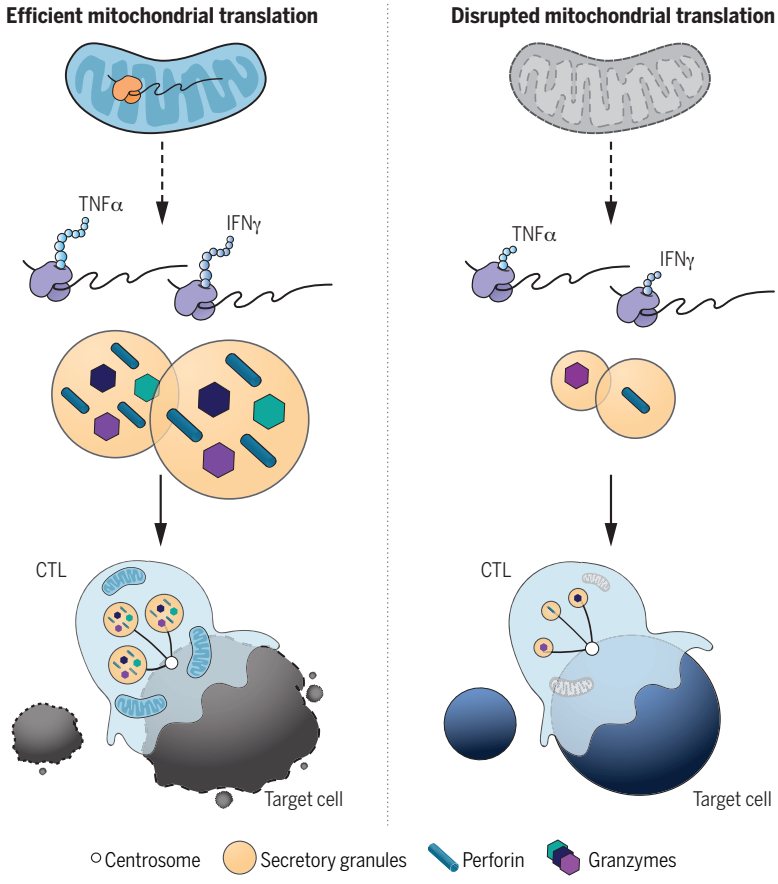
IMMUNOLOGY

# Mitochondrial translation is required for sustained killing by cytotoxic T cells

**Miriam Lisci, Philippa R. Barton, Lyra O. Randzavola, Claire Y. Ma, Julia M. Marchingo, Doreen A. Cantrell, Vincent Paupe, Julien Prudent, Jane C. Stinchcombe, Gillian M. Griffiths***

**INTRODUCTION:** Cytotoxic T lymphocytes (CTLs) are important cells within the immune system that are able to recognize and destroy cancer and virally infected cells. Killing is mediated by the release of specialized cytolytic proteins, including perforin and granzyme B, from stored secretory granules. One feature that makes CTLs particularly effective killers is their ability to carry out sustained, serial killing, with a single CTL attacking multiple targets, one after another. Although mitochondrial mass has been correlated with CTL antitumor activity, CTLs show an increased reliance on glycolysis, suggesting a decreased dependence on mitochondrial respiration. Whether, how, or why mitochondria contribute as CTLs seek, recognize, and kill their targets is not well understood.

**RATIONALE:** USP30 (ubiquitin carboxyl-terminal hydrolase 30), a deubiquitinase known to inhibit mitophagy, was identified as a regulator of CTL killing in a large-scale screen of single-gene–deleted mice. These results suggested that mitochondria may play a previously unappreciated role in CTL biology. Thus, we generated CTLs from USP30-deficient mice to study the nature of this defect and to understand how it affects killing.




**Mitochondria regulate sustained killing by CTLs.** Sustained killing of targets by CTLs requires replenishment of cytolytic proteins, including perforin and granzymes, in secretory granules and cytokines including TNF-α and IFN-γ. This is controlled through mitochondrial translation, selective disruption of which prevents refueling, leaving CTLs unable to sustain killing.

**RESULTS:** T cell development was unaffected in USP30-deficient mice. However, upon activation, $CD8^+$ T cells generated CTLs with an acute loss of mitochondria and impaired killing. The cytotoxicity of USP30-deficient CTLs diminished with time, indicating a defect in sustained killing. Despite the loss of mitochondria and decreased oxidative phosphorylation in *Usp30*$^{-/-}$ CTLs, motility, signaling, and secretion—which are required for killing—were all intact. However, we found that secretory granule size was reduced in *Usp30*$^{-/-}$ CTLs, with a reduction in newly synthesized intermediates of key cytolytic proteins, perforin and granzyme B. This suggested an underlying defect in the de novo protein synthesis, which we showed to be required for sustained killing.

Monitoring translation revealed a marked reduction in protein synthesis in *Usp30*$^{-/-}$ CTLs. Using mass spectrometry to ask whether the translational defect affected all proteins equally, we found that only a subset of cytosolic proteins was affected, including key mediators of the killing response [granzyme B, perforin, and the cytokines tumor necrosis factor–α (TNF-α) and interferon-γ (IFN-γ)]. CTLs showed no requirement for efficient oxidative phosphorylation to sustain cytotoxicity. However, selective inhibition of mitochondrial translation with doxycycline or chloramphenicol attenuated translation of cytolytic proteins, indicating an essential role for mitochondrial translation in sustained CTL killing.

How might mitochondrial translation selectively affect translation of cytolytic proteins and regulate the killing capacity of CTLs? Neither mTOR (mechanistic target of rapamycin) signaling nor activation of the integrated stress response caused loss of cytosolic translation and killing in *Usp30*$^{-/-}$ CTLs. However, there was an altered expression of metabolic enzymes that can moonlight as RNA-binding proteins (RBPs) in both USP30-depleted and doxycycline-treated CTLs. Thus, our findings suggest that post-transcriptional regulation by RBPs, a well-described phenomenon in CTLs, may mediate the selective down-regulation of protein synthesis detected upon loss of USP30.

**CONCLUSION:** Our study highlights a role for mitochondria as homeostatic regulators of CTL killing, with the production of new cytolytic proteins synchronized with mitochondrial translation. In this way, the energy requirements of protein synthesis can be fine-tuned to meet the demands of serial killing during immune challenges. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: gg305@cam.ac.uk
Cite this article as M. Lisci *et al.*, *Science* **374**, eabe9977 (2021). DOI: 10.1126/science.abe9977

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abe9977

RESEARCH ARTICLE

IMMUNOLOGY

# Mitochondrial translation is required for sustained killing by cytotoxic T cells

Miriam Lisci[1], Philippa R. Barton[1], Lyra O. Randzavola[1]†, Claire Y. Ma[1], Julia M. Marchingo[2], Doreen A. Cantrell[2], Vincent Paupe[3], Julien Prudent[3], Jane C. Stinchcombe[1], Gillian M. Griffiths[1]*

T cell receptor activation of naïve CD8$^+$ T lymphocytes initiates their maturation into effector cytotoxic T lymphocytes (CTLs), which can kill cancer and virally infected cells. Although CTLs show an increased reliance on glycolysis upon acquisition of effector function, we found an essential requirement for mitochondria in target cell–killing. Acute mitochondrial depletion in USP30 (ubiquitin carboxyl-terminal hydrolase 30)–deficient CTLs markedly diminished killing capacity, although motility, signaling, and secretion were all intact. Unexpectedly, the mitochondrial requirement was linked to mitochondrial translation, inhibition of which impaired CTL killing. Impaired mitochondrial translation triggered attenuated cytosolic translation, precluded replenishment of secreted killing effectors, and reduced the capacity of CTLs to carry out sustained killing. Thus, mitochondria emerge as a previously unappreciated homeostatic regulator of protein translation required for serial CTL killing.

CD8$^+$ cytotoxic T lymphocytes (CTLs) are key players of the adaptive immune response. Activation of T cell receptors (TCRs) on naïve CD8$^+$ T cells converts them into effector CTLs able to kill tumorigenic and virally infected cells. CTLs initiate cell death by secreting cytotoxic proteins, including perforin and granzymes, which selectively trigger caspase activation and subsequent apoptosis (*1*). An important attribute of CTLs is their ability to carry out "serial killing," with a single CTL attacking multiple targets one after another, facilitating sustained killing (*2*–*4*). Thus, CTL function relies on motility to find and engage targets, TCR signaling for activation, and regulated secretion to effect killing.

Mitochondria play crucial roles in the adaptive immune system, mediating the development, metabolism, and activation of T cells (*5*–*7*). Although mitochondrial mass is correlated with CTL antitumor activity and increases after TCR-mediated activation (*8*–*10*), paradoxically CTLs show enhanced aerobic glycolysis (*11*, *12*). How or why mitochondria affect target cell–killing by differentiated CTLs is not well understood.

The removal of mitochondria through autophagy (mitophagy) is an important contributor to mitochondrial abundance and quality control (*13*, *14*). The autophagic machinery can be recruited to mitochondria through an ubiquitin coat that forms in response to mitochondrial stress. Promiscuous ubiquitination is thought to be minimized by USP30 (ubiquitin carboxyl-terminal hydrolase 30), a transmembrane deubiquitinase (DUB) localized on the outer mitochondrial membrane (*15*–*17*). Thus, the loss of USP30-mediated deubiquitination results in increased mitophagy, reducing cellular mitochondrial content.

A large-scale screen of single-gene deletion mice for immunological function identified USP30 as a regulator of CTL killing (*18*). T cell development was normal in *Usp30*$^{-/-}$ (KO) mice, but CTLs generated after activation of naïve CD8$^+$ T cells showed reduced killing of target cells. The role of USP30, and by inference mitochondria, in CTL function was intriguing because there is an increased reliance on glycolysis in activated effector T cells (*11*). These observations raised the possibility that mitochondria play a previously unappreciated role in CTL biology and pointed to USP30 as a starting point for investigating this connection.

## Results

### *Activation leads to an acute loss of mitochondria in USP30-deficient CD8$^+$ T cells*

T cell development was unaffected in KO mice (*18*), with normal T cell populations in the spleen (fig. S1) that could then be used to generate CTLs, acquiring cytolytic activity 5 to 7 days after activation (Fig. 1A). Using tetramethylrhodamine ethyl ester (TMRE) (Fig. 1B) and MitoTracker Green (Fig. 1C) to label mitochondrial membrane potential and mass, respectively, wild-type (WT) and KO signals overlapped before activation (day 0). Upon activation, both WT and KO signals increased by day 3, which is consistent with an increase in mitochondria in activated T cells (*9*, *10*). However, by days 5 to 7, there was a loss of signal in KO but not WT CTLs, suggesting that mitochondria were selectively affected in KO CTLs after TCR activation. Using immunofluorescence microscopy, we found that by day 5 after activation, KO CTLs showed reduced levels of both the mitochondrial outer-membrane protein TOM20 as well as the mitochondrial matrix enzyme pyruvate dehydrogenase (PDH) compared with those of WT CTLs (Fig. 1, D and E, and fig. S2). Using transmission electron microscopy (TEM), mitochondria in day 5 KO CTLs exhibited absent or disorganized cristae, with an empty appearance compared with mitochondria in WT CTLs (Fig. 1F), which is consistent with a role for USP30 in safeguarding the mitochondrial import protein TOM20 (*19*). Quantitation of TEM images demonstrated that 32% of KO CTLs showed disrupted mitochondrial morphology, whereas 60% lacked mitochondria completely. Only 8% of KO CTLs showed normal mitochondrial morphology, compared with 73% of WT CTLs (Fig. 1G). Furthermore, the fluorescent, pH-dependent mitochondrial-Keima mitophagy reporter (mt-Keima) (*20*) revealed an increased mitophagic signal (red) in KO CTLs. Thus, TCR activation of CD8$^+$ T cells lacking USP30 appeared to induce substantial changes in mitochondrial number and morphology (Fig. 1, H and I, and movies S1 and S2).

Given the altered mitochondrial morphology, we asked whether mitochondrial function was affected. Oxidative phosphorylation was markedly reduced in day 5 KO CTLs, showing significant differences between both basal and maximal rates of respiration compared with those of WT CTLs (Fig. 1J). Moreover, the decrease in oxygen consumption rate (OCR) was accompanied by an increase in extracellular acidification rate (ECAR), suggesting an even greater reliance on glycolysis in KO compared with that of WT CTLs (Fig. 1K). Thus, activation of CD8$^+$ T cells lacking USP30 results in the generation of CTLs with substantial changes in mitochondrial number and morphology, which are associated with impaired function.

### *Mitochondrial depletion inhibits killing*

We found that day 5 KO CTLs exhibited a reduced killing capacity, particularly when required to kill for extended periods or when the number of CTLs was limiting. In a short-term killing assay, KO CTLs failed to kill targets until the ratio of CTL to targets was >7 and only killed 40% of target cells when CTLs outnumbered targets by 25:1. By contrast, WT CTLs killed targets within 2.5 hours when the ratio of CTL to targets was >1 and killed all targets when CTLs outnumbered targets by

[1]Cambridge Institute for Medical Research, University of Cambridge, Cambridge Biomedical Campus, CB2 0XY, UK. [2]Cell Signalling and Immunology Division, School of Life Sciences, University of Dundee, Dundee DD1 5EH, UK. [3]MRC Mitochondrial Biology Unit, University of Cambridge, Cambridge Biomedical Campus, CB2 0XY, UK.
*Corresponding author. Email: gg305@cam.ac.uk
†Present address: Department of Immunology and Inflammation, Imperial College London, London W12 0NN, UK.

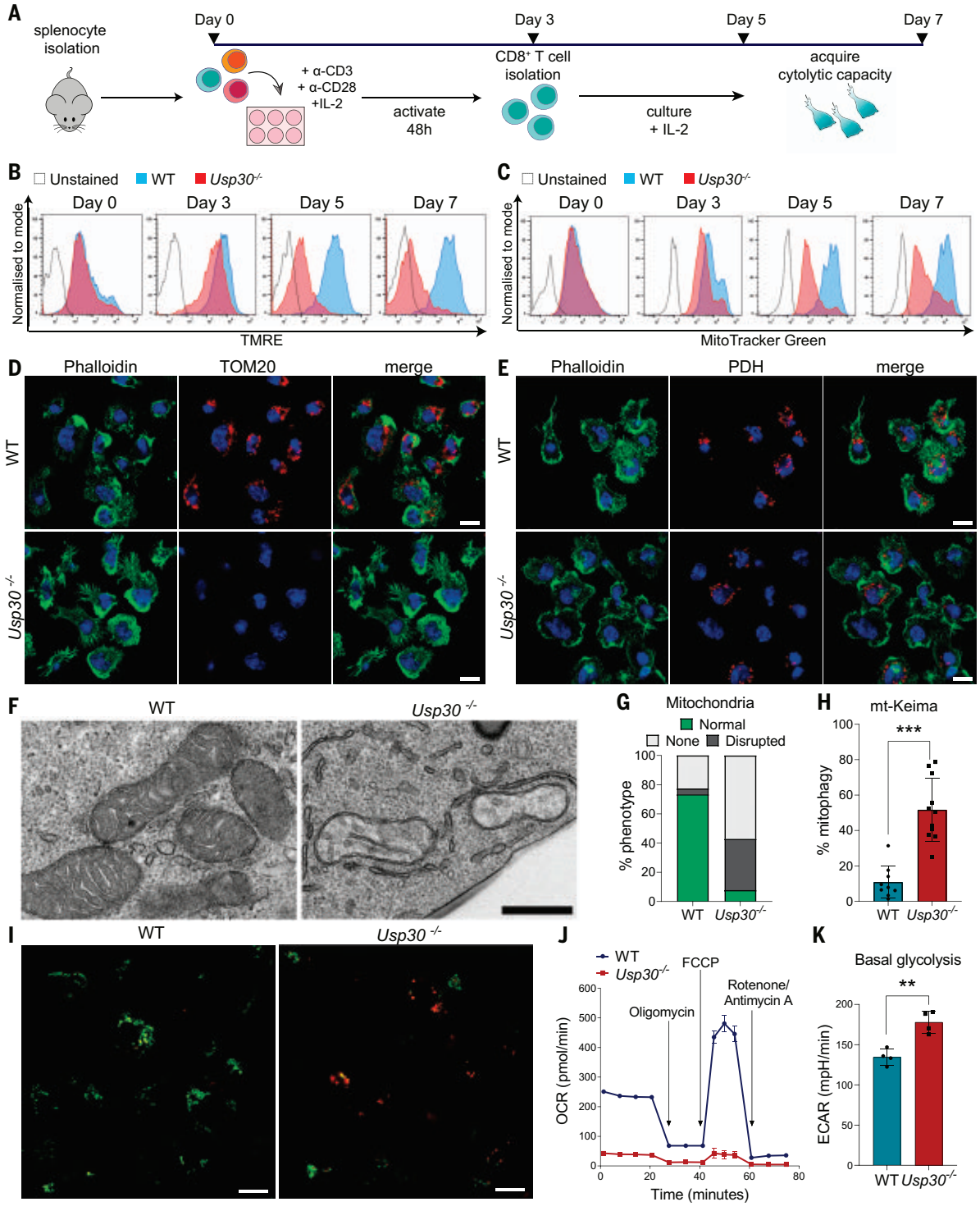



**Fig. 1. *Usp30* deletion leads to mitochondrial depletion in effector CD8$^+$ T cells.** (**A**) Isolation and stimulation of mouse primary CD8$^+$ T cells. (**B** and **C**) Time course of mitochondria labeled with (B) TMRE and (C) MitoTracker Green in WT and KO naïve CD8$^+$ T cells (day 0) as they mature into CTLs after stimulation (day 3 to 7). (**D** and **E**) Immunofluorescence staining of (D) TOM20 (outer mitochondrial membrane) and (E) PDH (mitochondrial matrix) in day 5 WT and KO CTLs. Scale bars, 10 μm. (**F**) TEM sections (70 nm) of mitochondria in day 5 WT and KO CTLs. Scale bars, 500 nm. (**G**) Quantitation of TEM images from 128 WT and 154 KO cell profiles. (**H**) Quantitation of live cell imaging of day 5 WT and KO CTLs expressing mt-Keima. Each data point shows percentage of mitolysosomes (red mt-Keima signal) in a 63X field, indicative of mitophagy (>40 cells per genotype in each independent repeat). (**I**) Representative images of mt-Keima signal in day 5 WT and KO CTLs. Scale bars, 15 μm. (**J**) Mitochondrial respiratory capacity is indicated by OCR in day 5 WT and KO CTLs. (**K**) Basal glycolysis in day 5 WT and KO CTLs measured with ECAR values at resting state. $^{**}P < 0.01$, $^{***}P < 0.001$ (two-tailed unpaired Student's *t* test). Error bars show mean values ± SD of technical replicates in one representative experiment. All data are representative of at least three independent biological replicates.

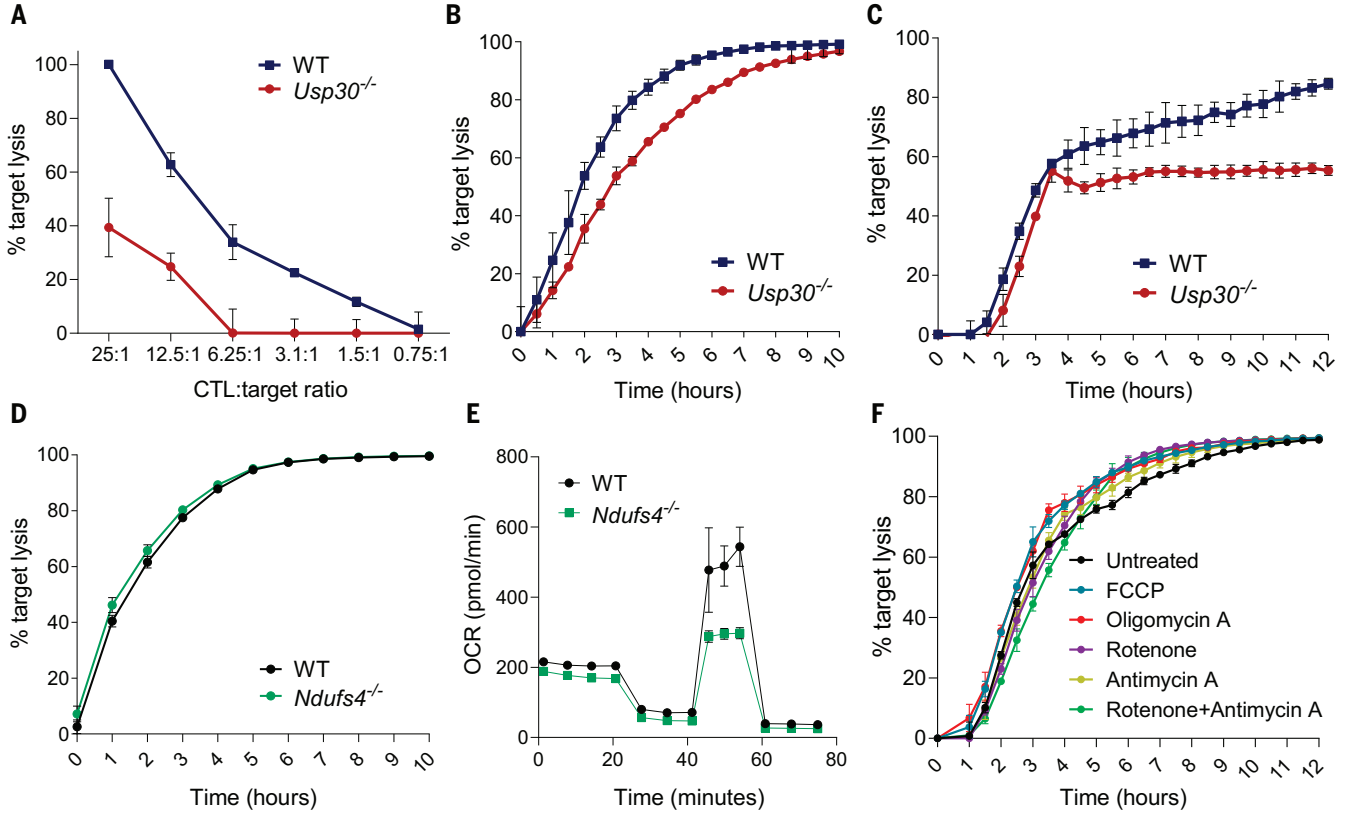



**Fig. 2. Mitochondrial depletion leads to oxidative phosphorylation–independent inhibition of killing.** (**A**) Short-term (2.5 hours) killing assay with day 5 WT and KO CTLs, varying CTL-to-target ratio. (**B**) Long-term (10 hours) killing assay showing percentage of target cell lysis over time (CTL-to-target ratio, 10:1) with day 5 WT and KO CTLs. (**C**) Long-term (12 hours) killing assay showing percentage of target cell lysis over time (CTL-to-target ratio, 1:1) with day 5 WT and KO CTLs. (**D**) Long-term (10 hours) killing assay showing percentage of target cell lysis over time (CTL-to-target ratio, 10:1) with day 5 WT and *Ndusf4*$^{-/-}$ CTLs. (**E**) Mitochondrial respiratory capacity shown by OCR in day 5 WT and *Ndufs4*$^{-/-}$ CTLs. (**F**) Long-term (12 hours) killing assay showing percentage of target cell lysis over time (CTL-to-target ratio, 1:1), with day 5 WT CTL treated with inhibitors used in the Seahorse assay: Complex I (rotenone), III (antimycin A), V (oligomycin), and the uncoupling reagent, FCCP. Errors bars indicate mean values ± SD of (A) to (C) and (F) technical replicates, (D) and (E) biological replicates. All data are representative of at least two independent biological replicates.

25:1 (Fig. 2A). Furthermore, a longer-term assay with 10 CTLs per target showed that KO CTLs took almost twice as long as WT CTLs to kill all targets (Fig. 2B), demonstrating a reduced killing capacity in KO CTLs. Reducing the number of CTLs per target to 1:1 revealed similar levels of killing for WT and KO CTLs over the first 4 hours, after which KO CTLs showed no further killing, suggesting that KO CTLs lost the capacity to kill over time (Fig. 2C).

### *ATP production, motility, signaling, and secretion are unaffected by loss of mitochondria*

Mitochondria could contribute to many of the different steps required for killing as CTL seek, recognize, and kill their targets. Therefore, we sought to identify which were affected in KO CTLs. First, we asked whether reduced oxidative phosphorylation affected killing by testing day 5 CTLs from *Ndufs4*$^{-/-}$ mice in which complex I was disrupted (*21*). Despite a defect in maximal mitochondrial respiration, there was no defect in T cell development in *Ndufs4*$^{-/-}$ mice (fig. S3) or in CTL killing by using either this genetic model; inhibitors of complex I, III, and V; or disruption of mitochondrial membrane potential (Fig. 2, D to F). Because mitochondrial adenosine 5′-triphosphate (ATP) has been shown to play a role in lymphocyte migration (*22*), we asked whether motility or ATP levels were affected in day 5 KO CTLs. We found no difference in motility (Fig. 3A and movies S3 and S4) nor overall ATP concentrations between WT and KO CTLs, with ATP concentrations reduced in both WT and KO upon selective inhibition of oxidative phosphorylation (oligomycin) or glycolysis [2-deoxy-D-glucose (2-DG)] (Fig. 3B). We found no defect in conjugate formation between day 5 KO CTLs and target cells or in polarization of the centrosome or cytolytic granules to the synapse formed between CTLs and targets (Fig. 3, C and D). Furthermore, day 5 KO CTLs showed no defects in either TCR-stimulated calcium fluxes nor the release of granules in response to TCR activation (Fig. 3, E and F). Proximal and distal signaling events driven by TCR engagement also showed no difference between day 5 WT and KO CTLs (fig. S4). Thus, ATP levels, motility, polarization, signaling, and secretion were all unaffected in KO CTLs.

### *Mitochondrial depletion leads to translation attenuation and loss of cytolytic proteins*

Given that all steps to secretion appeared to be intact, we asked whether the secretory granules that store the cytolytic proteins required for killing were affected in day 5 KO CTLs. There were no gross changes in the morphology of KO compared with WT granules. However, although similar in number, the granules were smaller (quantitated by reduced surface area) in KO CTLs (Fig. 4, A to C). There was no evidence of granule autophagy, suggesting that the reduced size might rather reflect a reduction in granule content. Two of the key cytolytic proteins stored in granules are granzyme B and the pore-forming protein perforin (*1*). Granzyme B expression in KO CTLs was reduced relative to that in WT CTLs at days 5 and 7 after activation (Fig. 4D). Furthermore,

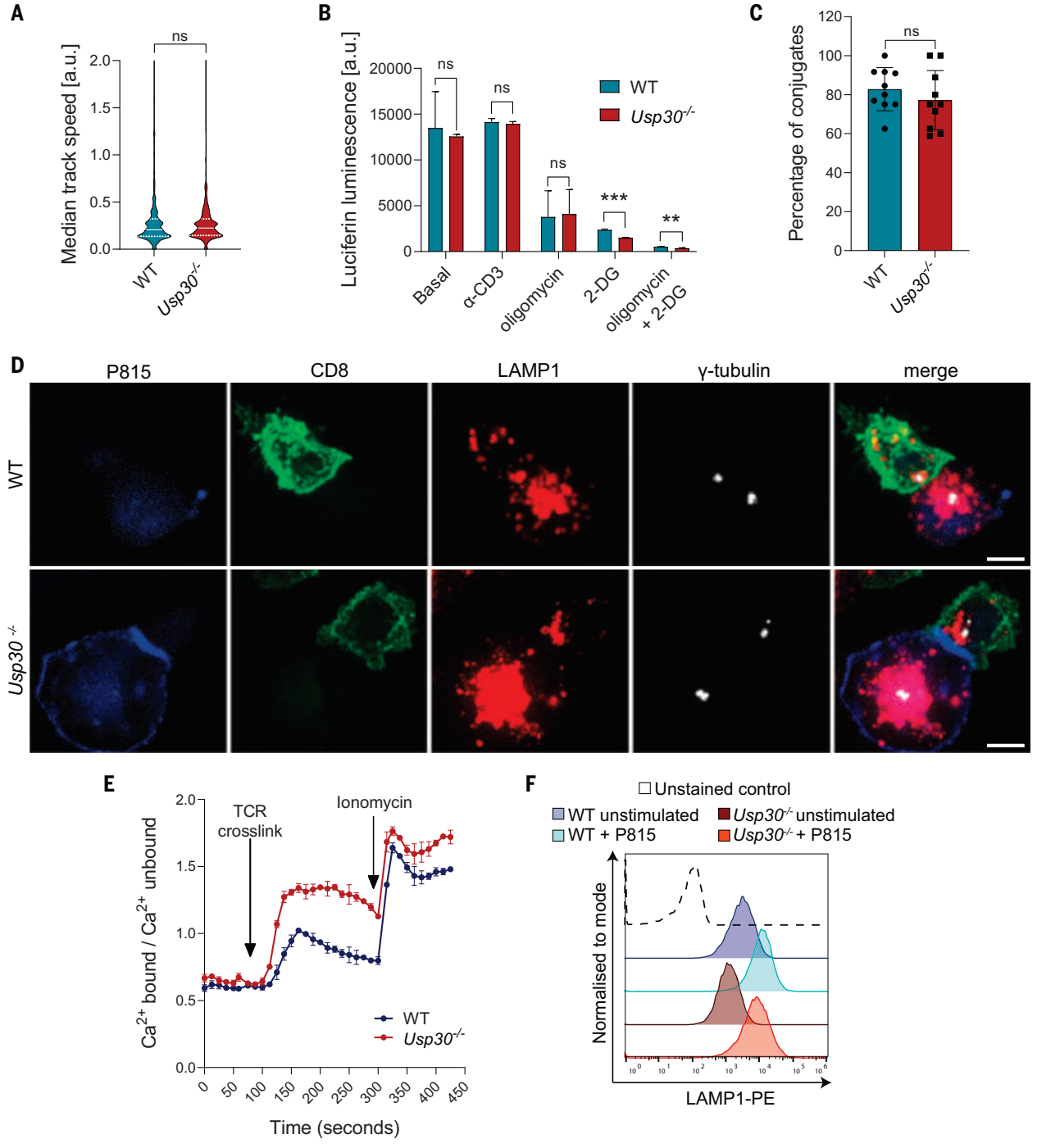



**Fig. 3. Loss of USP30 does not impair ATP production, migration, target recognition or secretion.** (**A**) Track speed analysis of migrating day 5 WT and KO CTLs (>200 cells per genotype per repeat; solid lines, median; dotted lines, quartiles). (**B**) ATP quantitation in day 5 WT and KO CTLs, with (anti-CD3) or without (basal) TCR activation or treated with oligomycin and/or 2-DG to account for mitochondria-derived and glycolysis-derived ATP respectively. (**C**) Quantitation of conjugate formation between day 5 WT and KO CTLs and P815 target cells (>70 cells per genotype per independent repeat). Each dot indicates the average percentage of conjugates in a 63X field. (**D**) Conjugates formed by day 5 WT or KO CTLs (green) with P815 targets (blue) showing polarization of cytolytic granules (red) and centrosome (white). Scale bars, 5 μm. (**E**) INDO-1 calcium assay showing changes in intracellular calcium concentration in day 5 WT and KO CTLs before [time ($t$) = 0 to 60 s] and after TCR crosslinking ($t$ = 60 s) or ionomycin addition ($t$ = 300 s). (**F**) Histograms representing surface LAMP1-PE fluorescence intensity in day 5 WT and KO CTLs before and after incubation with P815 target cells. a.u., arbitrary unit; ns, not significant; $^{**}P < 0.01$, $^{***}P < 0.001$ (unpaired two-tailed Student's $t$ test). All error bars show mean values ± SD of technical replicates. Data are representative of three independent biological replicates.

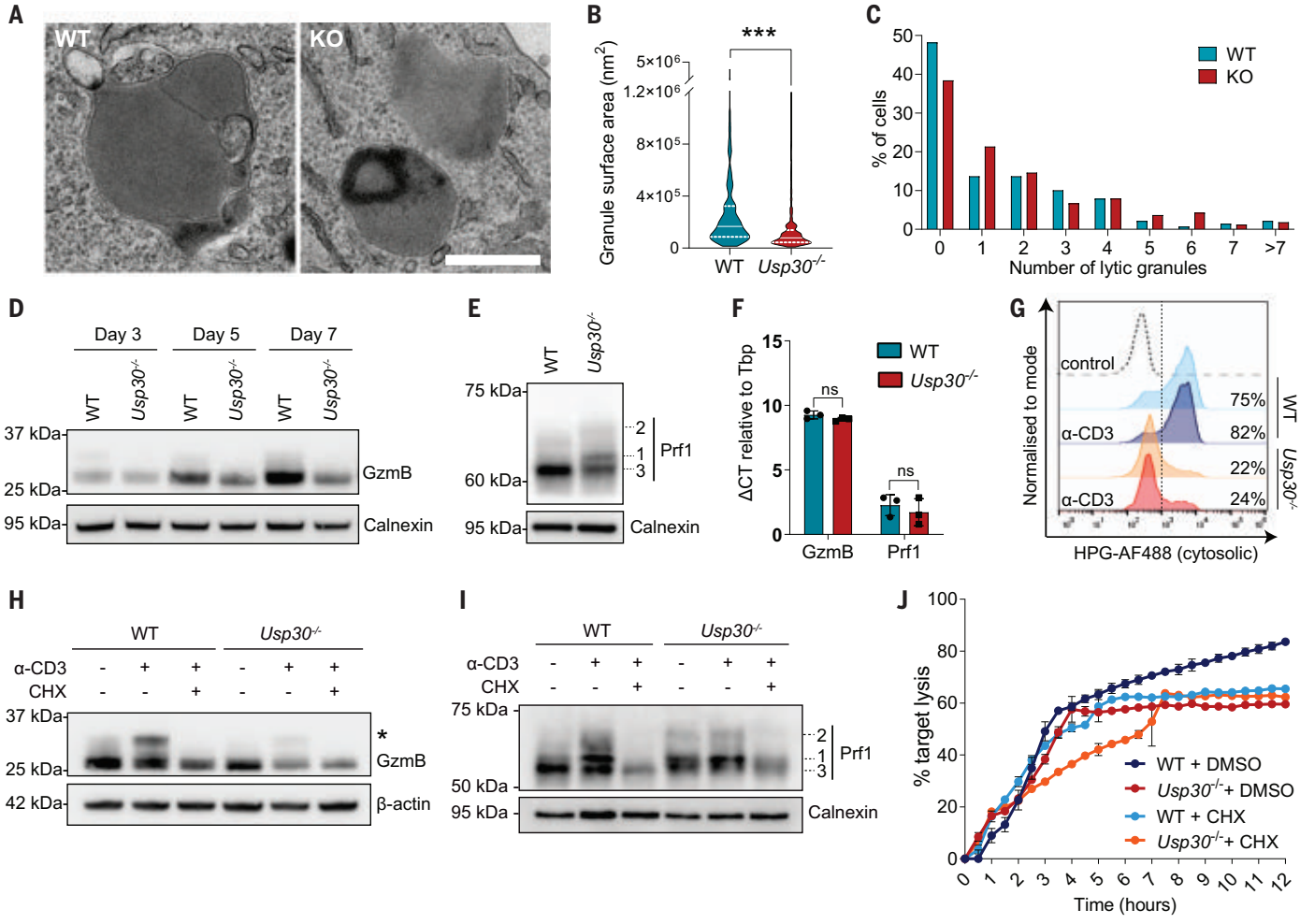



**Fig. 4. Mitochondrial depletion leads to translation attenuation and reduced expression of key cytolytic proteins.** (**A**) TEM sections (70 nm) of lytic granules in day 5 WT and KO CTLs. Scale bar, 500 nm. (**B** and **C**) Quantitation of (B) cross-sectional surface areas, and (C) lytic granule number (solid line, median; dotted lines, quartiles) from 139 (WT) and 164 (KO) profiles. ***$P < 0.001$ (two-tailed unpaired Student's $t$ test). (**D** and **E**) Immunoblots showing (D) granzyme B (GzmB) and (E) perforin (Prf1) in WT and KO CTLs. Immature (1), intermediate (2), and mature (3) forms of perforin are labeled. (**F**) GzmB and Prf1 mRNA in day 5 WT and KO CTLs. ΔCT is normalized to TATA-binding protein mRNA (Tbp). Error bars indicate mean values ± SD of three biological replicates. ns, not significant (two-tailed paired Student's $t$ test). (**G**) HPG-AF488 incorporation (1 hour) into day 5 WT and KO CTLs, ± anti-CD3 (TCR activation) where shown, or CHX-treated CTLs in which cytosolic translation is inhibited (control). (**H** and **I**) Immunoblots showing (H) GzmB and (I) Prf1 in day 5 WT and KO CTLs ± anti-CD3 or CHX. Newly synthesized GzmB is indicated with an asterisk. (**J**) Long-term (12 hours) killing assay with day 5 WT and KO CTLs (CTL-to-target, 1:1) ± CHX. Error bars show mean values ± SD of technical replicates in one representative experiment. All data are representative of at least three independent biological replicates.

although perforin was rapidly processed to its mature active form in WT CTLs, incompletely processed forms were evident in day 5 KO CTLs with a reduced level of mature perforin (Fig. 4E) (*23*). Loss of protein was not due to changes in transcription because mRNA expression was unchanged in day 5 KO relative to WT CTLs (Fig. 4F), suggesting that a defect in protein synthesis may underlie the loss of cytolytic proteins in the KO.

Protein synthesis, monitored through incorporation of the methionine analog homopropargylglycine (HPG) (*24*), was markedly reduced in day 5 KO CTLs relative to WT (Fig. 4G). Mimicking target encounter by cross-linking TCR (using anti-CD3ε) led to a modest increase in protein synthesis. This included renewed synthesis of granzyme B and perforin with newly synthesized intermediates of higher molecular weight (*25*), appearing in day 5 WT CTLs that could be abolished by treatment with the cytosolic protein synthesis inhibitor cycloheximide (CHX) (Fig. 4, H and I). By contrast, new synthesis was much reduced in the KO, suggesting that a defect in mRNA translation of cytolytic proteins might cause the loss of killing observed in KO CTLs. To test this, we treated day 5 WT and KO CTLs with CHX in a long-term killing assay that required sustained killing. WT CTLs killed for the first 4 hours, after which CHX treatment reduced cytotoxicity, mirroring the pattern of killing by KO CTLs (Fig. 4J). In KO CTLs treated with CHX, target cell–killing showed a further decrease between 2 and 8 hours, which is consistent with the low levels of cytosolic protein synthesis observed. Thus, de novo protein synthesis is required for sustained CTL killing and is impaired in KO CTLs.

### *Cytolytic proteins are selectively affected by the translational defect in USP30-deficient CTLs*

To determine whether the translational defect affected all proteins equally, we used mass spectrometry to compare the proteomes of day 5 WT and KO CTLs before and after TCR activation. Mass spectrometry data was normalized to the number of cells by using a

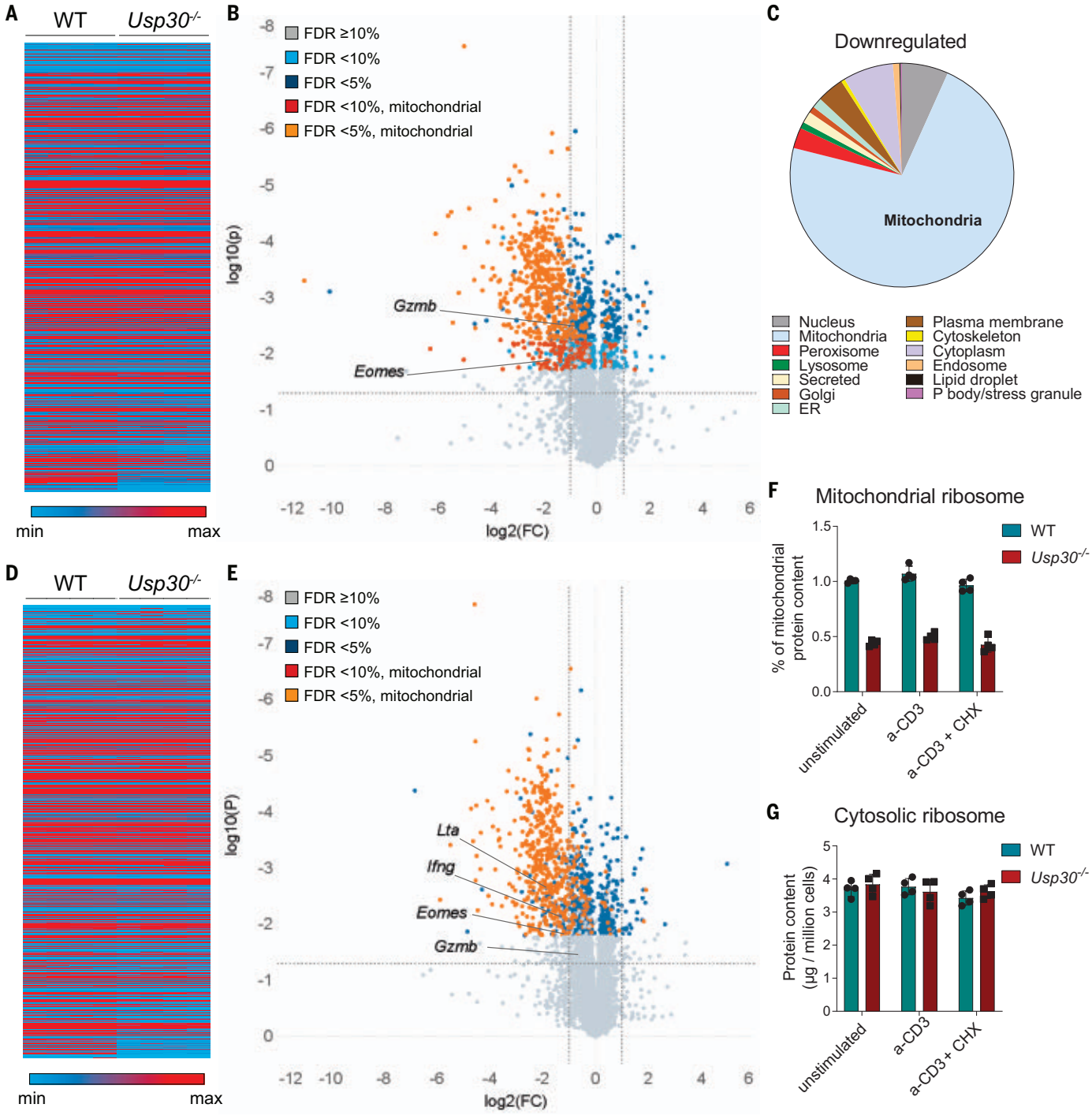



**Fig. 5. Proteomic analysis of mitochondrially depleted KO CTLs.** (**A**) Heat map showing global protein expression (protein copy number) in day 5 WT and KO CTLs. (**B**) Volcano plot of up-regulated (right quadrant) and down-regulated (left quadrant) proteins in day 5 KO CTLs compared with WT. Vertical bars denote $\log_2$(FC) = 1,–1, and the horizontal bar corresponds to the uncorrected $P = 0.05$ threshold. Colored dots indicate hits within the 5 and 10% FDR range. Gene Ontology Cellular Component–defined mitochondrial proteins are highlighted according to FDR. (**C**) Cellular localization of down-regulated proteins in day 5 KO CTLs (5% FDR). (**D**) Heat map displaying global protein expression (protein copy number) in anti-CD3–stimulated (TCR activated, 4.5 hours) day 5 WT and KO CTLs. (**E**) Volcano plot of up-regulated and down-regulated proteins in anti-CD3–stimulated day 5 KO CTLs compared with that of WT. Quadrants and color scheme are as in (B). Highlighted proteins are referenced by gene names in (B) and (E). (**F**) Percentage of mitochondrial ribosomal protein out of total mitochondrial peptides detected in unstimulated, anti-CD3–treated, and CHX-treated day 5 CTLs. (**G**) Cytosolic ribosome protein expression in unstimulated, anti-CD3–treated, and CHX-treated day 5 CTLs. Error bars indicate mean values ± SD from biological replicates [(F) and (G)]. All data are representative of four biological replicates.

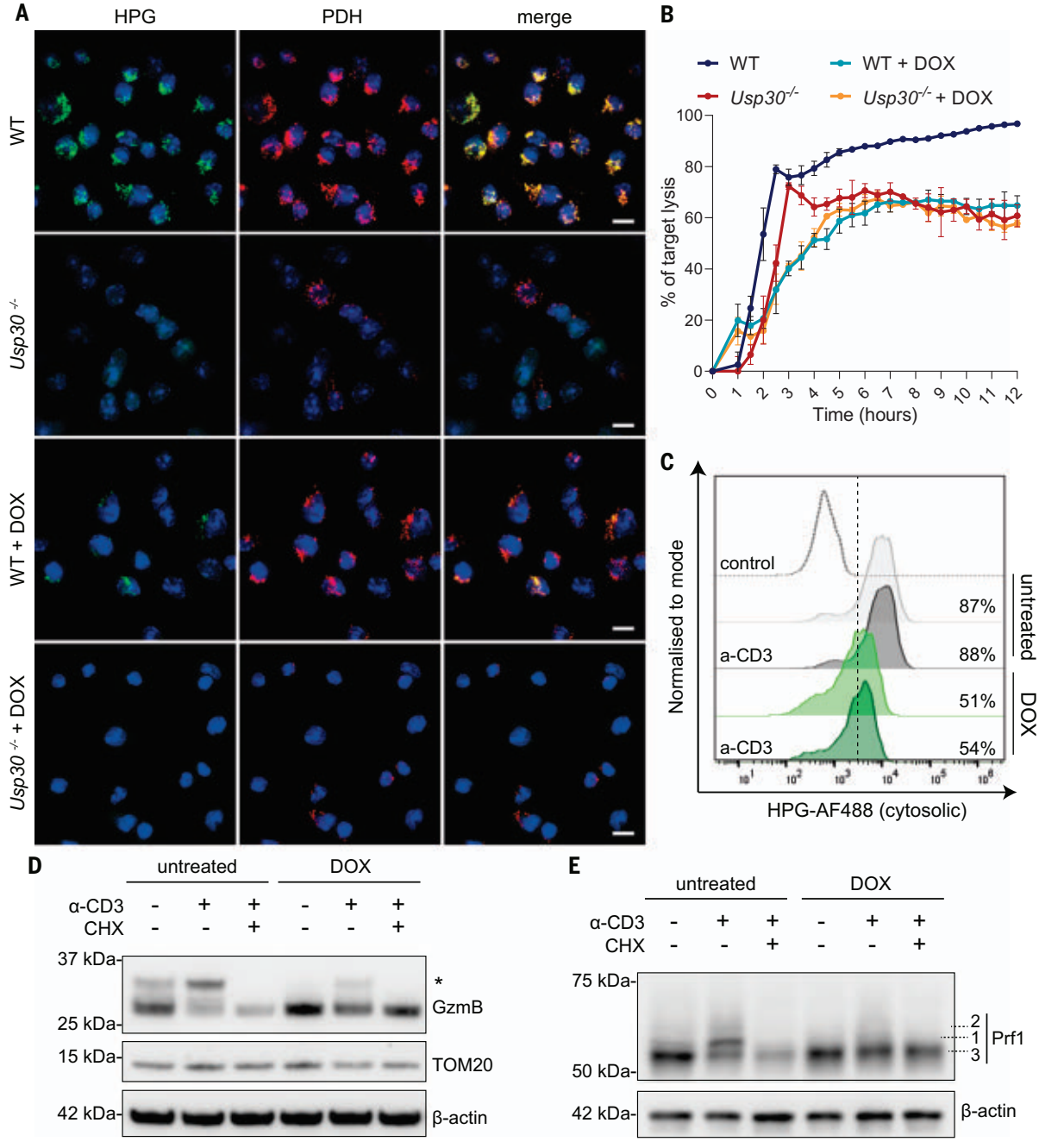



**Fig. 6. Mitochondrial translation is required for sustained T cell killing.** (**A**) Immunofluorescence imaging of mitochondrial translation by using 1-hour HPG-AF488 incorporation into day 5 WT and KO CTLs treated with CHX to inhibit cytosolic translation. CTLs were either left untreated or treated with 10 μg/ml of DOX to selectively inhibit mitochondrial translation as indicated. Scale bars, 7 μm. (**B**) Long-term (12 hours) killing assay (CTL-to-target ratio, 1:1) showing percentage target lysis with day 5 WT and KO CTLs treated with DOX. Error bars indicate mean values ± SD of technical replicates. (**C**) HPG-AF488 incorporation into day 5 WT CTLs treated for 1 hour with DOX to inhibit mitochondrial translation or CHX to inhibit cytosolic translation (control), stimulated with anti-CD3 (TCR) where shown. (**D** and **E**) Immunoblots showing (D) GzmB and (E) Prf1 in day 5 WT CTLs treated with DOX and/or stimulated through TCR activation (anti-CD3) as indicated for 4 hours. Newly synthesized GzmB is indicated with an asterisk. Immature (1), intermediate (2), and mature (3) forms of perforin are indicated. Error bars indicate mean values ± SD of technical replicates in one representative experiment. All data are representative of at least three independent biological replicates.

histone ruler approach in order to estimate protein mass and copy number per cell. Analysis of these data not only confirmed the absence of USP30 (table S1) but also revealed that the decrease in protein synthesis did not affect the abundance of all proteins equally. Only a subset of proteins (557) was found to have a significantly lower expression in KO CTLs [fold change > 2, false discovery rate (FDR) < 10%] (Fig. 5, A to C). Most affected were mitochondrial proteins (72%) (Fig. 5, B and C, and table S1), presumably reflecting their degradation by promiscuous mitophagy in the absence of USP30. Other highly down-regulated proteins included granzyme B and

eomesodermin, a transcription factor that regulates expression of cytotoxic proteins, including perforin, granzyme B, and interferon-γ (IFN-γ) (Fig. 5B and table S1) (*26*). A smaller number of proteins (35) were found to be significantly up-regulated in KO CTLs (fold change > 2, FDR < 10%), including 11 proteins encoded from genes linked to the integrated stress response (ISR) (*Cebpb*, *Hmox1*, *Avil*, *Bcl2l11*, *Slca9*, *Niban1*, *Ifrd1*, *Cth*, *Bbc3*, *Ca2*, and *Plin2*) and one involved in autophagy (*p62*/*Sqstm1*) (Fig. 5B and table S1). Upon TCR activation of CTLs, additional cytolytic proteins were down-regulated in KO CTLs (Fig. 5D and table S2), including IFN-γ, tumor necrosis factor–α (TNF-α), and TNF-β (Fig. 5E, fig. S5, and table S2). Thus, in addition to the loss of mitochondrial proteins, cytolytic proteins were particularly affected, decreasing the ability of KO CTLs to mount an effective response. Mitochondrial ribosomal proteins represented one of the most down-regulated subsets within the mitochondrial compartment both before and after TCR stimulation (Fig. 5F). Cytosolic ribosomes were unaffected, indicating that the translation attenuation observed in the KO CTLs was not caused by a defect in cytosolic ribosomal content (Fig. 5G). Thus, mitochondria may be modulating the killing capacity of CTLs by regulating levels of cytolytic proteins.

### *Mitochondrial translation is required for sustained CTL killing*

It is now well established that mitochondrial translation can regulate the synthesis of cytosolic proteins (*27*, *28*). In addition, recent in vivo studies suggest roles for mitochondrial translation in T cell function (*29*, *30*). We therefore asked whether mitochondrial protein synthesis was disrupted in day 5 KO CTLs by examining HPG incorporation when cytosolic translation was inhibited by CHX (Fig. 6A) (*31*). Mitochondrial translation was evident in WT CTLs, with HPG colocalizing with the mitochondrial marker, PDH. By contrast, there was little HPG incorporation in KO CTLs (Fig. 6A). To directly address the role of mitochondrial protein translation in CTL killing, we disrupted mitochondrial protein translation pharmacologically using doxycycline (DOX) (Fig. 6). DOX treatment inhibited CTL cytotoxicity after 4 hours, which replicated the phenotype of KO CTL. There was a small additive effect of DOX treatment in KO CTLs, which is consistent with the inhibition of any remaining mitochondria in KO CTLs (Fig. 6B). Furthermore, cytosolic protein translation was inhibited after 1 hour of DOX treatment (Fig. 6C), and de novo synthesis of granzyme B and perforin after TCR activation (4 hours) was also reduced (Fig. 6, D and E). These results were recapitulated with chloramphenicol inhibition of mitochondrial protein synthesis (fig. S6). Thus, the selective inhibition of mitochondrial translation is sufficient to attenuate selective cytosolic protein translation and to impair CTL cytotoxicity.

## Discussion

Inhibition of mitochondrial translation is known to affect cytosolic translation by way of several pathways, including mechanistic target of rapamycin (mTOR) and the ISR (*27*, *28*). However, these pathways do not appear to be causative in the mitochondrial regulation of CTL killing. The mTOR target and ribosomal protein S6 was phosphorylated in both KO and WT CTLs (fig. S4). In addition, although we identified up-regulation of some ISR proteins in KO CTLs, the ISR inhibitor (ISRIB) only partially restored protein synthesis and was unable to restore killing in KO CTLs (fig. S7). Two distinctive features of the KO CTLs were decreased oxidative phosphorylation (and increased reliance on glycolysis) (Fig. 1, J and K) together with selectivity in the down-regulation of cytosolic protein synthesis (Fig. 5). Our proteomic analyses suggested an alternative mechanism that may contribute to the loss of killing when mitochondrial disruption triggers changes in metabolic pathways in the cell. Many metabolic enzymes can "moonlight" as RNA-binding proteins (RBPs), which regulate the translation of selected mRNAs (*32*). The moonlighting functions of RBP metabolic enzymes can be induced by changes in CTL metabolism and can regulate expression of cytolytic proteins, including granzyme B, TNF-α, and IFN-γ (*33*–*35*). There was a selective down-regulation of proteins in KO CTLs compared with CHX-treated CTLs (fig. S8A and tables S1 and S3). Moreover, the expression of RBP enzymes was altered in both KO CTLs and CTLs in which mitochondrial translation was inhibited (fig. S8, B to D, and tables S1 and S4). Using Gene Ontology (GO) enrichment analysis, we observed that 68% of commonly down-regulated proteins in KO and DOX-treated CTLs were involved in cellular metabolism. We propose that this metabolic rewiring leads to the selective decrease in proteins required for sustained killing when mitochondrial translation is impaired.

Our study highlights a role for mitochondria as homeostatic regulators of CTL killing. Early in an immune response, antigen-specific CTLs will be limiting, and the ability of individual CTLs to carry out sustained, serial killing will be critical to the success of the response. Although previous studies have shown that the translational landscape of $CD8^+$ T cells is especially active as they proliferate and acquire effector capacity (*24*), we found that de novo protein synthesis is also required to sustain ongoing killing. Given the high energy demands required for protein synthesis (*36*), CTLs that need to continue killing have to rapidly rebalance their overall energy needs. We found that this is mediated by mitochondria with the production of new cytolytic proteins synchronized with mitochondrial translation. In this way, mitochondria act as homeostatic regulators in times of high metabolic demand and serve a crucial function in sustained CTL killing.

## Materials and methods

(Reagents used in this study are listed in table S5)

### *Mice*

WT C57BL/6N and *Usp30*$^{-/-}$ ($Usp30^{tm2b(EUCOMM)Hmgu}$) mice were bred and housed in the University of Cambridge Facility. This research has been regulated under the Animals (Scientific Procedures) Act 1986 Amendment Regulations 2012 following ethical review by the University of Cambridge Animal Welfare and Ethical Review Body (AWERB). Age- and gender-matched animals of 8-20 weeks of age were used in experiments. Spleens from *Ndufs4*$^{-/-}$ mice were supplied by C.Viscomi and M. Zeviani (University of Padua) from mice at 4 weeks of age.

### *Cell culture, and mitochondrial inhibitors*

Single-cell suspension of splenocytes, isolated using 70-μm cell strainers, were activated on plates coated with 0.5 μg/ml of anti-mouse CD3ε and 1 μg/ml of CD28 in CTL culture medium (RPMI 1640, 10% FBS, 2 mM glutamine, 1 mM sodium pyruvate, 100 U/ml of penicillin, 1 mg/ml of streptomycin, 50 μM β-mercaptoethanol, and 100 U/ml of mouse IL-2) for 48 hours at 37°C, and grown further in CTL culture medium. On day 3, $CD8^+$ T cells were purified following manufacturer's instructions for the $CD8a^+$ T Cell Isolation Kit and cultured in CTL medium. P815 and P815-blue (*37*) target cells were grown in DMEM, 10% FBS, penicillin-streptomycin. P815-NucLight transduced cells (*38*) were maintained under 1 μg/ml puromycin selection.

Mitochondrial translation was inhibited by supplementing the cell culture media with 10 μg/ml of doxycycline or 50-500 μg/ml of chloramphenicol throughout the killing (12 hours) and translation (1 hour) assays. CTLs used for immunoblots were treated with 10 μg/ml of doxycycline or 500 μg/ml of chloramphenicol for 4.5 hours before cell lysis. The ISR was inhibited using 100 nM ISRIB throughout killing and translation assays. Mitochondrial respiration was inhibited in the Mito Stress (Seahorse) and killing (IncuCyte) assays using 1 μM FCCP, 1 μM oligomycin A, 0.5 μM rotenone, and 0.5 μM antimycin A.

### *Killing assays*

CTLs (5 days after activation) were plated at the CTL-to-target ratios shown in phenol red-free

RPMI 1640 (2% FBS, 1% penicillin, and streptomycin) in round-bottomed 96-well plates with $10^4$ P815 per well in the presence of 1 μg/ml of anti-mouse CD3ε. Cultures were then incubated at 37°C, 10% $CO_2$ for 2.5 hours (short-term killing). The percentage of target cell lysis was determined using the CytoTox 96 assay following manufacturer's instructions.

Long-term killing assays were imaged in the IncuCyte S3 Live Cell Analysis System (Sartorius) that contains a Basler Ace 1920-155um compact camera and 4X objective, resolution 2.82 microns/pixel. CTLs (5 days after activation) were mixed with P815-NucLight targets in CTL culture medium with 1 μg/ml of anti-mouse CD3ε at the CTL-to-target ratios shown in ultra-low attachment round-bottomed 96-well plates (Corning, 7007). Plates were scanned using a 4X objective lens in both the bright-field and red (excitation: 655 nm; emission: 681 nm) channels and analyzed using IncuCyte S3 v2018-2021 software (Sartorius) to determine loss of target cells over time based on decrease of red fluorescence intensity.

### Granule release assay

CTLs (5 days after activation) were incubated with anti-CD107a-PE (LAMP1) in CTL culture medium in the presence or absence of P815 target cells with 1 μg/ml of anti-mouse CD3ε for 2.5 hours at 37°C, 10% $CO_2$. Cells were resuspended in ice-cold FACS buffer (PBS, 10% FBS), labeled for 20 min at 4°C with Brilliant Violet 711-conjugated anti-CD8 and Zombie Yellow, washed, and analyzed on an Attune NxT (ThermoFisher).

### Calcium flux assay

CTLs (5 days after activation) were incubated with 1 μM INDO-1 for 30 min at 37°C, washed and incubated with 1 μg/ml of APC-conjugated anti-CD8 and 5 μg/ml of anti-CD3ε for 15 min at room temperature. CTLs were resuspended in CTL culture medium and basal INDO-1 MFI was recorded for 1 min on an LSR Fortessa (Becton Dickinson). INDO-1 was excited at 355 nm and emission was collected at 450 nm ($Ca^{2+}$-unbound) and at 379 nm ($Ca^{2+}$-bound). TCR was crosslinked by addition of 50 μg/ml of goat anti-hamster antibodies and subsequent calcium fluxes were recorded for 5 min, after which 1 μg/ml of ionomycin was added to determine maximal calcium flux for 2 min. The MFI ratio of $Ca^{2+}$-bound to $Ca^{2+}$-unbound INDO-1 (379/450) was calculated to display changes in intracellular calcium concentration.

### SDS-PAGE and immunoblotting

Cells were lysed in RIPA buffer (50 mM Tris HCl pH 8.0, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, and 0.1% SDS) at $2\times10^7$ cells/ml. Lysates were resolved using 4 to 12% Bis-Tris or 7.5% SDS gels for perforin, before transfer to nitrocellulose filters. Membranes were blocked in 5% milk or 5% BSA (as indicated by antibody manufacturer) for 1 hour at room temperature before overnight incubation at 4°C with primary antibodies against β-actin, calnexin, granzyme B, perforin, TOM20, or phospho-S6 (Ser240/244). Membranes were washed in PBS, 0.1% Tween, incubated with goat anti-rabbit, mouse or rat HRP-conjugated antibodies for 1 hour in 5% milk or 5% BSA before washing and developing using ECL prime detection reagents (GE Healthcare, RPN2232) and imaged using a ChemiDoc MP Imaging System (Bio-Rad).

### RNA extraction and real-time quantitative PCR

Total RNA was extracted from CTLs (5 days after activation) using the RNeasy mini kit and QIAshredder columns and 100 ng per sample was used to synthesize cDNA using the Accuscript high fidelity first strand cDNA synthesis kit. Granzyme B and perforin gene expression were measured using the TaqMan Gene Expression Assay with quantitative PCR probes for mouse transcripts of granzyme B, perforin, and TATA-binding protein (Tbp) as a control, using the 7500 Real-Time PCR System (ThermoFisher). mRNA abundance was quantified by calculating the $C_T$ of the target of interest and normalizing to Tbp ($\Delta C_T$).

### Measurement of mitochondrial mass and membrane potential

Mouse CTLs were labeled with Zombie Yellow and Brilliant Violet 711-conjugated anti-CD8a for 15 min at room temperature, washed, and then incubated with 100 nM MitoTracker Green and 100 nM TMRE for 30 min at 37°C, 10% $CO_2$. Samples were analyzed on day 0 (naïve T cells) and days 3, 5, and 7 post activation using an LSR Fortessa.

### Immunofluorescence imaging

CTLs (5 days after activation) were washed in serum-free RPMI 1640 and allowed to adhere to slides at 37°C, 10% $CO_2$ for 10 min before fixation with 5% paraformaldehyde (PFA) for 15 min at 37°C and quenching in 50 mM ammonium chloride at room temperature. Cells were permeabilized for 10 min in either 0.1% Triton X-100 for PDH staining, or 0.1% saponin for TOM20 staining and blocked using PBS, 1% BSA for PDH staining, or 1% BSA, 0.1% saponin for TOM20 staining. Cells were labeled using anti-TOM20 or PDH for 2 hours washed in PBS, 1% BSA (PDH) or PBS, 1% BSA, 0.1% saponin (TOM20), incubated with goat anti-mouse Alexa Fluor (AF) 546 for 1 hour, washed in PBS, Hoechst 33342 and mounted in ProLong Diamond. Samples were imaged using a Zeiss 63X 1.4NA PlanApochromat objective on a Zeiss LSM 880 confocal microscope used in Airyscan mode and processed using Zen Black software (Zeiss).

CTL-target conjugates were formed by mixing equal numbers of day 5 post-activation CTLs and P815-blue target cells in the presence of 1 μg/ml of anti-CD3ε for 20 min at 37°C, 10% $CO_2$. Cells were allowed to adhere on slides, fixed in ice-cold methanol for 5 min, blocked in PBS, 1% BSA and incubated with anti-γ-tubulin, anti-LAMP1 (supernatant, 1D4B), and Alexa Fluor 488-conjugated anti-CD8. After washing cells were labeled with Alexa Fluor 546-conjugated anti-mouse IgG and Alexa Fluor 647-conjugated anti-rabbit IgG. Conjugate formation was imaged using an Andor Revolution spinning disk microscope with a Leica DMi8 microscope, Yokogawa CSU-X1 spinning disk, Andor Technology iXon Ultra 888 EMCCD camera and Fusion software. Samples were imaged using an Olympus Plan Apochromat 100 Å~ 1.4NA oil objective and $z$-step distances of 0.5 μm.

### Nucleofection with Lifeact-mApple and mt-Keima

CTLs were nucleofected with 5 μg of plasmid DNA in P3 Primary cell nucleofector solution using the primary mouse T cell program of the 4D-Nucleofector system. Post-nucleofection, cells were incubated for at 37°C, 10% $CO_2$ in mouse T Cell nucleofector medium with 10% FBS, 2 mM L-glutamine, and 0.01% DMSO. Cells were split after 4 hours into mouse T cell media and allowed to recover for at least 18 hours at 37°C, 10% $CO_2$.

### Live-cell microscopy and quantification of cell migration and mitophagy

CTLs (5 days post activation) were plated in dishes with a coverslip insert (pre-coated with 0.5 μg/ml of ICAM-1/CD45 Fc) in $CO_2$-independent media with 10% FBS and 2 mM L-glutamine, and imaged for 5 min at 37°C. All live microscopy samples were imaged on a Zeiss LSM 780 with Zen Black software, using a Zeiss 63X 1.4NA PlanApochromat objective. Cell migration was quantified in CTLs (5 days after activation) expressing Lifeact-mApple, with nuclei stained with Hoechst, excited at 561 nm and 405 nm, respectively. The speed of CTL migration was quantified using the TrackMate plugin of ImageJ v1.52 (NIH), to track the motion of nuclei as an indicator of cell migration. Mitophagy was detected in CTLs expressing mt-Keima, excited at 488 nm (optimal excitation at pH = 8) and 561 nm (optimal excitation at pH = 4.5). The Imaris spot tracking function was used to filter the mt-Keima signal and calculate the mean fluorescence intensity (MFI) for each object. Spots were defined as "red" (mitolysosomes) when red MFI > 2X green MFI. The percentage mitophagy was calculated from the number of red spots divided by the sum of green and red spots.

### TEM

$CD8^+$ cells were purified using flow cytometry-based sorting 4 days post activation, using

Zombie Aqua and APC-conjugated anti-CD8 and allowed to recover for up to 24 hours in CTL media at 37°C, 10% $CO_2$. CTL were washed and resuspended in serum-free RPMI, plated in four-well plastic tissue culture plates (Nunc) at $2 \times 10^6$ per sample and allowed to adhere at 37°C for 20 min. CTLs were fixed in 2% PFA/1.5% glutaraldehyde, washed and post-fixed in 1% osmium followed by 0.5% uranyl acetate, and processed for EPON embedding as described (*39*). Samples were from three independent biological replicates with two or more technical replicates analyzed per preparation. Sections (60 to 80 nm) were post-stained with lead citrate and viewed using a FEI Tecnai G2 Spirit BioTWIN transmission EM. Images were captured using a Gatan 4K US1000 CCD camera and FEI TIA software (v 4.7 SP3). For quantitation of mitochondria phenotype, images were collected of every cell profile within four (WT = 128 cells total) or five (KO = 154 cells total) grid squares of a single thin section (60 to 80 nm) of each of the WT and KO CTL samples. Each cell profile was scored for mitochondrial morphology as follows: none (no mitochondria), normal (normal mitochondria) and disrupted (mitochondria with disrupted cristae). For quantitation of lytic granule number and cross-sectional area, images were collected of every cell profile within four (WT = 139 cells total) or five (KO = 164 cells total) grid squares of a single thin section for WT and KO CTL samples. Cross sectional area ($nm^2$) was calculated using the lasso tool in Adobe Photoshop (CC 2019).

### Seahorse Mito Stress Assay

Analysis of mitochondrial respiration and glycolytic rate of day 5 post-activation CTLs was performed using a Seahorse XFe96 extracellular flux analyzer with a Seahorse XF Cell Mito Stress Test Kit following manufacturer's instructions. The Seahorse XFe96 FluxPak cartridge was hydrated with distilled water overnight at 37°C (no $CO_2$) and equilibrated for 1 hour on the day of the assay in Seahorse XF Calibrant solution. Vessels were pre-coated with 22.4 μg/ml of Cell-Tak for 20 min at 37°C to ensure continued adhesion of CTLs on the plate during the assay. DMEM used in the assay was supplemented with 10 mM glucose, 1 mM sodium pyruvate, and 2 mM L-glutamine, with pH adjusted to 7.4. To test the OCR and ECAR, 1 μM oligomycin, 1 μM FCCP, and 0.5 μM rotenone/antimycin A were loaded at the indicated times. Seahorse data were collected using Wave Controller 2.6 (Agilent).

### HPG protein synthesis assay

Protein synthesis was assayed using the Click-iT HPG-AF488 Protein Synthesis Assay Kit. CTLs (5 days after activation) were washed and incubated for 30 min at 37°C in methionine-free RPMI supplemented as for CTL culture medium. Protein synthesis was measured by addition of 100 μM Click-IT L-homoproparglycine (HPG) to the media. CHX (100 μg/ml) was added to control samples to inhibit cytosolic translation. Cells were plated ±2 μg/ml of anti-CD3ε at 37°C, 10% $CO_2$ for 1 hour to allow for HPG incorporation.

Cells were then labeled with Zombie Violet and Brilliant Violet 711-conjugated anti-CD8 for 20 min at 4°C. Samples were fixed in 3.7% PFA for 15 min at room temperature, washed in PBS, 3% BSA and permeabilized in PBS, 0.5% Triton X-100 and washed again in PBS, 3% BSA before HPG labeling through click-chemistry using azide-conjugated Alexa Fluor 488 and washed in rinse buffer as indicated in manufacturer's protocol. Cells were resuspended in PBS and analyzed on an Attune NxT.

Mitochondrial protein synthesis was assayed by washing and incubating CTLs (5 days after activation) in CTL culture medium using methionine-free RPMI and FBS dialysed against PBS (to remove methionine). After 30 min at 37°C in methionine-free media, 100 μg/ml of CHX was added to inhibit cytosolic translation and 100 μM Click-IT L-homoproparglycine (HPG) was added for 1-1.5 hours at 37°C, 10% $CO_2$ to label protein synthesis in the mitochondria. CTLs were then washed in serum-free RPMI and allowed to adhere to slides at 37°C, 10% $CO_2$ for 10-15 min before fixation with 3.7% PFA for 15 min at 37°C. Cells were permeabilized for 20 min in 0.5% Triton X-100 and HPG was labeled through click-chemistry using azide-conjugated AF488. Samples were then washed in PBS and mitochondria were labeled using anti-PDH (2 hours) and anti-mouse-AF546 for 1 hour, at room temperature. Nuclei were stained using 200 ng/ml Hoechst 33342, mounted in ProLong Diamond and imaged using the Andor Revolution spinning disk microscope as described above.

### Intracellular cytokine staining

WT and *Usp30*$^{-/-}$ CTLs (5 days after initial activation) incubated ±2 μg/ml of plate-bound anti-CD3ε for 4.5 hours at 37°C, 10% $CO_2$ were treated with GolgiStop (monensin), washed in PBS at 4°C, stained with Zombie Violet and Brilliant Violet 711-conjugated anti-CD8a for 20 min at 4°C, washed in PBS, 10% FBS and fixed and permeabilized using the Foxp3 Transcription Factor Staining Buffer following the manufacturer's protocol and labeled with antibodies to IFN-γ and TNF-α for 30 min at room temperature. Cells were rinsed in permeabilization buffer and PBS, resuspended in PBS and analyzed using an Attune NxT.

### ATP assay

ATP measurements were performed using the ATP Determination Kit. WT and *Usp30*$^{-/-}$ CTLs (5 days after activation) were incubated ±2 μg/ml of plate-bound anti-CD3ε for 4.5 hours at 37°C, 10% $CO_2$. Two million cells per sample were washed in PBS and lysed in distilled water by boiling at 90°C for 5 min as described before (*40*) before the supernatant was cleared by centrifugation at 18,800*g* for 10 min at 4°C. Control samples were treated with 25 μM oligomycin and 100 mM 2-DG to distinguish mitochondrial derived ATP from glycolysis-derived ATP, respectively, and lysed as above. ATP standards and ATP reaction buffer were prepared as indicated by manufacturer. The ATP concentration was quantified by measured by luminescence at 560 nm, using a SpectraMax M5 spectrophotometer and SoftMaxPro 5.4.1 (Molecular Devices).

### Cytof sample processing and analysis

CTLs (5-7 days after activation) ±1 μg/ml of plate-bound CD3ε (145-2C11) for 1 hour at 37°C were stained using MaxPar reagent kits (Fluidigm) using metal-conjugated antibodies as described (*41*).

### Mass spectrometry processing and analysis

CTLs (5 days after activation) were incubated ±2 μg/ml of plate-bound CD3ε (500A2) for 4.5 hours at 37°C. Cytosolic translation was inhibited with 100 μg/ml of CHX and mitochondrial translation with 10 μg/ml of doxycycline (DOX) for 4-4.5 hours. Five million CTLs per sample were then washed twice in ice-cold PBS and flash-frozen on dry ice prior to processing for mass spectrometry.

Protein lysates were generated from cell pellets as per (*42*) but using a lysis buffer of 5% SDS, 50 mM TEAB pH 8.5, and 10 mM TCEP. Protein clean up and peptide generation was performed by the S-Trap method (*43*) as per manufacturer protocol (Protifi) and as previously described (*44*).

For WT/KO and CHX-treated samples, 2 μg of peptide was analyzed per sample. Samples were injected onto a nanoscale C18 reverse-phase chromatography system (UltiMate 3000 RSLC nano, ThermoFisher) then electrosprayed into an Q Exactive HF-X Mass Spectrometer (Thermo Scientific) as described in (*44*). Analysis for doxycycline-treated samples had the following differences: 1.5 μg peptide was injected per sample, analysis was performed on Orbitrap Exploris 480 Mass Spectrometer (ThermoFisher), sample were loaded at 10 μl/min onto trap column and the easy spray source operated in positive mode with spray voltage at 2.445 kV, and the ion transfer tube temperature at 250°C.

The data were processed, searched, and quantified with Spectronaut (*45*) 14 using the directDIA option using the same analysis settings as (*44*). The directDIA data were searched against a mouse hybrid database from databases in July 2019 Uniprot release described in (*42*). Protein copy number quantification was performed in the Perseus

software package, version 1.6.6.0. Mean copy number per cell was calculated using the "proteomic ruler" plugin as previously described (*46*). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (*47*) partner repository with the dataset identifiers PXD021508 (*Usp30*$^{-/-}$, CHX) and PXD026948 (DOX).

### *Flow cytometry and statistical analysis*

All flow cytometry data was acquired using the BD FACSDIVA (BD Biosciences) or Attune NxT software (ThermoFisher) and analyzed using FlowJo v10. Microscopy images were rendered and analyzed using Imaris 9 software (Bitplane). Student's *t* test was used to determine statistical significance as indicated in the figure legends. Graphs were generated using Prism 8 (GraphPad Software).

For mass spectrometry data, copy number values were used to calculate fold change (FC). *P* values were calculated using a two-tailed paired Student's *t* test and applying a FDR correction based on the Benjamini–Hochberg method. Analysis was performed in Microsoft Excel and RStudio. Copy number values were displayed as a heat map using Morpheus (Broad), whereas pie charts indicating cellular compartments were visualized using Prism 8 (GraphPad). Volcano plots were generated by plotting $\log_2$(FC) and $\log_{10}(P)$ for all peptides on Tableau Interactive Dashboard 2020 and highlighting FDR based on the Benjamini–Hochberg method.

Protein cellular localization was manually curated based on UniProt annotations (www.uniprot.org). Mass contribution of proteins (grams per cell) was calculated as (protein copy number/cell) × (molecular weight (Daltons))/(Avogadro's constant). The percentage of mitochondrial protein content of mitochondrial ribosomes was calculated by defining total mitochondrial proteins as the proteins from in the Mouse MitoCarta 2.0 database (*48*) and mitochondrial ribosomal proteins on the protein description terms "ribosomal protein" and "mitochondrial". Cytosolic ribosomal proteins were defined by the KEGG term "ribosome". Moonlighting metabolic enzymes were based on their RNA-binding protein (RBP) ability, as previously reported (*32*, *33*, *49*–*51*). For gene ontology (GO) enrichment analysis, values for FC and FDR correction were compared for proteins detected in the mass spectrometry of both KO vs WT and DOX-treated vs untreated CTLs (all unstimulated). Proteins with corrected *P*-values below the 10% FDR statistical threshold were selected for further analysis. GO enrichment (http://geneontology.org) (*52*, *53*) was analyzed by detecting overrepresented "biological process" in the list of commonly down-regulated proteins in KO and DOX-treated CTLs.

## REFERENCES AND NOTES


1. P. Golstein, G. M. Griffiths, An early history of T cell-mediated cytotoxicity. *Nat. Rev. Immunol.* **18**, 527–535 (2018). doi: 10.1038/s41577-018-0009-3; pmid: 29662120
2. C. J. Sanderson, The mechanism of T cell mediated cytotoxicity. II. Morphological studies of cell death by time-lapse microcinematography. *Proc. R. Soc. London* **192**, 241–255 (1976). pmid: 3791
3. T. L. Rothstein, M. Mage, G. Jones, L. L. McHugh, Cytotoxic T lymphocyte sequential killing of immobilized allogeneic tumor target cells measured by time-lapse microcinematography. *J. Immunol.* **121**, 1652–1656 (1978). pmid: 309477
4. A. Wiedemann, D. Depoil, M. Faroudi, S. Valitutti, Cytotoxic T lymphocytes kill multiple targets simultaneously via spatiotemporal uncoupling of lytic and stimulatory synapses. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10985–10990 (2006). doi: 10.1073/pnas.0600651103; pmid: 16832064
5. M. Fischer *et al*., Early effector maturation of naïve human CD8$^+$ T cells requires mitochondrial biogenesis. *Eur. J. Immunol.* **48**, 1632–1643 (2018). doi: 10.1002/eji.201747443; pmid: 30028501
6. M. D. Buck *et al*., Mitochondrial dynamics controls T cell fate through metabolic programming. *Cell* **166**, 63–76 (2016). doi: 10.1016/j.cell.2016.05.035; pmid: 27293185
7. G. J. van der Windt *et al*., CD8 memory T cells have a bioenergetic advantage that underlies their rapid recall ability. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 14336–14341 (2013). doi: 10.1073/pnas.1221740110; pmid: 23940348
8. N. E. Scharping *et al*., The tumor microenvironment represses T cell mitochondrial biogenesis to drive intratumoral T cell metabolic insufficiency and dysfunction. *Immunity* **45**, 701–703 (2016). doi: 10.1016/j.immuni.2016.08.009; pmid: 27653602
9. A. D. D'Souza, N. Parikh, S. M. Kaech, G. S. Shadel, Convergence of multiple signaling pathways is required to coordinately up-regulate mtDNA and mitochondrial biogenesis during T cell activation. *Mitochondrion* **7**, 374–385 (2007). doi: 10.1016/j.mito.2007.08.001; pmid: 17890163
10. A. J. M. Howden *et al*., Quantitative analysis of T cell proteomes and environmental sensors during T cell differentiation. *Nat. Immunol.* **20**, 1542–1554 (2019). doi: 10.1038/s41590-019-0495-x; pmid: 31591570
11. E. L. Pearce, M. C. Poffenberger, C. H. Chang, R. G. Jones, Fueling immunity: Insights into metabolism and lymphocyte function. *Science* **342**, 1242454 (2013). doi: 10.1126/science.1242454; pmid: 24115444
12. A. V. Menk *et al*., Early TCR signaling induces rapid aerobic glycolysis enabling distinct acute T cell effector functions. *Cell Rep.* **22**, 1509–1521 (2018). doi: 10.1016/j.celrep.2018.01.040; pmid: 29425506
13. M. Giacomello, A. Pyakurel, C. Glytsou, L. Scorrano, The cell biology of mitochondrial membrane dynamics. *Nat. Rev. Mol. Cell Biol.* **21**, 204–224 (2020). doi: 10.1038/s41580-020-0210-7; pmid: 32071438
14. L. Montava-Garriga, I. G. Ganley, Outstanding questions in mitophagy: What we do and do not know. *J. Mol. Biol.* **432**, 206–230 (2020). doi: 10.1016/j.jmb.2019.06.032; pmid: 31299243
15. N. Nakamura, S. Hirose, Regulation of mitochondrial morphology by USP30, a deubiquitinating enzyme present in the mitochondrial outer membrane. *Mol. Biol. Cell* **19**, 1903–1911 (2008). doi: 10.1091/mbc.e07-11-1103; pmid: 18287522
16. B. Bingol *et al*., The mitochondrial deubiquitinase USP30 opposes parkin-mediated mitophagy. *Nature* **510**, 370–375 (2014). doi: 10.1038/nature13418; pmid: 24896179
17. E. Marcassa *et al*., Dual role of USP30 in controlling basal pexophagy and mitophagy. *EMBO Rep.* **19**, e45595–e45595 (2018). doi: 10.15252/embr.201745595; pmid: 29895712
18. L. Abeler-Dörner *et al*., High-throughput phenotyping reveals expansive genetic and structural underpinnings of immune variation. *Nat. Immunol.* **21**, 86–100 (2020). doi: 10.1038/s41590-019-0549-0; pmid: 31844327
19. L. Phu *et al*., Dynamic regulation of mitochondrial import by the ubiquitin system. *Mol. Cell* **77**, 1107–1123.e10 (2020). doi: 10.1016/j.molcel.2020.02.012; pmid: 32142684
20. N. Sun *et al*., Measuring In Vivo Mitophagy. *Mol. Cell* **60**, 685–696 (2015). doi: 10.1016/j.molcel.2015.10.009; pmid: 26549682
21. C. A. Ingraham *et al*., NDUFS4: Creation of a mouse model mimicking a Complex I disorder. *Mitochondrion* **9**, 204–210 (2009). doi: 10.1016/j.mito.2009.02.001; pmid: 19460290
22. S. Campello *et al*., Orchestration of lymphocyte chemotaxis by mitochondrial dynamics. *J. Exp. Med.* **203**, 2879–2886 (2006). doi: 10.1084/jem.20061877; pmid: 17145957
23. R. Uellner *et al*., Perforin is activated by a proteolytic cleavage during biosynthesis which reveals a phospholipid-binding C2 domain. *EMBO J.* **16**, 7287–7296 (1997). doi: 10.1093/emboj/16.24.7287; pmid: 9405358
24. K. Araki *et al*., Translation is actively regulated during the differentiation of CD8$^+$ effector T cells. *Nat. Immunol.* **18**, 1046–1057 (2017). doi: 10.1038/ni.3795; pmid: 28714979
25. S. Isaaz, K. Baetz, K. Olsen, E. Podack, G. M. Griffiths, Serial killing by cytotoxic T lymphocytes: T cell receptor triggers degranulation, re-filling of the lytic granules and secretion of lytic proteins via a non-granule pathway. *Eur. J. Immunol.* **25**, 1071–1079 (1995). doi: 10.1002/eji.1830250432; pmid: 7737276
26. E. L. Pearce *et al*., Control of effector CD8$^+$ T cell function by the transcription factor Eomesodermin. *Science* **302**, 1041–1043 (2003). doi: 10.1126/science.1090148; pmid: 14605368
27. U. Topf, B. Uszczynska-Ratajczak, A. Chacinska, Mitochondrial stress-dependent regulation of cellular protein synthesis. *J. Cell Sci.* **132**, jcs226258 (2019). doi: 10.1242/jcs.226258; pmid: 31028152
28. F. Boos, J. Labbadia, J. M. Herrmann, How the mitoprotein-induced stress response safeguards the cytosol: A unified view. *Trends Cell Biol.* **30**, 241–254 (2020). doi: 10.1016/j.tcb.2019.12.003; pmid: 31964548
29. D. O'Sullivan *et al*., Fever supports CD8$^+$ effector T cell responses by promoting mitochondrial translation. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2023752118 (2021). doi: 10.1073/pnas.2023752118; pmid: 34161266
30. L. Almeida *et al*., Ribosome-targeting antibiotics impair T cell effector function and ameliorate autoimmunity by blocking mitochondrial protein synthesis. *Immunity* **54**, 68–83.e6 (2021). doi: 10.1016/j.immuni.2020.11.001; pmid: 33238133
31. M. Zorkau, C. A. Albus, R. Berlinguer-Palmini, Z. M. A. Chrzanowska-Lightowlers, R. N. Lightowlers, High-resolution imaging reveals compartmentalization of mitochondrial protein synthesis in cultured human cells. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2008778118 (2021). doi: 10.1073/pnas.2008778118; pmid: 33526660
32. A. Castello, M. W. Hentze, T. Preiss, Metabolic enzymes enjoying new partnerships as RNA-binding proteins. *Trends Endocrinol. Metab.* **26**, 746–757 (2015). doi: 10.1016/j.tem.2015.09.012; pmid: 26520658
33. M. Turner, M. D. Díaz-Muñoz, RNA-binding proteins control gene expression and cell fate in the immune system. *Nat. Immunol.* **19**, 120–129 (2018). doi: 10.1038/s41590-017-0028-4; pmid: 29348497
34. F. Salerno, M. Turner, M. C. Wolkers, Dynamic post-transcriptional events governing CD8$^+$ T cell homeostasis and effector function. *Trends Immunol.* **41**, 240–254 (2020). doi: 10.1016/j.it.2020.01.001; pmid: 32007423
35. C. H. Chang *et al*., Posttranscriptional control of T cell effector function by aerobic glycolysis. *Cell* **153**, 1239–1251 (2013). doi: 10.1016/j.cell.2013.05.016; pmid: 23746840
36. F. Buttgereit, M. D. Brand, A hierarchy of ATP-consuming processes in mammalian cells. *Biochem. J.* **312**, 163–167 (1995). doi: 10.1042/bj3120163; pmid: 7492307
37. A. T. Ritter *et al*., Actin depletion initiates events leading to granule secretion at the immunological synapse. *Immunity* **42**, 864–876 (2015). doi: 10.1016/j.immuni.2015.04.013; pmid: 25992860
38. L. O. Randzavola *et al*., Loss of ARPC1B impairs cytotoxic T lymphocyte maintenance and cytolytic activity. *J. Clin. Invest.* **129**, 5600–5614 (2019). doi: 10.1172/JCI129388; pmid: 31710310
39. M. R. Jenkins, A. Tsun, J. C. Stinchcombe, G. M. Griffiths, The strength of T cell receptor signal controls the polarization of cytotoxic machinery to the immunological synapse. *Immunity* **31**, 621–631 (2009). doi: 10.1016/j.immuni.2009.08.024; pmid: 19833087
40. H. S. Garewal, F. R. Ahmann, R. B. Schifman, A. Celniker, ATP assay: Ability to distinguish cytostatic from cytocidal anticancer drug effects. *J. Natl. Cancer Inst.* **77**, 1039–1045 (1986). pmid: 3464795
41. C. Y. Ma, J. C. Marioni, G. M. Griffiths, A. C. Richard, Stimulation strength controls the rate of initiation but not the molecular organisation of TCR-induced signalling. *eLife* **9**, e53948 (2020). doi: 10.7554/eLife.53948; pmid: 32412411
42. J. M. Marchingo, L. V. Sinclair, A. J. Howden, D. A. Cantrell, Quantitative analysis of how Myc controls T cell proteomes and metabolic pathways during T cell activation. *eLife* **9**, e53725 (2020). doi: 10.7554/eLife.53725; pmid: 32022686
43. A. Zougman, P. J. Selby, R. E. Banks, Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. *Proteomics* **14**, 1006–1000 (2014). doi: 10.1002/pmic.201300553; pmid: 24678027

44. L. Reyes *et al*., A type I IFN, prothrombotic hyperinflammatory neutrophil signature is distinct for COVID-19 ARDS. *Wellcome Open Res.* **6**, 38 (2021). doi: 10.12688/wellcomeopenres.16584.2; pmid: 33997298
45. R. Bruderer *et al*., Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Mol. Cell. Proteomics* **14**, 1400–1410 (2015). doi: 10.1074/mcp.M114.044305; pmid: 25724911
46. J. R. Wiśniewski, M. Y. Hein, J. Cox, M. Mann, A "proteomic ruler" for protein copy number and concentration estimation without spike-in standards. *Mol. Cell. Proteomics* **13**, 3497–3506 (2014). doi: 10.1074/mcp.M113.037309; pmid: 25225357
47. Y. Perez-Riverol *et al*., The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **47** (D1), D442–D450 (2019). doi: 10.1093/nar/gky1106; pmid: 30395289
48. S. E. Calvo, K. R. Clauser, V. K. Mootha, MitoCarta2.0: An updated inventory of mammalian mitochondrial proteins. *Nucleic Acids Res.* **44** (D1), D1251–D1257 (2016). doi: 10.1093/nar/gkv1003; pmid: 26450961
49. B. M. Beckmann *et al*., The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nat. Commun.* **6**, 10127 (2015). doi: 10.1038/ncomms10127; pmid: 26632259
50. E. D. Garcin, GAPDH as a model non-canonical AU-rich RNA binding protein. *Semin. Cell Dev. Biol.* **86**, 162–173 (2019). doi: 10.1016/j.semcdb.2018.03.013; pmid: 29574117
51. J. Cieśla, Metabolic enzymes that bind RNA: Yet another level of cellular regulatory network? *Acta Biochim. Pol.* **53**, 11–32 (2006). doi: 10.18388/abp.2006_3360; pmid: 16410835
52. M. Ashburner *et al*., Gene ontology: Tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000). doi: 10.1038/75556; pmid: 10802651
53. Gene Ontology Consortium, The Gene Ontology resource: Enriching a GOld mine. *Nucleic Acids Res.* **49**, D325–D334 (2021). doi: 10.1093/nar/gkaa1113; pmid: 33290552


## ACKNOWLEDGMENTS

We thank R. Hegde, D. Ron, A. Brenes, M. Clague, M. Minczuk, A. Richard, and S. Urbé for their advice and helpful discussions and C. Viscomi and M. Zeviani for spleens from *Ndufs4*$^{-/-}$ mice. We thank the Microscopy and Flow Cytometry facilities at the Cambridge Institute for Medical Research and the Fingerprints Proteomics Facility at the University of Dundee for technical assistance. **Funding:** This research was funded, in whole or in part, by the Wellcome Trust [grants 217100, 103930, 108415, 220543, 100156, and 205023]. For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. M.L. is supported by a PhD studentship from the Cambridge Institute for Medical Research, J.P. is supported by the Medical Research Council (Core grant MC_UU_0015/7), and J.M.M. is supported by an Australian National Health and Medical Research Council CJ Martin early Career Fellowship. **Author contributions:** M.L. carried out all experimental work with the following exceptions: P.R.B. identified USP30 as a regulator of cytotoxicity, and L.O.R. initiated characterization; C.Y.M. provided CyTOF analysis; and J.M.M. and D.A.C. support with analysis of mass spectrometry data. V.P. provided training for Seahorse assays, and J.P. helpful discussions throughout. J.C.S. carried out electron microscopy and contributed throughout. M.L. and G.M.G. conceived the study, designed the experiments, and wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Mass spectrometry data are available from ProteomeXchange with identifier PXD021508 (*Usp30*$^{-/-}$) and PXD026948 (DOX). All other data are available in the main text or supplementary materials.

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abe9977
Figs. S1 to S8
Tables S1 to S5
Movies S1 to S4
MDAR Reproducibility Checklist

27 September 2020; resubmitted 30 June 2021
Accepted 3 September 2021
10.1126/science.abe9977

RESEARCH ARTICLE SUMMARY

NEUROSCIENCE

# Astrocyte $Ca^{2+}$-evoked ATP release regulates myelinated axon excitability and conduction speed

**Jonathan Lezmy*, I. Lorena Arancibia-Cárcamo, Tania Quintela-López, Diane L. Sherman, Peter J. Brophy, David Attwell***

**INTRODUCTION:** Astrocytes support neuronal function throughout the central nervous system. In the gray matter, they regulate synapse number during development, remove synaptically released neurotransmitters to terminate their action and prevent excitotoxicity, control the extracellular potassium concentration to prevent hyperexcitability, regulate blood flow to ensure an adequate energy supply, provide lactate to neurons for energy, and respond to rises of intracellular calcium concentration ($[Ca^{2+}]_i$) by releasing adenosine triphosphate (ATP) and other gliotransmitters that act on neuronal receptors to modulate information processing. However, their role is unclear in the white matter, which transmits information rapidly between gray matter areas using axons wrapped with capacitance-reducing myelin (although they have been suggested to regulate myelination during development and during normal function).

**RATIONALE:** Recently, it has been suggested that learning and memory may reflect not only changes in synaptic function in the gray matter, but also changes in white matter function. In particular, neural circuit function might be regulated by changes in the conduction speed of myelinated axons that result in an altered arrival time of action potentials at a distant neuron. These speed changes might be brought about by alterations of the properties of the passively conducting myelinated internodes or of the intervening excitable nodes of Ranvier, where the action potential is generated. We applied immunohistochemistry to assess how astrocytes interact with myelinated axons, neuronal stimulation and light-evoked calcium uncaging in astrocytes to evoke $Ca^{2+}$-dependent release of gliotransmitters, and electrophysiology and pharmacology to characterize how astrocyte-released substances might affect the axon initial segment (AIS) and nodes of Ranvier of myelinated neurons. Measurements of conduction velocity and computer modeling allowed us to interpret the results.

**RESULTS:** Astrocytes closely approach the axons of myelinated neurons in layer V of the cerebral cortex that enter the corpus callosum. Uncaging $Ca^{2+}$ within astrocytes or stimulating spike trains in neurons evoked a rise of astrocyte $[Ca^{2+}]_i$ that triggered the release of ATP-containing vesicles from these cells. This evoked an inward current in the AIS and nodes of Ranvier of the pyramidal neurons. Pharmacology showed that this was mediated by the activation of $G_s$-linked adenosine $A_{2a}$ receptors ($A_{2a}$Rs), implying that the released ATP was converted to adenosine by extracellular enzymes. The $A_{2a}$Rs raise the intracellular concentration of cyclic AMP, which activates hyperpolarization-activated cyclic nucleotide–gated (HCN) channels mediating the inward hyperpolarization-activated current ($I_h$) and thus depolarizes the cell. In the AIS, the activation of $A_{2a}$Rs alters excitability and hence action potential generation, whereas in the nodes of Ranvier, it decreases the conduction speed of the action potential along the axon.

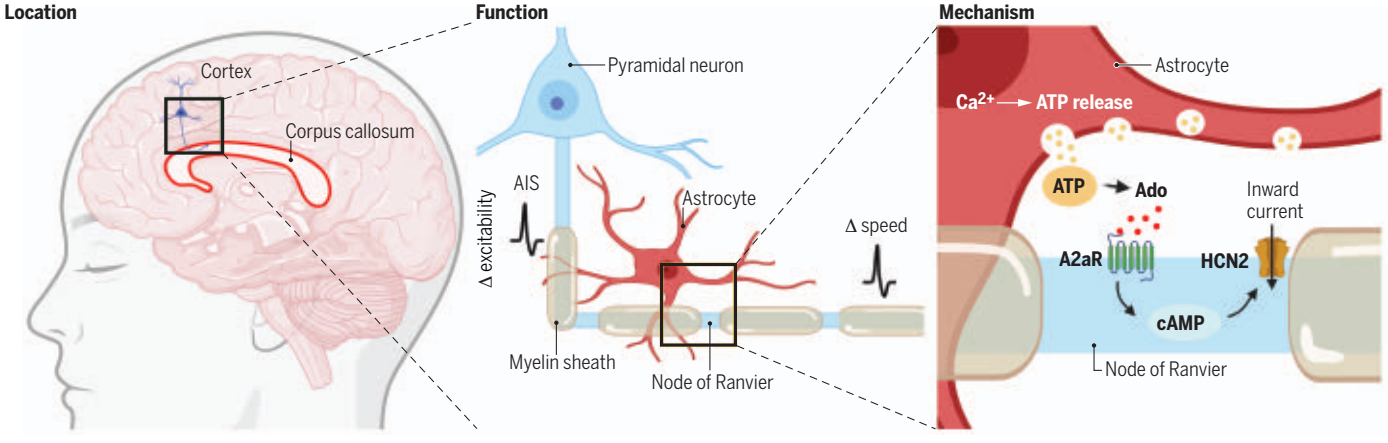
**CONCLUSION:** As in the gray matter, astrocyte $[Ca^{2+}]_i$ regulates the release of ATP into the extracellular space in the white matter. After conversion to adenosine, this regulates the excitability and conduction speed of myelinated axons. The changes in excitability at the AIS will lead to changes in the relationship between the synaptic input and action potential output of the cell. The altered conduction speed of the myelinated axon may change neural circuit function by changing the action potential arrival time at the cell's output synapses, thus altering the integration of signals in postsynaptic neurons. Variations in astrocyte-derived adenosine level can occur between wake and sleep states, and the extracellular adenosine concentration rises during energy deprivation conditions. These changes in adenosine level could thus control white matter information flow and neural circuit function. ■

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

*Corresponding author. Email: j.lezmy@ucl.ac.uk (J.L.); d.attwell@ucl.ac.uk (D.A.)

Cite this article as J. Lezmy *et al.*, *Science* **374**, eabh2858 (2021). DOI: 10.1126/science.abh2858

READ THE FULL ARTICLE AT https://doi.org/10.1126/science.abh2858




**Astrocytes regulate myelinated axon excitability and conduction speed.** For cortical neurons with myelinated axons crossing the corpus callosum (left), astrocytes regulate AIS excitability and axonal conduction speed (middle). Increases of astrocyte $[Ca^{2+}]_i$ release ATP, which, after conversion to adenosine (Ado) extracellularly, activates $A_{2a}$Rs that raise the intracellular cyclic AMP concentration and thus generate an inward current through HCN2 channels in the AIS and nodes of Ranvier (right). Image was created with BioRender.com.

RESEARCH ARTICLE

NEUROSCIENCE

# Astrocyte $Ca^{2+}$-evoked ATP release regulates myelinated axon excitability and conduction speed

Jonathan Lezmy[1]*, I. Lorena Arancibia-Cárcamo[1,2], Tania Quintela-López[1], Diane L. Sherman[3], Peter J. Brophy[3], David Attwell[1]*

In the brain's gray matter, astrocytes regulate synapse properties, but their role is unclear for the white matter, where myelinated axons rapidly transmit information between gray matter areas. We found that in rodents, neuronal activity raised the intracellular calcium concentration ($[Ca^{2+}]_i$) in astrocyte processes located near action potential–generating sites in the axon initial segment (AIS) and nodes of Ranvier of myelinated axons. This released adenosine triphosphate, which was converted extracellularly to adenosine and thus, through $A_{2a}$ receptors, activated HCN2-containing cation channels that regulate two aspects of myelinated axon function: excitability of the AIS and speed of action potential propagation. Variations in astrocyte-derived adenosine level between wake and sleep states or during energy deprivation could thus control white matter information flow and neural circuit function.

Astrocytes support neuronal function throughout the central nervous system. In the gray matter, they regulate synapse number during development, remove synaptically released neurotransmitters, control the extracellular $[K^+]$ ($[K^+]_o$), regulate blood flow, and provide lactate to neurons for energy (*1*). However, their role is unclear in the white matter, which transmits information rapidly between gray matter areas using myelinated axons. Wrapping of these axons with myelin by oligodendrocytes reduces the axon capacitance and thus confers a high conduction speed for the action potential, which, once generated at the axon initial segment (AIS) (*2*), is maintained as it propagates by sodium influx at the nodes of Ranvier occurring between the myelinated internodes. Because oligodendrocytes isolate almost all of the axon from the extracellular space, they are thought to be the main mediators of $[K^+]_o$ control and energy provision to axons (*3*, *4*), making the role of astrocytes uncertain despite astrocyte processes occurring close to nodes of Ranvier (*5*).

Labeling for glial fibrillary acidic protein (GFAP) revealed astrocytes throughout the gray matter and white matter (Fig. 1A). On patch-clamping layer V cortical pyramidal cells or oligodendrocytes, we observed that astrocyte processes were aligned and intimately associated with the myelinated axon and its internodal sheaths (Fig. 1, B and C, and fig. S1, A to F). We used double whole-cell patch-clamping to load an astrocyte with a $Ca^{2+}$-sensing dye (Fluo-4) and to depolarize a pyramidal cell to stimulate action potentials. When the pyramidal cell was driven to fire briefly (30 Hz for 1 s), the astrocyte intracellular calcium concentration ($[Ca^{2+}]_i$) rose more in processes near the neuronal dendrites [within 12.7 ± 2.5 s from start of stimulus to peak fractional change in the fluorescence signal ($\Delta F/F$); mean $\Delta F/F$ = 0.20 ± 0.03, $n$ = 5 neuron-astrocyte pairs] than in processes near the axon ($\Delta F/F$ = 0.03 ± 0.02, $P$ = 0.023; Fig. 1E). However, if prolonged spiking was evoked (10 s), $[Ca^{2+}]_i$ rose in both locations (Fig. 1, D and E, and movie S1; $\Delta F/F$ = 0.19 ± 0.08 and 0.17 ± 0.05, respectively, $P$ = 0.63, $n$ = 3). Uncaging $Ca^{2+}$ in the astrocyte soma also raised astrocyte $[Ca^{2+}]_i$ (Fig. 1F, $\Delta F/F$ = 1.06 ± 0.51 near dendrite, $\Delta F/F$ = 1.05 ± 0.53 near axon, $P$ = 0.86, $n$ = 6), and neuronal activity evoked a further $[Ca^{2+}]_i$ rise superimposed on this ($\Delta F/F$ = 0.39 ± 0.09 near dendrite, $\Delta F/F$ = 0.01 ± 0.06 near axon, $P$ = 0.018).

In the gray matter, astrocytes can modulate neuronal function by releasing adenosine triphosphate (ATP) (*6*). We imaged putative vesicles containing ATP in astrocytes around myelinated axons of layer V pyramidal neurons using quinacrine (*7*). Uncaging $Ca^{2+}$ in the astrocyte soma evoked a $Ca^{2+}$ wave that propagated into processes along the axon (movie S2) and triggered the loss of 43% of quinacrine-labeled vesicles (Fig. 1, G to I; $P$ = 0.0053), a loss that was not seen when the two-photon uncaging excitation was insufficient (see the materials and methods, fig. S2) to evoke a detectable rise of $[Ca^{2+}]_i$ nor when assessing ATP vesicle number in areas 5 μm away (Fig. 1, J and K). The resulting ATP release into the extracellular solution was detected using luciferin-luciferase (see the materials and methods). Puffing 1 mM ATP, but not artificial cerebrospinal fluid (aCSF) from a pipette into the sensing solution evoked a large increase of luciferin luminescence (Fig. 2A), and uncaging $Ca^{2+}$ in an astrocyte soma evoked a similar response that was not seen when the uncaging illumination was too weak to evoke a detectable $[Ca^{2+}]_i$ rise (Fig. 2, B and C).

On release, ATP is rapidly hydrolyzed to adenosine by ecto–adenosine triphosphatases (ecto-ATPases) located on microglia and astrocytes (*8*). Microglia and astrocytes associate with both the AIS and nodes of Ranvier (*9*, *10*). Adenosine receptors have previously been reported at synapses. We used immunohistochemistry to identify adenosine receptors on layer V neuron myelinated axons. Neither $A_1$ nor $A_{2b}$ receptors were detected (fig. S1, G and H), but $A_{2a}$ receptors ($A_{2a}$Rs) were found at 92% of AISs (Fig. 2D, E and H) and at 85% of nodes of Ranvier (Fig. 2, F to H). Overview images of the cortex (fig. S3D) showed $A_{2a}$R–expressing AISs leading toward the corpus callosum, implying that these were excitatory neurons. $A_{2a}$Rs were detected in fewer nodes of the cerebellar white matter, which contains excitatory mossy and climbing fiber axons as well as inhibitory Purkinje cell axons (70%; fig. S3F), and were absent from AISs of cerebellar Purkinje cells (fig. S3E), implying a neuron type–specific expression of $A_{2a}$Rs in myelinated axons. $A_{2a}$Rs raise the cyclic AMP level, which can affect cell excitability by promoting the opening of the hyperpolarization-activated cyclic nucleotide–gated (HCN) type channels (*11*, *12*) present in axons (*13*–*15*). We detected HCN2 channel subunits at 51% of AISs and at 64% of nodes of Ranvier (Fig. 2, I to M). HCN1 subunits were observed in pyramidal cell somata and in nodes of Ranvier (fig. S1, G to I). Adenosine receptors and HCN channels have not previously been reported at the node of Ranvier.

To examine the effect of activation of these $A_{2a}$Rs, we patch-clamped layer V pyramidal neurons and applied adenosine (100 μM from a puffer pipette) or another agonist for $A_{2a}$Rs (CGS 21680, 0.5 μM) to the AIS (Fig. 3A). Activating $A_{2a}$Rs in this way depolarized the cell by 6 to 7 mV (Fig. 3, A, B, and E; aCSF had no effect). The action potential response to small injected currents was increased in frequency (Fig. 3, C, F, and G), but at high injected currents was reduced, presumably as a result of increased $Na^+$ channel inactivation (Fig. 3, D, F, and H). Hyperpolarizing the cell evoked a time-dependent inward current increase (Fig. 3I), which was mediated by $I_h$ (HCN) channels, because it was blocked by the $I_h$ blocker ZD7288 (fig. S4). Applying CGS 21680 increased the amplitude of $I_h$ tail currents, reflecting an increase in magnitude of the fully activated conductance of 51% in 6 cells ($P$ = 0.029), and shifted the 50% point of the activation curve derived from these

[1]Department of Neuroscience, Physiology and Pharmacology, University College London, London WC1E 6BT, UK. [2]UK Dementia Research Institute, Francis Crick Institute, London NW1 1AT, UK. [3]Centre for Discovery Brain Sciences, University of Edinburgh, Edinburgh EH16 4SB, UK.
*Corresponding author. Email: j.lezmy@ucl.ac.uk (J.L.); d.attwell@ucl.ac.uk (D.A.)

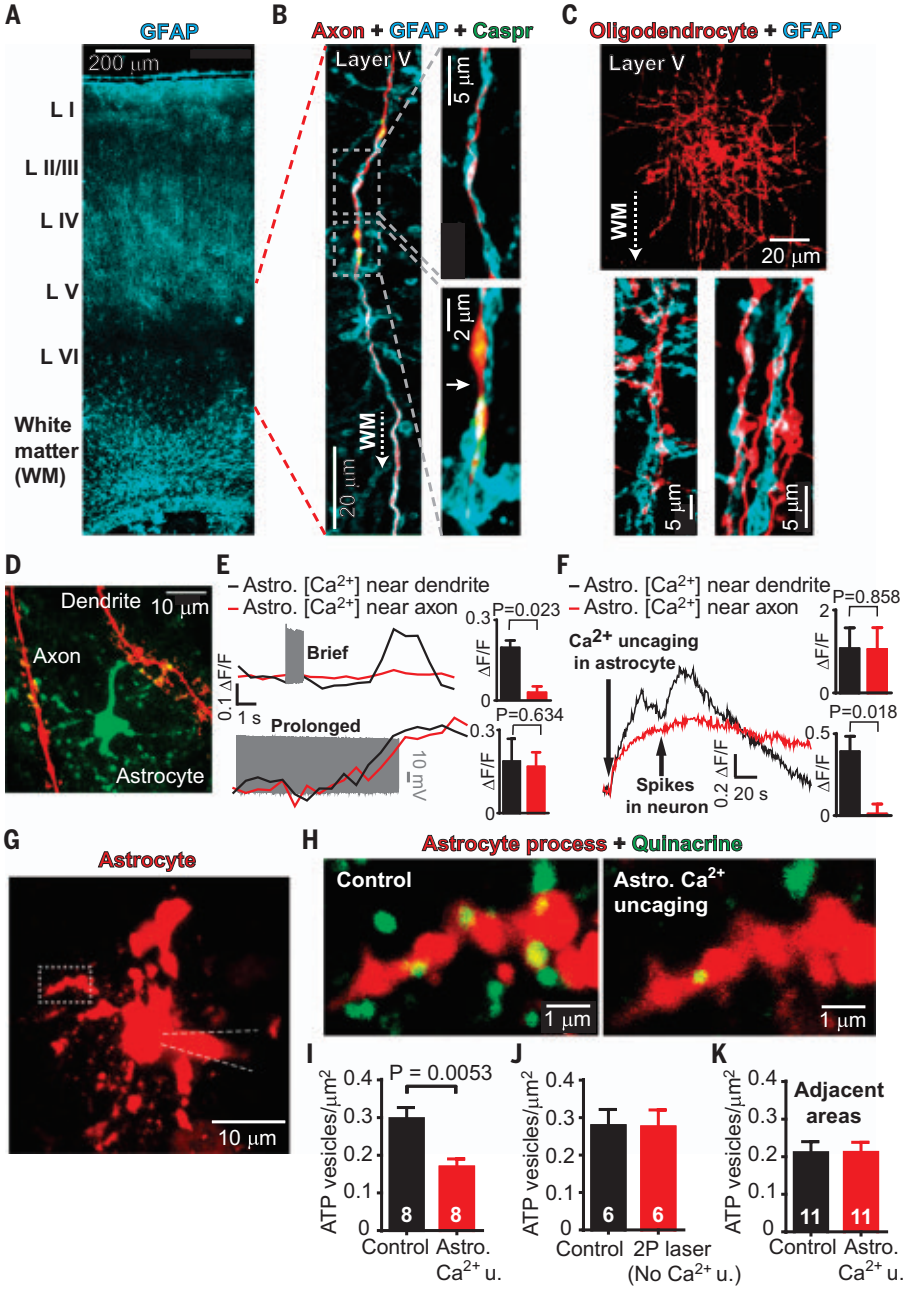



**Fig. 1. Astrocyte processes near myelinated axons release ATP in response to $[Ca^{2+}]_i$ rises.** (**A**) GFAP labeling of mouse coronal slice showing astrocytes in gray matter and white matter. (**B**) Patch-clamp–loaded Alexa Fluor 594 labels axon of layer V pyramidal cell. Expanded internodal (top) and nodal (bottom, identified by Caspr labeling, node is at arrow) regions reveal astrocytes at both locations. (**C**) Patch-clamped oligodendrocyte in layer V with insets below showing GFAP labeling around myelinated internodes. (**D**) Dendrite and axon of patch-clamped rat layer V pyramidal cell near an astrocyte loaded with the $Ca^{2+}$ sensor Fluo-4. (**E**) $[Ca^{2}]_i$ response in astrocyte processes near dendrite (black) and axon (red) to neuron depolarization with 500 pA for 1 s (top, $n$ = 5) or 10 s (bottom, $n$ = 3) to evoke spiking (gray). (**F**) $[Ca^{2+}]_i$ response in astrocyte processes to $Ca^{2+}$ uncaging and to brief neuronal spike trains ($n$ = 6). Top bar chart shows the $Ca^{2+}$ response just before spiking; bottom bar chart shows $Ca^{2+}$ response after neuron spiking evoked by 500 pA for 1 s. (**G**) Patch-clamped rat astrocyte loaded with $Ca^{2+}$ cage NP-EGTA and Rhod-2 to measure $[Ca^{2}]_i$; the region imaged for (H) is shown. (**H**) Quinacrine-labeled puncta in astrocyte process are depleted on uncaging $Ca^{2+}$ (Rhod 2 may not enter the smallest astrocyte processes, explaining why some puncta appear outside the astrocyte). (**I** to **K**) Quantification of ATP vesicles present per square micrometer of astrocyte process before and after uncaging [Astro. $Ca^{2+}$ u.] (I); before and after excitation that did not evoke a $[Ca^{2}]_i$ rise in the astrocyte [No $Ca^{2+}$ u.; see fig. S2] (J); and in regions outside the astrocyte (5 μm away) before and after uncaging (K). Numbers of processes shown on bars. Processes came from 12 cells. Data shown in (A) to (C) are from mice; those in (D) to (K) are from rats.

currents in the depolarizing direction by +8.2 mV ($P$ = 0.045) so that more $I_h$ is activated in the physiological range (Fig. 3, I to K). This shift was mimicked by including 50 μM cAMP in the patch pipette (shifted by +8.6 mV in 6 cells, $P$ = 0.026), and in these cells, CGS 21680 evoked no further depolarizing shift (–1.77 mV in 6 cells, $P$ = 0.087; Fig. 3, J to L).

To assess the function of $A_{2a}$Rs at the nodes of Ranvier, we patch-clamped layer V pyramidal cells. Dye filling revealed an enlarged bleb at the end of the axon where it was cut during the brain-slicing process. This allowed patch-clamping of the bleb (see the materials and methods), which was ~3 (mean 3.2 ± 0.4 in 5 cells) nodes away from the soma (Fig. 4A). In Thy1-Caspr-GFP mice, the nodes of layer V pyramidal neurons could be seen during the experiment, from paranodal GFP fluorescence and because axon branches often occur at nodes, and postrecording immunohistochemistry revealed $A_{2a}$Rs at the AIS and nodes (Fig. 4A). Puffing CGS 21680 onto a node (Fig. 4A) did not significantly depolarize the pyramidal cell soma (0.99 ± 0.46 mV in 5 cells, $P$ = 0.1), implying that the puffed drug did not reach the soma or AIS. Injecting current (500 pA for 5 msec, repeated at 2 Hz) into the soma evoked an action potential. This evoked a delayed action potential in the axonal bleb, which became more delayed when CGS 21680 was puffed at an intervening node (Fig. 4B; >100 responses were averaged). Plots of rate of change of signal (d(signal)/d$t$) versus signal (Fig. 4C) were used to estimate when the action potentials in each location showed an accelerating onset phase (soma depolarization >1 mV/msec and a bleb inward current increase more negative than –0.5 pA/msec, shown as dashed lines in Fig. 4B). The latency of the response at the bleb increased from 108 ± 31 μsec in control conditions to 320 ± 89 μsec when CGS 21680 was puffed ($P$ = 0.03, paired $t$ test, $n$ = 5 cells; Fig. 4D) and the spike width recorded at the bleb increased by 0.28 msec from 0.30 ± 0.10 msec to 0.58 ± 0.14 msec ($P$ = 0.015, paired $t$ test, $n$ = 5).

To convert this latency increase to a change of conduction speed, we need to assume where in the AIS the action potential is initiated and how fast it propagates backward to the soma compared with forward to the bleb (see the materials and methods). If the spike starts at the middle of the AIS and the forward speed is twice the backward speed (*16*), then CGS 21680 reduced the forward speed from 1.21 ± 0.23 to 0.43 ± 0.08 m/sec (64% reduction; filled circles in Fig. 4E). Alternatively, if the spike starts at the end of the AIS and the forward speed is three times faster than the backward speed (*16*), then the speed is reduced from 0.68 ± 0.16 to 0.31 ± 0.13 m/sec (54% reduction; open circles in Fig. 4E). Thus, for a range of assumptions, activation of nodal $A_{2a}$Rs produces

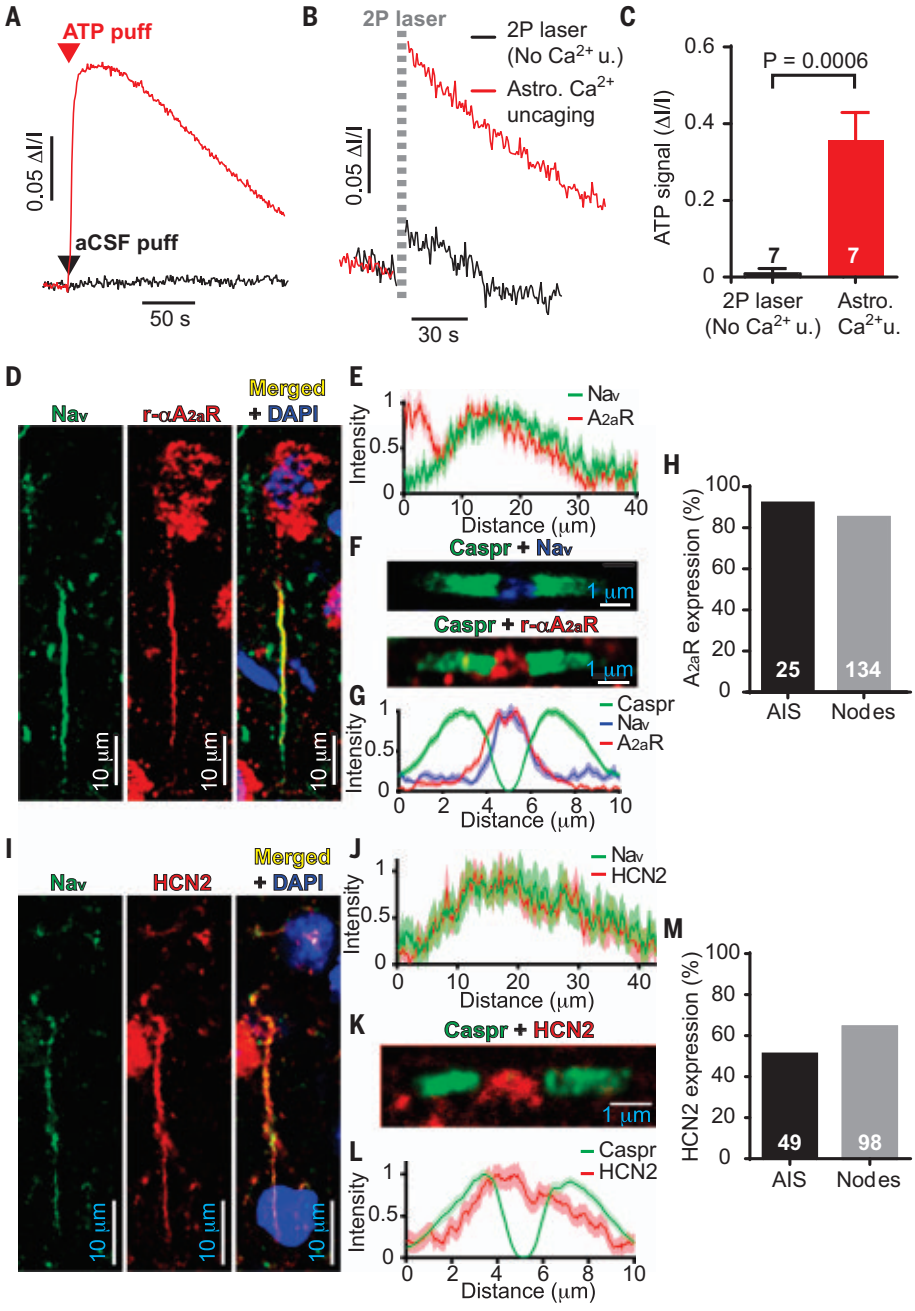



**Fig. 2. ATP release from astrocytes may target adenosine receptors on myelinated axons of layer V pyramidal neurons.** (**A**) Detection of ATP puffed into extracellular solution using luciferin-luciferase. (**B**) Two-photon excitation uncaging $Ca^{2+}$ in astrocytes evoked a luciferin-luciferase signal, unless the excitation failed to raise astrocyte $[Ca^{2+}]_i$ (No $Ca^{2+}$ u.; see fig. S2). (**C**) Quantification of experiments in (B) in seven cells. (**D** to **H**) $A_{2a}$Rs are present in the AIS (D), where they overlap with $Na_v$ expression [mean of 14 $Na_v$ and 25 $A_{2a}$R profiles (E)] and at the node of Ranvier (F), where they overlap with $Na_v$ and are flanked by Caspr labeling [mean of 48 Caspr, 48 $A_{2a}$R and 24 $Na_v$ profiles (G)]. (H) Percentage of 25 AISs and 134 nodes that express $A_{2a}$Rs. (**I** to **M**) HCN2 channel subunits are present in the AIS (I), where they overlap with $Na_v$ expression [mean of 19 $Na_V$ and 16 HCN2 profiles (J)], and at the node of Ranvier (K), where they overlap with $Na_v$ and are flanked by Caspr labeling [mean of 53 Caspr and HCN2 profiles (L)]. (M) Percentage of 49 AISs and 98 nodes that express HCN2. Data shown in (A) to (E), (I), and (J) are from rats; those in (F), (G), (K), and (L) are from mice; and those in (H) and (M) combine rat AIS and mouse node data.

a very significant reduction of conduction velocity ($P = 0.0017$ comparing the ratio of the conduction velocities).

To explore how $A_{2a}$Rs and $I_h$ affect excitability and conduction velocity, we used a MATLAB model of myelinated axons in the corpus callosum (*17*), which was adapted to mimic either (i) a soma attached to a truncated axon with three internodes and a terminal bleb (Fig. 5A) or (ii) an infinitely long axon (see the materials and methods). The AIS and nodes of Ranvier contained voltage-gated $Na^+$ and $K^+$ channels to generate action potentials, whereas the soma and bleb lacked these channels. We modeled the $I_h$ current in the absence of $A_{2a}$R activation as being present in the soma and proximal AIS only and unaffected by the puff application of $A_{2a}$R-activating drugs at the distal AIS (and the [cAMP] rise that they produce), consistent with the presence in the somatodendritic compartment (fig. S1G) of $I_h$ channels composed of HCN1 subunits (*18*), which are relatively insensitive to cAMP (*19*). The increase in $I_h$ evoked by $A_{2a}$R activation in the AIS was modeled as the addition of an extra current mediated by cAMP-activated HCN2 channels located in the distal AIS. The maximum conductance of this extra current was set to increase the total maximum $I_h$ conductance (measured at the soma) by 51% as observed (Fig. 3I), and the midpoint of the activation curve was set at −80 mV to reproduce the positive shift for the total $I_h$ apparent activation curve in Fig. 3J. Adenosine-activated $I_h$ was also added to the nodes of Ranvier at a density 1.41-fold higher than in the distal AIS, as measured immunohistochemically (see the materials and methods).

Adding the adenosine-activated conductance to the distal AIS in the model depolarized the soma by 5.9 mV, as seen experimentally (Figs. 3E and 5B). It also increased the firing evoked by a small injected current (Fig. 5, C and D) while decreasing that evoked by a large current (Fig. 5, C and D), also as seen experimentally (Fig. 3, C to H). However, the simulated depression of action potential amplitude at large currents was larger than that observed experimentally. These changes were largely a result of the $I_h$ added to the AIS, because adding it solely to the nodes had only a minor effect on the resting potential at the soma (Fig. 5B). Adding $I_h$ to the nodes of Ranvier decreased the conduction speed (Fig. 5E), as observed when puffing CGS 21680 onto the nodes (Fig. 4E). The predicted reduction was smaller than that observed experimentally, presumably because of the mild depolarization generated at the nodes (+3.8 mV). Using the infinite cable model, adding the adenosine-activated $I_h$ depolarized the nodes by +11.4 mV, reduced the predicted conduction speed by 48% from 2.23 to 1.17 m/s (Fig. 5F), and increased the axonal spike width by 0.22 ms (mainly by increasing

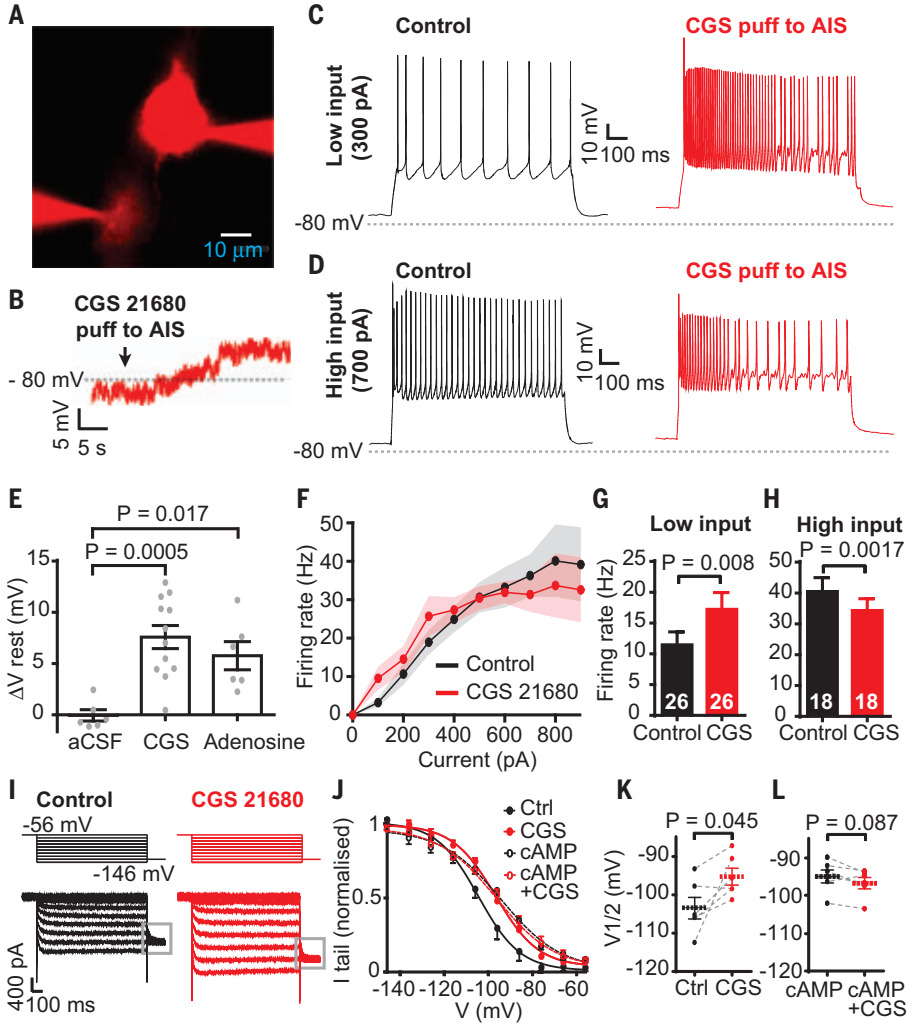



**Fig. 3. $A_{2a}$ receptors, through cAMP and $I_h$, modulate excitability at the AIS.** (**A**) Patch-clamped layer V pyramidal cells loaded with Alexa Fluor 594. Lower pipette puffing adenosine at distal AIS contains Alexa Fluor 594 to delineate the region affected by adenosine. (**B**) Depolarization of soma evoked by puffing the $A_{2a}$R agonist CGS 21680 (0.5 µM). (**C** and **D**) Voltage response to 300 pA (C) or 700 pA (D) injected current with and without CGS 21680 application. (**E**) Mean resting potential change when puffing aCSF ($n$ = 6), 0.5 µM CGS 21680 ($n$ = 12), or 100 µM adenosine ($n$ = 6) onto the AIS. (**F**) Firing rate averaged over 1 s as a function of injected current, in control conditions or with CGS 21680 applied to the AIS ($n$ = 9; SEM shown faint). (**G** and **H**) Firing rate change for "low" [averaged over 100 to 300 pA, 26 current steps from nine cells (G)] or "high" (700 to 900 pA, 18 current steps from eight cells (H)] input currents. (**I**) Specimen currents on stepping from −56 mV to various pulse potentials and then to −136 mV to evoke tail currents, allowing construction of the activation curve in control conditions and with CGS 21680 puffed at the AIS. (**J**) $I_h$ activation curves for normal conditions, puff application of CGS 21680 ($n$ = 6), and both with 50 µM cAMP included in the patch pipette ($n$ = 6). (**K**) $V_{1/2}$ values (50% activation voltage of fitted Boltzmann curve) before (Ctrl) and during CGS 21680 application. (**L**) As in (K) but with cAMP in patch pipette. All data are from rats.

$Na^+$ channel inactivation at the resting potential because this was mimicked by decreasing the peak $Na^+$ conductance ~twofold), similar to the 0.28 msec reported experimentally above.

Because astrocyte $[Ca^{2+}]_i$ rises evoked the release of ATP (Fig. 1, G to K, and Fig. 2, A to C), which is expected to be converted to adenosine extracellularly and to act on $A_{2a}$Rs at the AIS and nodes of Ranvier, we tested whether uncaging $Ca^{2+}$ in astrocytes modulated the excitability and conduction speed of cortical layer V pyramidal cell myelinated axons. $Ca^{2+}$ was uncaged in an astrocyte with processes running near the AIS of a pyramidal cell (Fig. 6A and movie S3). Within 70 s of uncaging the $Ca^{2+}$, the resting potential of the neuron was depolarized (Fig. 6B) and the action potential response of the neuron was changed (Fig. 6C), exhibiting a higher firing rate to low injected currents ($P$ = 0.0175 when comparing firing rates normalized to the maximum rate evoked at high injected current) and a lower firing rate to high injected currents ($P$ = 0.0136). These changes, which are similar to those seen when activating AIS $A_{2a}$Rs (Fig. 3, B to H), were not mediated by glutamate release from astrocytes (fig. S5), were blocked by superfusion of the $A_{2a}$R blocker ZM 241385 (100 nM) or by intracellular dialysis of the neuron with the $I_h$ blocker ZD7288 (20 µM), and were not seen if the uncaging illumination failed to evoke a $[Ca^{2+}]_i$ rise (Fig. 6, B to F).

Uncaging $Ca^{2+}$ in astrocytes, which were shown by post hoc immunostaining to have processes close to nodes of Ranvier (Fig. 6, G and H), evoked a decrease of axonal conduction speed (Fig. 6I) in experiments monitoring the action potential propagation from the soma to a bleb ~4 (mean 3.8 ± 0.5 in 5 cells) nodes along the axon. These data are similar to those obtained by puffing an $A_{2a}$R agonist onto the nodes (Fig. 4) and, as for those experiments, the decrease of conduction speed calculated was dependent on the exact site of action potential initiation within the AIS and its forward and backward propagation speeds (Fig. 6I). Introducing ZD7288 (20 µM) into the targeted axons to block $I_h$ (HCN) channels prevented the speed reduction evoked by astrocyte $Ca^{2+}$ activity (fig. S6A). In unmyelinated axons and some other neuronal types, axonal HCN channels can speed, rather than slow, the action potential (*13*, *14*) (see fig. S6 legend). Inducing robust neuronal activity (30 Hz for 1 min) to raise astrocyte $[Ca^{2+}]_i$ (Fig. 1E) also induced a decrease of axonal conduction speed ($P$ = 0.014), which was abolished by superfusing the $A_{2a}$R blocker ZM 241385 (100 nM) (fig. S6B).

These data reveal a new modulation of neuronal circuit function by glial cells. Astrocyte modulation of neuronal synaptic function in the gray matter has been accepted as a major determinant of neuronal function (*6*, *20*, *21*), but the role of white matter astrocytes is less clear. Our data, summarized in fig. S7, demonstrate that astrocytes modulate the excitability of the AIS of excitatory neurons and regulate the conduction speed of myelinated axons. This regulation is a result of astrocytes raising the concentration of adenosine near the AIS or nodes of Ranvier (of a single axon or possibly several axons at once), which can increase or decrease AIS excitability and decreases axon conduction speed. Adenosine release can occur when axonal action potential firing raises $[Ca^{2+}]_i$ in adjacent astrocytes (*22*); this triggers the release of ATP from vesicles (Figs. 1 and 2), which is subsequently converted to adenosine by ecto-ATPases expressed by microglia, astrocytes, and oligodendrocyte lineage cells (*8*). Neuronal activity in cortical layer V

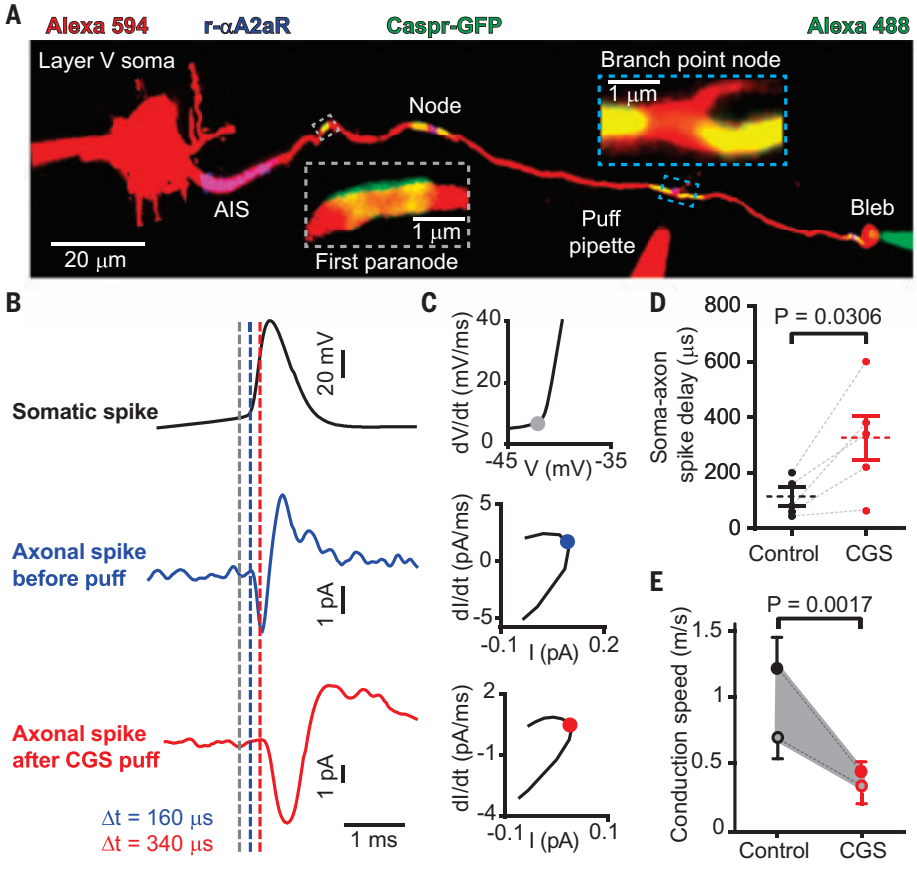



**Fig. 4. $A_{2a}$Rs in the node of Ranvier modulate conduction velocity.** (**A**) Myelinated axon in Thy1-Caspr-GFP mouse filled with Alexa Fluor 594 and patch-clamped at the cell soma and end-of-axon bleb. (**B**) Average of >100 evoked action potentials in the soma and bleb in control conditions and while puffing 0.5 μM CGS 21680 at a node of Ranvier. Dashed lines show times of initiation of action potential derived from threshold values of rate of change of voltage (d*V*/d*t*) or current (d*I*/d*t*) with time (see text). (**C**) Phase plane plots showing times indicated in (B) (dots). (**D**) Response latency in bleb. (**E**) Conduction velocities derived making assumptions discussed in the main text (closed circles assume spike starts at the middle of the AIS and the forward speed is twice the backward speed; open circles assume spike starts at the end of the AIS and the forward speed is three times faster than the backward speed). Data in (A) to (C) are from mice; those in (D) and (E) combine data from rats and mice (neither the initial speed nor the percentage change evoked by CGS 21680 differed significantly between rats and mice, $P$ = 0.43 and 0.15, respectively).

gray matter readily raised $[Ca^{2+}]_i$ in astrocyte processes near neuronal dendrites, presumably as a result of classical neurotransmitters or other signaling molecules acting on astrocytes, but more intense neuronal activity (or conceivably a large number of neurons firing at a low rate) was needed to raise it in the processes near axons (Fig. 1, E and F). Thus, the overall level of neuronal activity may modulate the conduction speed of white matter axons.

Changes in the ATP level in astrocytes or neurons may also alter the extracellular adenosine concentration as a result of intracellular interconversion of ATP and adenosine and export or import of adenosine across the cell membrane by the equilibrative nucleoside transporter. Across a wide area of neocortex, the intracellular ATP level decreases during neuronal activity, on passing from the awake state to non–rapid eye movement (non-REM) sleep, and on passing from non-REM sleep to REM sleep (*23*). In the basal forebrain, the accumulation of extracellular adenosine when awake has been proposed to generate pressure to sleep (*24*), and this adenosine is likely to be generated from ATP released by astrocytes (*25*, *26*). Although the arousal-modulating effects of adenosine have often been attributed to it acting on presynaptic $A_1$ receptors to suppress glutamate release at excitatory synapses, genetic evidence suggests that in fact these effects are generated by $A_{2a}$Rs (*27*). Thus, the effects of astrocyte-derived adenosine on AIS excitability and myelinated axon conduction speed that we have characterized may be crucial contributors to mediating changes between wake and sleep states.

GABA, dopamine, and serotonin receptors regulate neuronal excitability at the AIS (*15*, *28*, *29*). Our data demonstrate that the node of Ranvier can have its electrical function rapidly altered by substances released from surrounding cells. We have shown this for astrocyte-released ATP/adenosine, but oligodendrocyte precursor cells and microglia could theoretically exert similar effects either by releasing ATP or adenosine directly onto spike generation sites or by responding to astrocytic ATP release by releasing other gliotransmitters onto the axons. Neural circuits depend crucially on action potential arrival time for their function (*30*), and cognition depends on oscillations of neural firing probability that in turn depend on propagation time in myelinated neurons (*31*–*33*). Thus, the decrease of the conduction speed of myelinated axons that is evoked by adenosine may change information processing and cognition, when the adenosine level rises either during a prolonged awake period or in response to complex motor behavior and pathological conditions. Indeed, behavioral effects linked to exploratory activity, aggression and anxiety (*34*, *35*), and pathological effects in epilepsy, Parkinson's disease, Alzheimer's disease, depression, and autism have all been linked to impairments of adenosine signaling (*36*, *37*). Our results make the testable prediction that, when adenosine levels rise during prolonged wakefulness or in complex motor behavior and pathological conditions, the conduction speed of myelinated axons should decrease, resulting in delays to the action potential arrival time at downstream synapses and possible impairments of coincidence detection or oscillatory firing generation.

## Materials and Methods

### *Animals*

Sprague-Dawley rats or transgenic mice at postnatal days 28 to 32, housed on a 12-hour light/12-hour dark cycle were used in all experiments. At this age, myelination is advanced, but cells are still suitable for stable patch-clamp recordings. Animals were killed to make brain slices 4 hours after the room light was turned on. Each experiment was performed on brain slices from at least three animals and at least one of each sex. Expression of $A_{2a}$Rs and HCN channels in sites of spike generation appeared the same in mouse and rat and in males and females. Animal procedures were carried out in accordance with the guidelines of the UK Animals (Scientific Procedures) Act 1986 and subsequent amendments (under Project License 70/8976). Thy1-Caspr-GFP mice were generated as previously described (*38*).

### *Acute brain slice preparation*

Coronal cortical slices (300 μm thick) were prepared on a vibratome in ice-cold solution containing (in mM): 93 N-methyl-D-glucamine

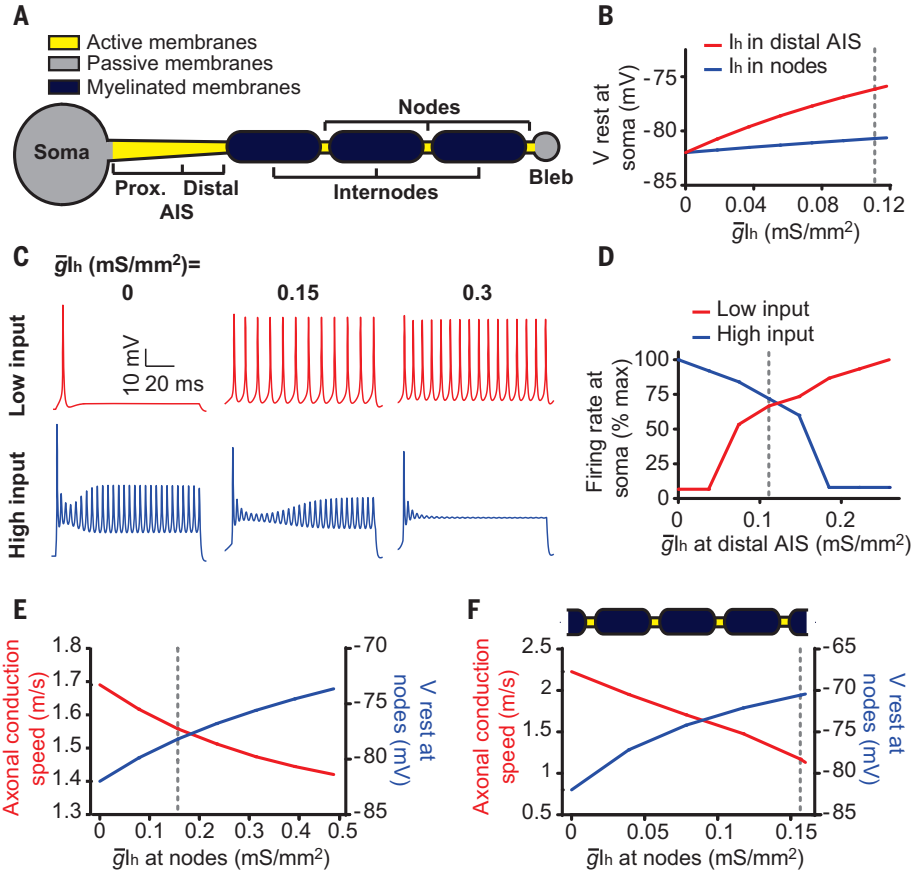



**Fig. 5. Computational modeling predicts the adenosine-evoked decrease of axonal conduction speed.** (**A**) Schematic diagram of the model of the experiment in Fig. 4. (**B**) Adding different densities of adenosine-sensitive maximal $I_h$ conductance ($\bar{g}I_h$) to the distal AIS evokes a larger soma depolarization than adding it to the nodes of Ranvier. Vertical dashed line shows measured maximal conductance (0.11 mS/mm$^2$). (**C**) Voltage response at soma to injecting 20 pA (top row) or 180 pA (bottom row) with three different levels of $\bar{g}I_h$ added (as indicated) to the distal AIS. As observed experimentally (Fig. 3, C to H), at low injected current, the action potential frequency is increased; at high injected current, it is decreased (because of the simulated decrease of action potential amplitude, we defined an action potential as occurring if the voltage crossed −50 mV). (**D**) Firing rate as a function of $\bar{g}I_h$ for simulations as in (C). (**E**) Action potential speed from the first to the last node in (A) and average resting potential of the three nodes as a function of $\bar{g}I_h$ added to each node (vertical dashed line shows the estimated physiological value of 0.1565 mS/mm$^2$; see the materials and methods). (**F**) Predictions of infinite axon model for conduction speed and node resting potential as a function of $\bar{g}I_h$.

(NMDG) chloride, 2.5 KCl, 30 $NaHCO_3$, 10 $MgCl_2$, 1.2 $NaH_2PO_4$, 25 glucose, 0.5 $CaCl_2$, 20 HEPES, 5 sodium ascorbate, 3 sodium pyruvate, and 1 kynurenic acid. The slices were incubated at 37°C in this solution for 20 min, then transferred to a similar solution containing (in mM): 93 NaCl, 1 $MgCl_2$, and 2 $CaCl_2$ instead of the NMDG chloride, $MgCl_2$, and $CaCl_2$, and incubated at room temperature until use. Experiments were performed in aCSF containing (in mM): 125 NaCl, 3 KCl, 26 $NaHCO_3$, 2 $MgCl_2$, 2 $CaCl_2$, 1.25 $NaH_2PO_4$, and 10 glucose heated to 37°C. All the solutions were gassed with 95% $O_2$ and 5% $CO_2$. In some experiments, 50 μM ZD7288, 100 nM ZM 241385, 20 μM D-AP5, 10 μM DNQX, 50 μM MSPG, and 1 μM NPS 2390 (all from Tocris Bioscience) were added to the aCSF.

### *Patch-clamping of neuronal somata*

Experiments were performed with an Olympus BX51WI microscope under an Olympus LUMPlanFI 40× lens. Layer V pyramidal cells were identified by their location and morphology. Microelectrodes with resistances of 5 to 6 MΩ were pulled from borosilicate glass capillaries (Harvard Apparatus) and filled with an intracellular solution containing (in mM): 145 K-gluconate, 2 $MgCl_2$, 0.5 $H_2$-EGTA, 2 MgATP, 0.2 $Na_2GTP$, and 10 HEPES, pH adjusted to 7.2 with KOH. Somata were patch-clamped in whole-cell configuration, and the signals were amplified using a Multiclamp 700B (Molecular Devices), filtered at 4 kHz, and digitized at 50 kHz. The series resistance was <20 MΩ, was compensated by ≥70%, and did not vary by >10% during experiments. The mean input resistance in whole-cell mode, after compensation for series resistance, was 130.6 ± 25.9 MΩ in control conditions and 84.8 ± 23.8 MΩ during CGS 21680 application to the AIS ($n$ = 6, $P$ = 0.049). Recordings were acquired with Clampex and analyzed with Clampfit software (Molecular Devices). In current-clamp mode, bridge balance and pipette capacitance were corrected. To generate graphs of firing rate versus injected current, current steps in increments of 100 pA were injected for 1 s. All membrane potentials were corrected for a pipette liquid junction potential of −16 mV. To analyze currents activated by hyperpolarization, tail currents were analyzed at −136 mV obtained after 1 s pretest pulses in 10-mV increments from −56 to −146 mV. The tail currents analyzed were primarily generated by HCN channels ($I_h$ current) because they were significantly diminished by adding 50 μM ZD7288 (Tocris Bioscience) (fig. S4). Tail currents were fitted with single exponential curves that were extrapolated back to the end of the pretest pulses. The resulting tail amplitude values were fit to a function:

$$I = I_{max} + (I_{min} - I_{max})/\{1 + \exp[k(V - V_{1/2})]\} \quad (1)$$

where $I_{max}$ is the predicted tail amplitude at −136 mV for a large depolarizing voltage step, $I_{min}$ is the tail amplitude for a large hyperpolarizing step, $V_{1/2}$ is the voltage midpoint of the activation curve, and $k$ is a factor defining the slope of the activation curve. The conductance conferred by the maximally activated current is given by

$$G_{max} = (I_{max} - I_{min})/[V_{rev} - (-136\ mV)] \quad (2)$$

where the reversal potential of $I_h$ was taken as −23 mV (*13*). In some experiments, 50 μM cAMP (Sigma-Aldrich) or 20 μM ZD7288 (Tocris Bioscience) was added to the intracellular solution. ZD7288 was previously shown to block $I_h$ when applied intracellularly (*15*).

### *Patch-clamp recording from axonal blebs*

Experiments were performed using a Zeiss LSM780 two-photon confocal microscope with a W Plan-Apochromat 20× objective. Blebs are patchable structures formed at the cut end of axons at the surface of slices (*39*). Alexa Fluor 594 (ThermoFisher Scientific) was allowed time to diffuse from the soma to the bleb. Action potentials were recorded as a current in voltage-clamp mode. To avoid phototoxicity, exposures to excitation lights were briefly used only to trace the axon and monitor the puffing area (see below). High-resolution images of whole axons were acquired after completion of the electrophysiology experiments. Exposure to the excitation lasers did not significantly affect

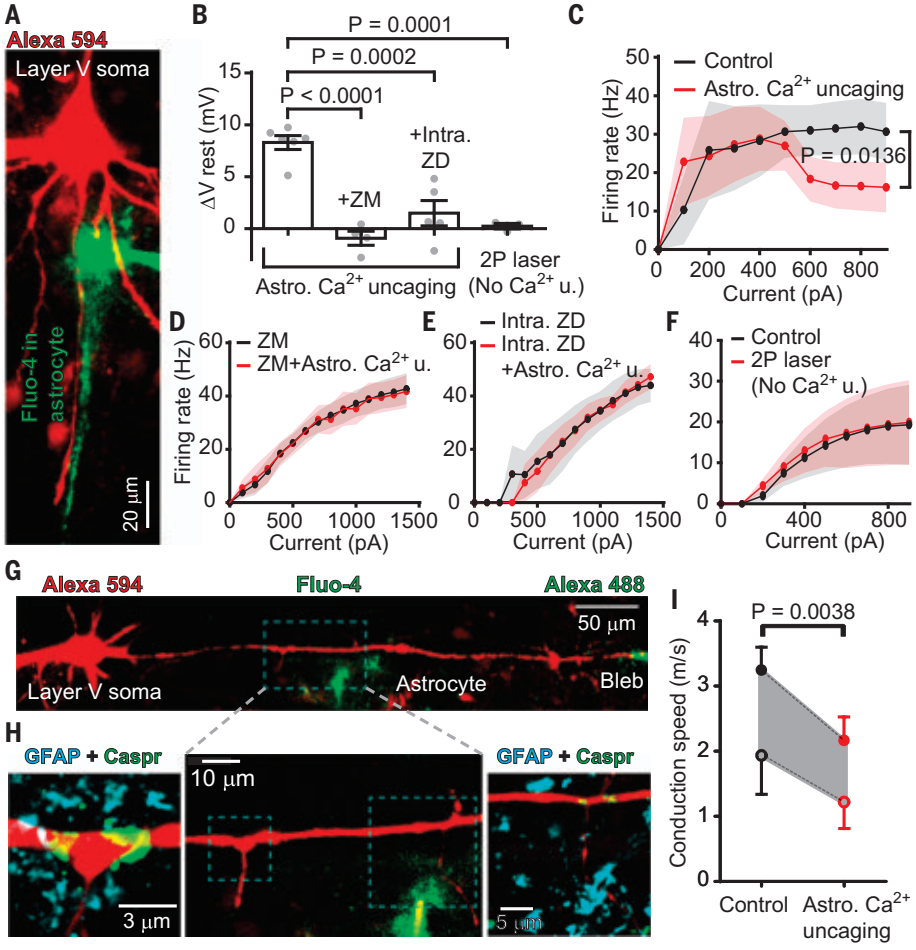



**Fig. 6. $Ca^{2+}$ concentration rises in astrocytes regulate pyramidal cell excitability and axonal conduction speed.** (**A**) Patch-clamped L5 pyramidal neuron (with Alexa Fluor 594 in the left pipette, red) and a periaxonal astrocyte filled with the $Ca^{2+}$ cage NP-EGTA and Fluo-4 (right pipette, green). $[Ca^{2+}]_i$ rises in astrocyte processes near the axon (see movie S3). (**B**) Resting potential ($V_{rest}$) depolarized after astrocyte $Ca^{2+}$ uncaging but not when blocking $A_{2a}$Rs with superfused ZM 241385 (100 nM) or $I_h$ channels with ZD7288 (20 µM in the pipette) nor when two-photon laser excitation failed to raise $[Ca^{2+}]_i$ (No $Ca^{2+}$ u.; see fig. S2) (Astro. $Ca^{2+}$ uncaging: $n$ = 6, +ZM: $n$ = 4, +Intra. ZD: $n$ = 5, 2P laser without uncaging: $n$ = 3; one-way ANOVA, $P < 0.0001$). (**C**) Neuronal firing rate evoked by injecting 1 s current steps in 100 pA increments before (black) and after astrocytic $Ca^{2+}$ uncaging (red). (**D** to **F**) As in (C) but with ZM 241385 [$n$ = 4 (D)], or ZD7288 [$n$ = 4 (E)], or with illumination that failed to uncage $Ca^{2+}$ [$n$ = 3 (F)]. (**G**) Live imaging of an L5 pyramidal neuron patch-clamped at the soma (left pipette, loading red Alexa Fluor 594) and the axon end (right pipette, green). A perinodal astrocyte (near axon branches) was patch-filled with NP-EGTA and Fluo-4. (**H**) Middle: high-resolution image of the dashed box in (G). Left and right images show the areas in the dashed boxes of the middle image after immunostaining for GFAP and Caspr. Nodes flanked by Caspr (green) are close to GFAP-positive astrocyte processes (cyan). (**I**) Estimated axonal conduction speed before and after astrocyte $Ca^{2+}$ uncaging (see text associated with Fig. 4E for assumptions made). Data in (A) to (H) are from rats; (I) combines data from rats and mice (neither the initial speed nor the percentage change evoked by $Ca^{2+}$ uncaging differed significantly between rats and mice, $P$ = 0.08 and $P$ = 0.93, respectively).

the waveform of action potentials. In Fig. 6, D and E, where the effects of adenosine receptors or HCN channels were blocked, there was no change in action potential width measured at the soma. For the first spike evoked by a 500-pA step, the full width at half-maximum with ZM 241385 was 1.15 ± 0.13 msec before and 1.18 ± 0.14 msec after $Ca^{2+}$ uncaging ($n$ = 4; not significantly different, $P$ = 0.12); with ZD7288, it was 1.25 ± 0.17 msec before and 1.21 ± 0.13 msec after $Ca^{2+}$ uncaging ($n$ = 4; not significantly different, $P$ = 0.44). Similarly, there was no difference in the action potential width measured at the bleb with ZD7288 inside the neurons (fig. S6A); the full width at half-maximum (the maximum being the amplitude of the inward current peak) was 0.53 ± 0.17 msec before and 0.53 ± 0.10 msec after $Ca^{2+}$ uncaging ($n$ = 5; not significantly different, $P$ = 0.99). Microelectrodes with resistances of 8 to 9 MΩ were filled with aCSF and 100 µM Alexa Fluor 488 (ThermoFisher Scientific). Blebs were patch-clamped in cell-attached configuration, and the signals were filtered at 2 kHz and digitized at 50 kHz. Action potentials at the soma were evoked and recorded, and action currents at the axon bleb were recorded simultaneously. To accurately determine the delay between the somatic and axonal signals, at least 100 somatic spikes were evoked. First derivatives of the somatic and axonal signals were generated and then time-shifted to align the peaks of the somatic spikes. The signals were then averaged, and the soma-axon delay was measured as the difference between the onset of the axonal spike (time point recorded before a current drop of >0.5 pA/ms) and the somatic spike (time point recorded before an increase of >1 mV/ms).

### *Local application of drugs to axons*

CGS 21680 (500 nM; Cayman Chemical) or 100 µM adenosine (Sigma-Aldrich) were puff-applied onto the AIS or nodes of Ranvier. The drugs were added to aCSF with 100 µM Alexa Fluor 594 to monitor the spread of the drugs. Two or three MΩ pipettes were filled with this solution and placed 10 µm away from the targeted axon. Positive pressure was either applied with a microinjector (PMI-100, Dagan) for 20 ms at 10 psi or controlled manually with a syringe. In both cases, positive pressure was calibrated and live monitored to eject drug over a radius of ≤20 µm (from the tip of the pipette set at saturating intensity to the edge of detectable fluorescence). The flow of the perfusion was set to wash out the puffed drug in the direction away from the targeted neuron. The AIS was detected by adding 100 µM Alexa Fluor 594 into the intracellular solution, and its diffusion into the axon was live monitored. The pipette puffing onto the AIS was placed near the distal AIS 25 to 30 µm away from the soma. The pipette puffing onto the nodes was placed near an axonal branch (detected from the Alexa Fluor 594 fill) or green fluorescent protein (GFP) signal detected along the axon of a Thy1-Caspr-GFP mouse. Caspr was post hoc immunolabeled to confirm the presence of nodes at the puffing sites.

### *Patch-clamping, calcium uncaging, and imaging of astrocytes*

Astrocytes were detected by their morphology and with the selective marker sulforhodamine 101 (SR 101). Before the experiments, the slices were incubated with 1 µM SR101 (Tocris Bioscience) added to the slicing solution for 20 min at 37°C. The presence of astrocytes was confirmed by post hoc fixing slices and immunolabeling for GFAP. Microelectrodes with resistances of 7 to 8 MΩ were filled with a K-methylsulfate intracellular solution containing (in mM): 100 $KMeSO_4$, 50 KCl, 2 $MgCl_2$, 4 MgATP,

0.3 $Na_2GTP$, and 10 HEPES. Fluo-4 (100 μM; ThermoFisher Scientific) was added to image calcium changes, and 5 mM NP-EGTA (ThermoFisher Scientific) was added to allow uncaging of calcium in single astrocytes. Images were acquired at 2 Hz with a confocal argon laser, and NP-EGTA photolysis was obtained by applying two-photon excitation at 720 nm to the soma (10 iterations, 2.55 μs pixel dwell time). Laser intensity was set at 10 mW and, if needed, was increased gradually until a calcium concentration rise was evoked (fig. S2). Intensities just below those needed to evoke a detectable $[Ca^{2+}]_i$ rise were used as control experiments to check that two-photon illumination alone was not responsible for the release of ATP and adenosine generation observed. In the experiments examining calcium activity at astrocyte processes, images were acquired at the plane of the axon or of a dendrite contacting astrocyte processes. Regions of interest (ROIs) were selected and $[Ca^{2+}]_i$ changes were measured as $\Delta F/F$ after background subtraction. At the start of the experiments, the baseline astrocyte $Ca^{2+}$ activity was recorded for 10 s, and, over this brief period, no spontaneous $Ca^{2+}$ transients were observed.

### Immunohistochemistry

Brains from Thy1-Caspr-GFP mice were perfusion-fixed in 4% paraformaldehyde (PFA) in 0.01 M phosphate-buffered saline (PBS), and fixed tissue was then cut into 70-μm-thick slices. Alternatively, 300-μm-thick acute slices from mouse and rat brains were immersion-fixed in a solution containing 4% PFA, 4% sucrose, and 0.1 M PBS. The slices were permeabilized with 0.2% Triton X-100 (Sigma-Aldrich) in a blocking solution (10% goat serum in 0.01 M PBS) for 1 hour. The slices were then incubated overnight at 4°C with the following primary antibodies, as required: rabbit anti-$A_{2a}R$ (Abcam, ab3461, 1:100), mouse anti-$A_{2a}R$ [Millipore, 05-717, which has been validated as giving no labeling in $A_{2a}R$ KO tissue (*40*), 1:200], rabbit anti-$A_1R$ (Abcam, ab82477, 1:100), rabbit anti-$A_{2b}R$ (Cohesion, CPA3755, 1:100), mouse anti-Ankyrin G (Neuromab, N106/36, 1:500), mouse anti-panNav (Sigma-Aldrich, K58/35, 1:100), mouse anti-Caspr (Neuromab, K65/35, 1:100), chicken anti-GFP (Millipore, AB16901, 1:1000), rabbit anti-GFAP (Millipore, AB5804, 1:500), rabbit anti-HCN1 (Alomone, APC 056, 1:100), and rabbit anti-HCN2 (Alomone, APC 030, 1:100). The slices were then incubated for 2 hours at room temperature with the following secondary antibodies, as required (ThermoFisher Scientific, 1:500): anti-chicken or anti-mouse Alexa Fluor 488, anti-rabbit or anti-mouse Alexa Fluor 546, and anti-rabbit Alexa Fluor 647. The slices were incubated with 4′,6-diamidino-2-phenylindole nuclear stain (1:50,000 in PBS from a 5 mg/ml stock concentration) for 10 min and mounted in Dako fluorescent mounting medium. Slices were washed with 0.01 M PBS three times with 10 min between each step. Negative control experiments were also performed to check for any labeling caused by unspecific binding of the secondary antibodies. For this, we followed the same protocol but omitted the incubation with the primary antibodies. Slices were imaged using a Zeiss LSM700 confocal microscope with a Zeiss Plan-Apochromat 63× oil immersion lens, and images were acquired with ZEN Microscope Software (Zeiss). *Z*-stacks of 1-μm intervals were imaged, and their maximum intensity was projected using ImageJ (FIJI). A line was drawn along the axons to plot intensity profiles. To plot node of Ranvier profiles, the width of the line was adjusted to fit the width of the Caspr-labeled paranodes. To analyze GFAP staining running parallel to axons, the width of the line was adjusted to 5 μm and centered on the axon.

### Quantification of putative ATP-containing vesicles

Before the experiments, the slices were incubated with 20 μM quinacrine dihydrochloride (Sigma-Aldrich) added to the slicing solution for 25 min at 37°C. Quinacrine is a green selective marker for ATP-containing vesicles (*7*) or lysosomes (*41*). Calcium was uncaged and imaged from single astrocytes as described above using the red $Ca^{2+}$ indicator Rhod-2 (50 μM, ThermoFisher Scientific) instead of Fluo-4. Images of astrocyte processes with ATP vesicles were acquired before and after $Ca^{2+}$ uncaging with two-photon excitation at 720 nm at the soma. Adjacent areas were also imaged. ROIs located a similar distance away from the astrocyte soma but without any detectable dye-filled astrocyte processes within a 5 μm perimeter minimum were selected. With ImageJ, a constant threshold was applied to the images and particle analysis was used to quantify the number of ATP vesicles. The mean Feret's diameter of ATP vesicles was 0.56 ± 0.03 μm ($n$ = 2227), similar to previous reports (*42*).

### Detection of extracellular ATP

Changes in extracellular ATP level were detected with a luciferin/luciferase–based chemiluminescence assay emitting green light. Perfusion of brain slices was halted and 50-μl aliquots containing 12 mg/ml luciferin (Sigma-Aldrich) and 5 mg/ml luciferase (Sigma-Aldrich) were added to a 1-ml bath chamber filled with aCSF at room temperature (because luciferase is not stable at higher temperatures). Calcium was uncaged and imaged from single astrocytes as described above using the red $Ca^{2+}$ indicator Rhod-2 (50 μM) instead of Fluo-4. A field of view containing an astrocyte and its processes was selected and ATP-derived bioluminescence was collected in darkness using highly sensitive GaAsP detectors with a Zeiss LSM780 two-photon/confocal microscope.

### Estimating axonal conduction velocity

We estimated axonal conduction velocities on the basis of parameters acquired experimentally from dual soma and axon bleb patch-clamp recordings, live confocal imaging, and post hoc immunostaining of layer V pyramidal neurons. In accordance with previous studies, two assumptions were made: (i) spikes initiate at the distal half of the AIS (the AIS being defined as being from the edge of the soma to the start of the first Caspr-labeled paranode) because in the distal AIS $Na^+$ channels are found in high density, have a low voltage threshold, and are isolated from the capacitive load of the soma (*43*, *44*) and (ii) forward conduction speed is two to three times faster than backward conduction speed because of the lower $Na^+$ channel density in the proximal AIS and the capacitive load of the soma (*16*, *44*). The axonal conduction speed can be then calculated as follows:

$$v = \frac{x_{\text{bleb}} - 2x_{\text{soma}}}{T} \text{ or } v = \frac{x_{\text{bleb}} - 3x_{\text{soma}}}{T} \quad (3)$$

where $v$ is the speed from the spike initiation site to the axon bleb, $T$ is the delay time from the somatic signal to the bleb signal (obtained from dual patch-clamp recordings), and $x_{\text{bleb}}$ and $x_{\text{soma}}$ are the distances from the spike initiation site to the recording sites on the axon bleb or on the soma, respectively (obtained from confocal imaging by assuming a position for spike initiation).

### Computer simulations of infinite corpus callosal axon

As in our previous studies (*17*, *45*, *46*), action potential conduction along myelinated axons was simulated using MATLAB. Electrophysiological parameters were based on the finite impedance double-cable model (model C) of Richardson *et al.* (*47*), except that the membrane capacitance was taken as the physiologically measured value of 0.9 $\mu F/cm^2$ (*48*). The differential equations of the model were derived and solved as in the myelinated axon model of Halter and Clark (*49*), in which the axon is divided into compartments representing the node, paranode, and internode. The MATLAB code, including all parameters and equations, is available (see the Acknowledgments). Conduction speed simulations of long callosal-like axons were performed as described previously (*17*), where the parameters used were based on experimental data. The speed was measured between nodes 20 and 30 in a uniform axon containing 51 nodes and 50 internodes of constant lengths (nodes: 1.5 μm, internodes: 81.7 μm) and diameters (nodes: 0.64 μm, internodes: 0.73 μm). The periaxonal space thickness, which has an

important effect on conduction speed (*17*, *50*), was set at 15 nm except in the paranodes (the 1.9-μm end parts of the internodes), where the width was reduced to 0.0123 nm (*17*). Assuming a myelin wrap periodicity of 15.6 nm, five myelin wraps were needed to set the g-ratio close to 0.8. Each node expressed fast voltage-gated $Na^+$ ($Na_V$) channels, persistent $Na_V$ channels, and slow voltage-gated potassium ($K_V$) channels at fully activated conductance densities of 10, 0.01, and 0.4 mS/mm$^2$, respectively, as well as adenosine-gated $I_h$ channels at a density of either 0 (for no $A_{2a}R$ activation) or 0.1565 (to mimic $A_{2a}R$ activation) mS/mm$^2$. This density was experimentally derived as follows: the maximum $I_h$ conductance (derived from tail current analysis as in Fig. 3, I and J) evoked in the ~10-μm length of the distal AIS was divided by the area of the distal AIS to obtain a fully-activated conductance density of 0.111 mS/mm$^2$. Immunohistochemical labeling of $A_{2a}$Rs then showed their density at nodes of Ranvier to be a factor of 1.41 higher than in the distal AIS (calculated by integrating the background-subtracted fluorescent signal across a small length $w$ of the axon and dividing by $\pi \bullet w \bullet d$, where $d$ is the diameter of the axon), implying a density at the nodes of 0.1565 mS/mm$^2$. The node resting potential in the absence of adenosine was set to −82 mV by adjusting the magnitude of a leak conductance (to 0.113 mS/mm$^2$) with a reversal potential of −84 mV (the reversal potential of the other $K^+$ currents in the model). The leak current represents $K^+$ leak channels such as TRAAK present in the nodes of Ranvier (*51*, *52*).

### Computer simulations of patch-clamped neuron with soma, AIS, and three internodes

This model was then modified so that the dimensions of a specimen patch-clamped and imaged neuron acquired for the present study (Fig. 4) were implemented (Fig. 5A). The model included "nodes" representing the soma, the proximal AIS, the distal AIS, three nodes of Ranvier, and a terminating bleb. To represent the area and thus the capacitance of the real soma, the diameter and length of the cylindrical "node" representing the soma were both set to 18 μm. The proximal and distal AIS had lengths of 21.9 and 10 μm, respectively, and diameters (measured experimentally as the average over each of these two zones) of 1.046 and 0.525 μm respectively. The soma and proximal AIS, and the proximal and distal AIS, were separated by extremely small ($10^{-7}$ μm long) "internodes" (which are essential for the alternating node-internode model to function in MATLAB; not shown in Fig. 5A) with diameters of 1.046 and 0.79 μm, respectively. Distal to the AIS are three internodes of lengths 24.91, 52.44, and 39.20 μm and diameters of 0.51, 0.56, and 0.53 μm respectively, each followed by nodes of Ranvier of lengths 1.126, 0.915, and 0.414 μm, and diameters of 0.445, 0.452, and 0.401 μm, respectively. The third node of Ranvier was followed by an extremely small ($10^{-7}$ μm long) "internode" of diameter 0.4 μm and then a terminal bleb of length 1.64 μm and diameter 3 μm. Each node of Ranvier expressed the same conductance as in the infinite cable model. The proximal and distal AIS compartments expressed fast $Na_V$, persistent $Na_V$, and slow $K_V$ channels at the same densities as in the infinite cable model, and in addition a low threshold $K_V$ channel with a maximum conductance of 0.4 mS/mm$^2$, which was needed to allow the generation of repetitive action potential trains in response to current injected at the soma. Adenosine-insensitive $I_h$ conductance (with a half-maximum activation voltage of −103.5 mV, as found on average in experiments as in Fig. 3J) was expressed in the soma and proximal AIS (as seen immunohistochemically for HCN1 channels) at a density of 0.0033 mS/mm$^2$ (set so as to reproduce the mean $I_h$ maximal conductance seen in the absence of adenosine in Fig. 3, I and J), whereas adenosine-sensitive $I_h$ was present in the distal AIS at a density of 0.111 mS/mm$^2$ and in the nodes of Ranvier at a density 1.41-fold higher (see above). The density used for the distal AIS was set to increase the maximum $I_h$ conductance measured at the soma by 51% as observed during activation of $A_{2a}$Rs in the distal AIS (Fig. 3, I and J), and the midpoint of its activation curve was at −80 mV to reproduce the positive shift for the total $I_h$ apparent activation curve in Fig. 3J. The terminal bleb expressed no voltage-gated conductances. To match the model to the experimental data, we wanted to reproduce the 5.8-mV depolarization recorded at the soma when puffed adenosine activated $I_h$ in the distal AIS (Fig. 3E). We found that if we set the resting potential to −82 mV in the soma and AIS using a leak conductance with a reversal potential of −84 mV (as in the nodes), then this introduced too much resting conductance, reducing the depolarization evoked at the soma when $A_{2a}$Rs were activated in the distal AIS to only +1.55 mV. Accordingly, in the soma, AIS, and terminal bleb, we used a nominal leak reversal potential of −4000 mV, converting the leak into an essentially voltage-independent (~zero conductance) current, similar conceptually to the constant outward current generated by the $Na^+/K^+$ pump. The adenosine-sensitive $I_h$ (when activated in the distal AIS) then generated a depolarization of 5.9 mV at the soma. To investigate the effect of the adenosine-evoked $I_h$ current, the action potential speed was measured between the first and the last nodes of Ranvier when current was injected briefly (1 nA for 100 μsec) at the soma. The model reproduced reasonably well the assumptions that we made to assess the conduction speed in our experiments: In the simulations, spikes evoked by current injection at the soma were initiated in the distal AIS "node", the forward conduction speed was faster than the backward speed (from distal AIS to bleb: 1.33 m/s; from distal AIS to soma: 0.23 m/s), and the conduction speed from the distal AIS to the bleb (1.33 m/s) was in the range estimated from the experiments in Figs. 4E and 6I (1.31 to 2.23 m/s).

### Code availability

The computer code used for the simulations described above has been deposited in Zenodo (*53*).

### Statistics

Statistical analyses and graphs were performed with GraphPad Prism 6. Data are presented as mean ± SEM. Data normality was assessed with D'Agostino & Pearson omnibus or Kolmogorov-Smirnov tests. Comparisons of normally distributed two groups were made using two-tailed Student $t$ tests. When treatments involved more than two independent groups, statistical comparisons were performed with one-way ANOVA and post hoc Tukey's multiple comparisons. Data that were not normally distributed were analyzed with Mann-Whitney test or Wilcoxon matched-pairs signed rank test. Assessment of whether the slope of linear regressions differed significantly from zero was obtained using the $t$ statistic for the slope. $P$ values within any figure panel were adjusted for multiple comparisons.

### Gender and species effects

Mice were used in experiments when it was essential to have the nodes labeled with Caspr-GFP. Rats were also used because they are more generally available at all ages in our university. We observed no differences between the two species, for example, in

(i) the presence of $A_{2a}$Rs and HCN channels at nodes of Ranvier and the AIS;

(ii) the initial speed and the percentage change evoked by CGS 21680 puff onto nodes, which did not differ significantly between rats and mice ($P$ = 0.43 and $P$ = 0.15, respectively); and

(iii) the initial speed and the percentage change evoked by $Ca^{2+}$ uncaging in perinodal astrocytes, which did not differ significantly between rats and mice ($P$ = 0.08 and $P$ = 0.93, respectively).

Animals of both sexes were used because of increasing awareness of the fact that gender can play a role in determining function. For major results, we found no differences between the two sexes, for example, in

(i) the presence of $A_{2a}$Rs and HCN channels at nodes of Ranvier and the AIS;

(ii) the initial resting membrane potential and the change evoked by CGS 21680 puff onto the AIS, which did not differ significantly

between males and females ($P$ = 0.69 and 0.75 respectively); and

(iii) the initial speed and the percentage change evoked by CGS 21680 puff onto nodes, which did not differ significantly between males and females ($P$ = 0.3 and $P$ = 0.45, respectively).

## REFERENCES AND NOTES


1. N. J. Allen, Astrocyte regulation of synaptic behavior. *Annu. Rev. Cell Dev. Biol.* **30**, 439–463 (2014). doi: 10.1146/annurev-cellbio-100913-013053; pmid: 25288116
2. M. H. P. Kole *et al*., Action potential generation requires a high sodium channel density in the axon initial segment. *Nat. Neurosci.* **11**, 178–186 (2008). doi: 10.1038/nn2040; pmid: 18204443
3. V. A. Larson *et al*., Oligodendrocytes control potassium accumulation in white matter and seizure susceptibility. *eLife* **7**, e34829 (2018). doi: 10.7554/eLife.34829; pmid: 29596047
4. A. S. Saab, I. D. Tzvetanova, K.-A. Nave, The role of myelin and oligodendrocytes in axonal energy metabolism. *Curr. Opin. Neurobiol.* **23**, 1065–1072 (2013). doi: 10.1016/j.conb.2013.09.008; pmid: 24094633
5. C. Ffrench-Constant, R. H. Miller, J. Kruse, M. Schachner, M. C. Raff, Molecular specialization of astrocyte processes at nodes of Ranvier in rat optic nerve. *J. Cell Biol.* **102**, 844–852 (1986). doi: 10.1083/jcb.102.3.844; pmid: 2419343
6. A. Araque *et al*., Gliotransmitters travel in time and space. *Neuron* **81**, 728–739 (2014). doi: 10.1016/j.neuron.2014.02.007; pmid: 24559669
7. S. Forveille *et al*., Quinacrine-mediated detection of intracellular ATP. *Methods Enzymol.* **629**, 103–113 (2019). doi: 10.1016/bs.mie.2019.05.050; pmid: 31727236
8. T. Pfeiffer, D. Attwell, Brain's immune cells put the brakes on neurons. *Nature* **586**, 366–367 (2020). doi: 10.1038/d41586-020-02713-7; pmid: 32999439
9. K. Baalman *et al*., Axon initial segment-associated microglia. *J. Neurosci.* **35**, 2283–2292 (2015). doi: 10.1523/JNEUROSCI.3751-14.2015; pmid: 25653382
10. J. Zhang, X. Yang, Y. Zhou, H. Fox, H. Xiong, Direct contacts of microglia on myelin sheath and Ranvier's node in the corpus callosum in rats. *J. Biomed. Res.* **33**, 192–200 (2019). pmid: 30403198
11. D. DiFrancesco, P. Tortora, Direct activation of cardiac pacemaker channels by intracellular cyclic AMP. *Nature* **351**, 145–147 (1991). doi: 10.1038/351145a0; pmid: 1709448
12. G. Vargas, M. T. Lucero, Dopamine modulates inwardly rectifying hyperpolarization-activated current ($I_h$) in cultured rat olfactory receptor neurons. *J. Neurophysiol.* **81**, 149–158 (1999). doi: 10.1152/jn.1999.81.1.149; pmid: 9914276
13. N. Byczkowicz *et al*., HCN channel-mediated neuromodulation can control action potential velocity and fidelity in central axons. *eLife* **8**, e42766 (2019). doi: 10.7554/eLife.42766; pmid: 31496517
14. F. C. Roth, H. Hu, An axon-specific expression of HCN channels catalyzes fast action potential signaling in GABAergic interneurons. *Nat. Commun.* **11**, 2248 (2020). doi: 10.1038/s41467-020-15791-y; pmid: 32382046
15. K. W. Ko, M. N. Rasband, V. Meseguer, R. H. Kramer, N. L. Golding, Serotonin modulates spike probability in the axon initial segment through HCN channels. *Nat. Neurosci.* **19**, 826–834 (2016). doi: 10.1038/nn.4293; pmid: 27110919
16. Y. Yu, Y. Shu, D. A. McCormick, Cortical action potential backpropagation explains spike threshold variability and rapid-onset kinetics. *J. Neurosci.* **28**, 7260–7272 (2008). doi: 10.1523/JNEUROSCI.1613-08.2008; pmid: 18632930
17. I. L. Arancibia-Cárcamo *et al*., Node of Ranvier length as a potential regulator of myelinated axon conduction speed. *eLife* **6**, e23329 (2017). doi: 10.7554/eLife.23329; pmid: 28130923
18. A. Lörincz, T. Notomi, G. Tamás, R. Shigemoto, Z. Nusser, Polarized and compartment-dependent distribution of HCN1 in pyramidal cell dendrites. *Nat. Neurosci.* **5**, 1185–1193 (2002). doi: 10.1038/nn962; pmid: 12389030
19. B. J. Wainger, M. DeGennaro, B. Santoro, S. A. Siegelbaum, G. R. Tibbs, Molecular mechanism of cAMP modulation of HCN pacemaker channels. *Nature* **411**, 805–810 (2001). doi: 10.1038/35081088; pmid: 11459060
20. N. B. Hamilton, D. Attwell, Do astrocytes really exocytose neurotransmitters? *Nat. Rev. Neurosci.* **11**, 227–238 (2010). doi: 10.1038/nrn2803; pmid: 20300101
21. B. S. Khakh, Astrocyte-neuron interactions in the striatum: Insights on identity, form, and function. *Trends Neurosci.* **42**, 617–630 (2019). doi: 10.1016/j.tins.2019.06.003; pmid: 31351745
22. N. Hamilton *et al*., Mechanisms of ATP- and glutamate-mediated calcium signaling in white matter astrocytes. *Glia* **56**, 734–749 (2008). doi: 10.1002/glia.20649; pmid: 18293404
23. A. Natsubori *et al*., Intracellular ATP levels in mouse cortical excitatory neurons varies with sleep-wake states. *Commun. Biol.* **3**, 491 (2020). doi: 10.1038/s42003-020-01215-6; pmid: 32895482
24. T. Porkka-Heiskanen *et al*., Adenosine: A mediator of the sleep-inducing effects of prolonged wakefulness. *Science* **276**, 1265–1268 (1997). doi: 10.1126/science.276.5316.1265; pmid: 9157887
25. M. M. Halassa *et al*., Astrocytic modulation of sleep homeostasis and cognitive consequences of sleep loss. *Neuron* **61**, 213–219 (2009). doi: 10.1016/j.neuron.2008.11.024; pmid: 19186164
26. L. I. Schmitt, R. E. Sims, N. Dale, P. G. Haydon, Wakefulness affects synaptic and network activity by increasing extracellular astrocyte-derived adenosine. *J. Neurosci.* **32**, 4417–4425 (2012). doi: 10.1523/JNEUROSCI.5689-11.2012; pmid: 22457491
27. Z. L. Huang *et al*., Adenosine A2A, but not A1, receptors mediate the arousal effect of caffeine. *Nat. Neurosci.* **8**, 858–859 (2005). doi: 10.1038/nn1491; pmid: 15965471
28. W. Wefelmeyer, D. Cattaert, J. Burrone, Activity-dependent mismatch between axo-axonic synapses and the axon initial segment controls neuronal output. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 9757–9762 (2015). doi: 10.1073/pnas.1502902112; pmid: 26195803
29. K. J. Bender, C. P. Ford, L. O. Trussell, Dopaminergic modulation of axon initial segment calcium channels regulates action potential initiation. *Neuron* **68**, 500–511 (2010). doi: 10.1016/j.neuron.2010.09.026; pmid: 21040850
30. P. König, A. K. Engel, W. Singer, Integrator or coincidence detector? The role of the cortical neuron revisited. *Trends Neurosci.* **19**, 130–137 (1996). doi: 10.1016/S0166-2236(96)80019-1; pmid: 8658595
31. E. Salinas, T. J. Sejnowski, Correlated neuronal activity and the flow of neural information. *Nat. Rev. Neurosci.* **2**, 539–550 (2001). doi: 10.1038/35086012; pmid: 11483997
32. C. Adaikkan, L.-H. Tsai, Gamma entrainment: Impact on neurocircuits, glia, and therapeutic opportunities. *Trends Neurosci.* **43**, 24–41 (2020). doi: 10.1016/j.tins.2019.11.001; pmid: 31836315
33. S. Pajevic, P. J. Basser, R. D. Fields, Role of myelin plasticity in oscillations and synchrony of neuronal activity. *Neuroscience* **276**, 135–147 (2014). doi: 10.1016/j.neuroscience.2013.11.007; pmid: 24291730
34. C. Ledent *et al*., Aggressiveness, hypoalgesia and high blood pressure in mice lacking the adenosine $A_{2a}$ receptor. *Nature* **388**, 674–678 (1997). doi: 10.1038/41771; pmid: 9262401
35. M. P. Kaster *et al*., Caffeine acts through neuronal adenosine $A_{2A}$ receptors to prevent mood and memory dysfunction triggered by chronic stress. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7833–7838 (2015). doi: 10.1073/pnas.1423088112; pmid: 26056314
36. Y.-J. Liu *et al*., Research progress on adenosine in central nervous system diseases. *CNS Neurosci. Ther.* **25**, 899–910 (2019). doi: 10.1111/cns.13190; pmid: 31334608
37. D. J. Hines, P. G. Haydon, Astrocytic adenosine: From synapses to psychiatric disorders. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130594 (2014). doi: 10.1098/rstb.2013.0594; pmid: 25225088
38. V. Brivio, C. Faivre-Sarrailh, E. Peles, D. L. Sherman, P. J. Brophy, Assembly of CNS nodes of Ranvier in myelinated nerves is promoted by the axon cytoskeleton. *Curr. Biol.* **27**, 1068–1073 (2017). doi: 10.1016/j.cub.2017.01.025; pmid: 28318976
39. Y. Shu, A. Hasenstaub, A. Duque, Y. Yu, D. A. McCormick, Modulation of intracortical synaptic potentials by presynaptic somatic membrane potential. *Nature* **441**, 761–765 (2006). doi: 10.1038/nature04720; pmid: 16625207
40. A. G. Orr *et al*., Astrocytic adenosine receptor $A_{2A}$ and $G_s$-coupled signaling regulate memory. *Nat. Neurosci.* **18**, 423–434 (2015). doi: 10.1038/nn.3930; pmid: 25622143
41. Z. Zhang *et al*., Regulated ATP release from astrocytes through lysosome exocytosis. *Nat. Cell Biol.* **9**, 945–953 (2007). doi: 10.1038/ncb1620; pmid: 17618272
42. C. Cserép *et al*., Microglia monitor and protect neuronal function through specialized somatic purinergic junctions. *Science* **367**, 528–537 (2020). doi: 10.1126/science.aax6752; pmid: 31831638
43. W. Hu *et al*., Distinct contributions of Na(v)1.6 and Na(v)1.2 in action potential initiation and backpropagation. *Nat. Neurosci.* **12**, 996–1002 (2009). doi: 10.1038/nn.2359; pmid: 19633666
44. G. Baranauskas, Y. David, I. A. Fleidervish, Spatial mismatch between the $Na^+$ flux and spike initiation in axon initial segment. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 4051–4056 (2013). doi: 10.1073/pnas.1215125110; pmid: 23341597
45. M. C. Ford *et al*., Tuning of Ranvier node and internode properties in myelinated axons to adjust action potential timing. *Nat. Commun.* **6**, 8073 (2015). doi: 10.1038/ncomms9073; pmid: 26305015
46. Y. Bakiri, R. Káradóttir, L. Cossell, D. Attwell, Morphological and electrical properties of oligodendrocytes in the white matter of the corpus callosum and cerebellum. *J. Physiol.* **589**, 559–573 (2011). doi: 10.1113/jphysiol.2010.201376; pmid: 21098009
47. A. G. Richardson, C. C. McIntyre, W. M. Grill, Modelling the effects of electric fields on nerve fibres: Influence of the myelin sheath. *Med. Biol. Eng. Comput.* **38**, 438–446 (2000). doi: 10.1007/BF02345014; pmid: 10984943
48. L. J. Gentet, G. J. Stuart, J. D. Clements, Direct measurement of specific membrane capacitance in neurons. *Biophys. J.* **79**, 314–320 (2000). doi: 10.1016/S0006-3495(00)76293-X; pmid: 10866957
49. J. A. Halter, J. W. Clark Jr.., A distributed-parameter model of the myelinated nerve fiber. *J. Theor. Biol.* **148**, 345–382 (1991). doi: 10.1016/S0022-5193(05)80242-5; pmid: 2016898
50. K. M. Young *et al*., Oligodendrocyte dynamics in the healthy adult CNS: Evidence for myelin remodeling. *Neuron* **77**, 873–885 (2013). doi: 10.1016/j.neuron.2013.01.006; pmid: 23473318
51. H. Kanda *et al*., TREK-1 and TRAAK are principal $K^+$ channels at the nodes of Ranvier for rapid action potential conduction on mammalian myelinated afferent nerves. *Neuron* **104**, 960–971.e7 (2019). doi: 10.1016/j.neuron.2019.08.042; pmid: 31630908
52. S. G. Brohawn *et al*., The mechanosensitive ion channel TRAAK is localized to the mammalian node of Ranvier. *eLife* **8**, e50403 (2019). doi: 10.7554/eLife.50403; pmid: 31674909
53. Code used in the simulations for: J. Lezmy, I. L. Arancibia-Cárcamo, T. Quintela-López, D. L. Sherman, P. J. Brophy, D. Attwell, Astrocyte $Ca^{2+}$-evoked ATP release regulates myelinated axon excitability and conduction speed, Zenodo (2021); doi: 10.5281/zenodo.5177076
54. Code used to simulate the action potential in myelinated axons for: J. Lezmy, I. L. Arancibia-Cárcamo, T. Quintela-López, D. L. Sherman, P. J. Brophy, D. Attwell, Astrocyte $Ca^{2+}$-evoked ATP release regulates myelinated axon excitability and conduction speed, Github (2021); https://github.com/AttwellLab/MyelinatedAxonModel.


## ACKNOWLEDGMENTS

We thank A. Gourine, M. Häusser, C. Madry, D. Rusakov, M. Shah, H. Shiina, and V. Vyazovskiy for comments on the manuscript. **Funding:** This work was supported by European Research Council (BrainEnergy) and Wellcome Trust Investigator Awards (099222/Z/12/Z and 219366/Z/19/Z) to D.A., an EMBO fellowship (ALTF 430-2019) to J.L., and a Wellcome Trust Investigator Award (107008) to P.J.B. This research was funded in part by grants 099222/Z/12/Z and 219366/Z/19/Z from the Wellcome Trust, a cOAlition S organization. The authors will make the Author Accepted Manuscript (AAM) version available under a CC BY public copyright license. **Author contributions:** J.L. performed all electrophysiological and live imaging experiments. T.Q. and J.L. performed the immunohistochemistry experiments. L.A., J.L., and D.A. performed computer simulations. D.L.S. and P.J.B. produced the Thy1-Caspr-GFP mouse. J.L. and D.A. conceived the project, analyzed the data, and wrote the first draft of the manuscript. All other authors commented on the manuscript. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. Thy1-Caspr-GFP mice are available under a material transfer agreement with the University of Edinburgh from P.J.B. and D.L.S. Code used for simulations is freely available (*53*, *54*).

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abh2858
Figs. S1 to S7
Movies S1 to S3
Reference (*55*)
MDAR Reproducibility Checklist

2 March 2021; resubmitted 5 July 2021
Accepted 1 September 2021
10.1126/science.abh2858

RESEARCH ARTICLES

CATALYSIS

# Univariate classification of phosphine ligation state and reactivity in cross-coupling catalysis

Samuel H. Newman-Stonebraker[1]†‡, Sleight R. Smith[2]†, Julia E. Borowski[1], Ellyn Peters[2], Tobias Gensch[2]§, Heather C. Johnson[3]¶, Matthew S. Sigman[2]*, Abigail G. Doyle[1]*‡

Chemists often use statistical analysis of reaction data with molecular descriptors to identify structure-reactivity relationships, which can enable prediction and mechanistic understanding. In this study, we developed a broadly applicable and quantitative classification workflow that identifies reactivity cliffs in 11 Ni- and Pd-catalyzed cross-coupling datasets using monodentate phosphine ligands. A distinctive ligand steric descriptor, minimum percent buried volume [%$V_{bur}$ (min)], is found to divide these datasets into active and inactive regions at a similar threshold value. Organometallic studies demonstrate that this threshold corresponds to the binary outcome of bisligated versus monoligated metal and that %$V_{bur}$ (min) is a physically meaningful and predictive representation of ligand structure in catalysis.

Chemists commonly use data-driven modeling to understand the many complex relationships connecting chemical structure to reactivity (*1*, *2*). For many reactions, a continuous dependence is observed between a descriptor—a mathematical way to describe a subunit or the entirety of a molecule—and chemical reactivity, as captured by linear free-energy relationships (*3*, *4*). However, some processes exhibit reactivity cliffs, wherein a criterion or threshold value of a given descriptor must be met for the reaction to occur (i.e., a binary response) (*5*). Identification of the molecular feature that is mechanistically linked to a reactivity cliff could enable the classification of molecules by structure, predict reaction outcomes for unseen examples, and reveal key mechanistic insights.

Transition metal–catalyzed cross-coupling represents a relevant case study for identifying reactivity cliffs. This reaction class is synthetically important because of its widespread use in the synthesis of pharmaceuticals (*6*) and materials (*7*). In addition, the success of these reactions is highly dependent on ancillary ligand identity, with monodentate phosphines among the most frequently used (*8*, *9*) (Fig. 1A). Given that thousands of unique structural examples exist, chemists have developed numerous tools to quantitatively describe the diverse steric and electronic properties of monodentate phosphines (*10*). These descriptors—including Tolman cone angle (*11*, *12*), solid angle (*13*), Sterimol (*14*), and percent buried volume (%$V_{bur}$) (*15*–*17*) (Fig. 1B)—have been used to correlate structure-reactivity trends in cross-coupling datasets (*2*, *18*). However, there are situations in which seemingly similar ligands afford substantially different reactivities (*19*), suggesting the presence of discontinuities in ligand reactivity (*20*).

Considering this, we hypothesized that these discontinuities were evidence of reactivity cliffs and linked to a phosphine structural feature yet to be defined. Identification of such a descriptor would not only allow for the development of a statistical means to classify active and inactive ligands but also provide a quantitative tool for mechanistically rationalizing ligand performance. Herein, we provide a workflow and analysis to achieve these goals. Using the organophosphorus(III) descriptor library recently developed by Gensch *et al.* (*kraken*) that possesses nearly 200 conformationally representative descriptors for each of several thousand monodentate phosphines (*21*) (Fig. 1C), we were able to classify 11 Ni- and Pd-catalyzed cross-coupling datasets into active and inactive regions of reactivity based on whether the metal preferentially coordinates to one or two phosphines (i.e., catalyst ligation state). Spectroscopic and crystallographic organometallic studies demonstrate that minimum percent buried volume [%$V_{bur}$ (min)] is the single descriptor able to achieve this classification. Although %$V_{bur}$ describes the steric bulk of any given ligand structure within 3.5 Å of the metal center, %$V_{bur}$ (min), a previously unexplored variant of the descriptor, quantifies the smallest %$V_{bur}$ among all of the ligand's energetically accessible conformers. Ultimately, this analysis revealed nonintuitive trends in organometallic chemistry and can thereby serve as an important mechanistic tool to understand and predict monodentate phosphine structure-reactivity relationships and catalyst ligation state in cross-coupling catalysis (Fig. 1D).

[1]Department of Chemistry, Princeton University, Princeton, NJ 08544, USA. [2]Department of Chemistry, University of Utah, Salt Lake City, UT 84112, USA. [3]Process Research and Development, Merck & Co., Inc., Rahway, NJ 07065, USA.
***Corresponding author. Email: sigman@chem.utah.edu (M.S.S.); agdoyle@chem.ucla.edu (A.G.D.)**
†These authors contributed equally to this work.
‡Present address: Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, USA.
§Present address: Department of Chemistry, TU Berlin, 10623 Berlin, Germany.
¶Present address: Drug Substance Technologies, Process Development, Amgen, Inc., Cambridge, MA 02142, USA.

### Exploration of phosphine steric descriptors in Ni catalysis

Our initial platform for probing the presence of ligand reactivity cliffs was inspired by a recent study from the Doyle group that identified a new class of monodentate phosphine ligands—the DinoPhos ligands [TriceraPhos and TyrannoPhos (Fig. 1A)]—as distinctively effective for the Ni-catalyzed cross-coupling of acetals with aryl boroxines (*19*). The characteristic that distinguished the DinoPhos ligands from those that were less successful in the reaction was found to be their high level of remote steric bulk, quantified by both a large cone angle and small %$V_{bur}$ value. The large cone angles of the DinoPhos ligands suggested that they may behave like the sterically bulky phosphines developed for Pd-catalyzed cross-couplings [e.g., P(*t*-Bu)$_3$ or CyJohnPhos (*t*-Bu, *tert*-butyl; Cy, cyclohexyl); Fig. 1A], where preferential formation of monoligated (denoted as L$_1$) Pd complexes has been shown to be critical for reaction success (*22*–*26*). However, the small %$V_{bur}$ characterized these ligands as similar to smaller structures [e.g., PPh$_3$ (Ph, phenyl); Fig. 1A] that tend to form bisligated (denoted as L$_2$) complexes. The observation that this distinctive steric profile was essential for reaction success in Ni catalysis raised questions about its impact on catalyst ligation state and prompted us to use Ni-catalyzed cross-coupling reactions as a case study for the identification of ligand reactivity cliffs.

The original study by the Doyle lab included 19 phosphines, which we supplemented with 15 additional ligands to span the entire range of %$V_{bur}$ (reaction I; Fig. 2A). We also collected data on Ni-catalyzed Csp$^2$–Csp$^2$ Suzuki-Miyaura coupling (SMC) reactions with aryl chlorides using high-throughput experimentation. The ligand set was assembled from a combined Doyle and Merck Sharp & Dohme inventory to encompass 90 monodentate phosphines that are included in the *kraken* virtual library (*21*), with *k*-means clustering applied to enable a selection of structures that broadly covered phosphine chemical space (*27*). These ligands were evaluated for four cross-coupling reactions wherein the aryl halide and aryl boronic acid were altered (reactions II to V; Fig. 2A). The coupling partners included a range of electronic and steric features that could affect different aspects of the elementary steps of the catalytic cycle. Notably, the DinoPhos ligands were among the highest yielding in each of these reactions, along with DrewPhos (*28*), a triarylphosphine ligand that possesses *t*-Bu groups at the remote 3,5 positions of the arenes, similar to the DinoPhos ligands.

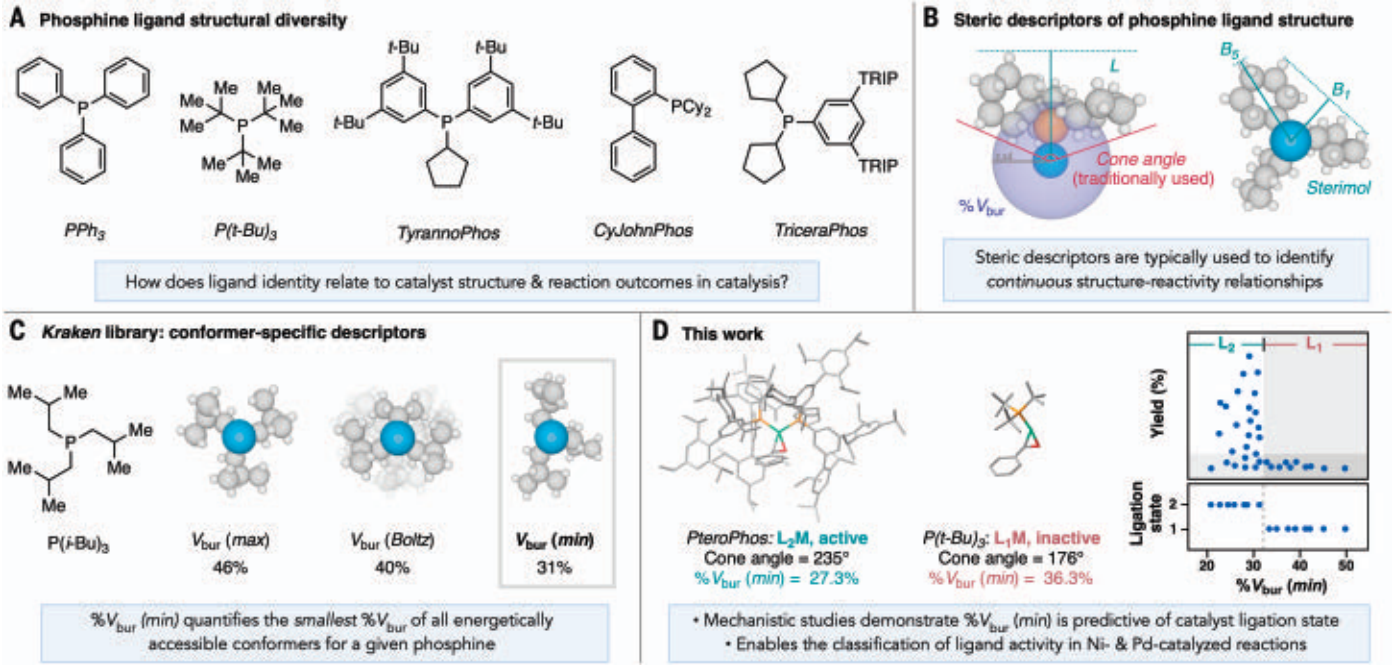



**Fig. 1. Studying phosphine ligand reactivity using structural descriptors.** (**A**) Examples of monodentate phosphines used in Ni and Pd cross-coupling reactions, including TyrannoPhos and TriceraPhos (DinoPhos ligands) recently reported by the Doyle lab (*19*). TRIP, 2,4,6-triisopropylphenyl. (**B**) Commonly used methods for quantifying phosphine steric properties. Cone angle is the traditionally used descriptor for quantifying monodentate phosphine steric bulk and is defined as the angular width (in degrees) of an imaginary cone needed to encapsulate the entire phosphine structure; the vertex of the cone is defined by a metal atom bound to the ligand with a bond length of 2.28 Å. $\%V_{bur}$, a more modern descriptor designed initially to study N-heterocyclic carbenes, is defined as the volume percent of the phosphine's atoms that fill an imaginary sphere of 3.5 Å radius that is centered on a metal atom bound to the phosphine with a bond length of 2.28 Å. Sterimol descriptors B1 and B5 describe the lowest and highest width of the ligand perpendicular to the metal-phosphorus axis, respectively, and Sterimol L describes the ligand's length along that axis. (**C**) Phosphine descriptor library (*kraken*) capturing multiple ligand conformers, with maximum, Boltzmann average, and minimum $\%V_{bur}$ values of the conformational ensemble of P(*i*-Bu)$_3$ shown. (**D**) This work. L1, one equivalent of ligand bound to metal; L2, two equivalents of ligand bound to metal; M, metal.

With these data, we sought to determine how the *kraken* steric descriptors relate to reaction performance. For each reaction, the yield was evaluated against representative descriptors in the library; as three examples, the cone angle (*29*), Boltzmann averaged $\%V_{bur}$ of the ligand's conformational ensemble [$\%V_{bur}$ (Boltz)], and $\%V_{bur}$ of the library conformer with the smallest buried volume [$\%V_{bur}$ (min)] are shown in Fig. 2B. Whereas the cone angle descriptors did not provide clear reactivity cutoffs, both $\%V_{bur}$ (Boltz) and $\%V_{bur}$ (min) allowed for classification of ligand performance into "active" and "inactive" bins (details of the classification tool will be described below for quantitative demarcation of the thresholds), though several outliers were present with the former. The $\%V_{bur}$ (min) descriptor revealed sharp cutoffs in reactivity, wherein nearly all ligands above 32% were found to be unreactive. Ligands that were unsuccessful at promoting reactivity below 32% $V_{bur}$ (min) were grouped in their own bin, wherein electronic features and catalyst-poisoning functionality (e.g., cyano, carbonyl, and halide groups) appeared to be largely responsible for their inactivity. The use of $\%V_{bur}$ (min), but not $\%V_{bur}$ (Boltz), allowed for various highly flexible scaffolds such as P(*i*-Bu)$_3$, PBn$_3$, and small (dialkyl)-*ortho*-biaryl (Buchwald-type) phosphines (*23*, *30*, *31*), to be successfully classified (*i*-Bu, isobutyl; Bn, benzyl). Thus, this steric descriptor best categorized phosphines into active and inactive bins (tables S15 and S16), with a reactivity threshold of ~32% $V_{bur}$ (min) appearing across all reactions studied.

### Mechanistic origin of $\%V_{bur}$ (min) reactivity thresholds

The presence of reaction-independent reactivity cliffs relative to $\%V_{bur}$ (min) in these data prompted us to investigate the mechanistic basis for this descriptor's importance. Given the importance of ligation state in cross-coupling catalysis, we questioned whether this descriptor is predictive of the preference for forming $L_2M$ versus $L_1M$ complexes (M, metal). Historically, the development of large ligands—intuitively captured through the cone angle descriptor—is an established principle to favor $L_1M$ and thereby promote reactivity in Pd catalysis (*22*–*26*). Although TriceraPhos and TyrannoPhos both have cone angle values higher than many phosphines known to form $L_1$ species, $\%V_{bur}$ (min) categorizes their reactivity with those that form $L_2$ Ni and Pd complexes (*22*, *32*). Thus, this interpretation of the observed reactivity cliffs would suggest that steric bulk within the metal's first coordination sphere primarily governs the ligation state outcome, regardless of the overall size of the phosphine.

To interrogate this hypothesis, we surveyed a subset (28 phosphines) of the ligands used across the Ni-catalyzed SMC reactions to determine ligation state spectroscopically. $L_nNi$(4-fluorobenzaldehyde) was chosen as a platform, because its three nuclear magnetic resonance (NMR) handles ($^1$H, $^{19}$F, and $^{31}$P) provided a readout of the in situ ligation state at the metal center (figs. S4 to S87) upon reaction of Ni(COD)$_2$ (COD, 1,5-cyclooctadiene) with the aldehyde and two equivalents of phosphine (reaction VI; Fig. 3A) (*33*, *34*). Each ligand was assigned as forming $L_2Ni$ or $L_1Ni$ complexes based on the observed spectra. The results of these experiments were then evaluated with steric features from the library to determine which ones classified the ligands into regions of similar ligation states (Fig. 3B; see fig. S3 for additional descriptors). As observed with the cross-coupling reaction yields, cone angle was unable to partition the ligands. Indeed, the DinoPhos ligands and DrewPhos all formed $L_2Ni$ complexes exclusively, despite having

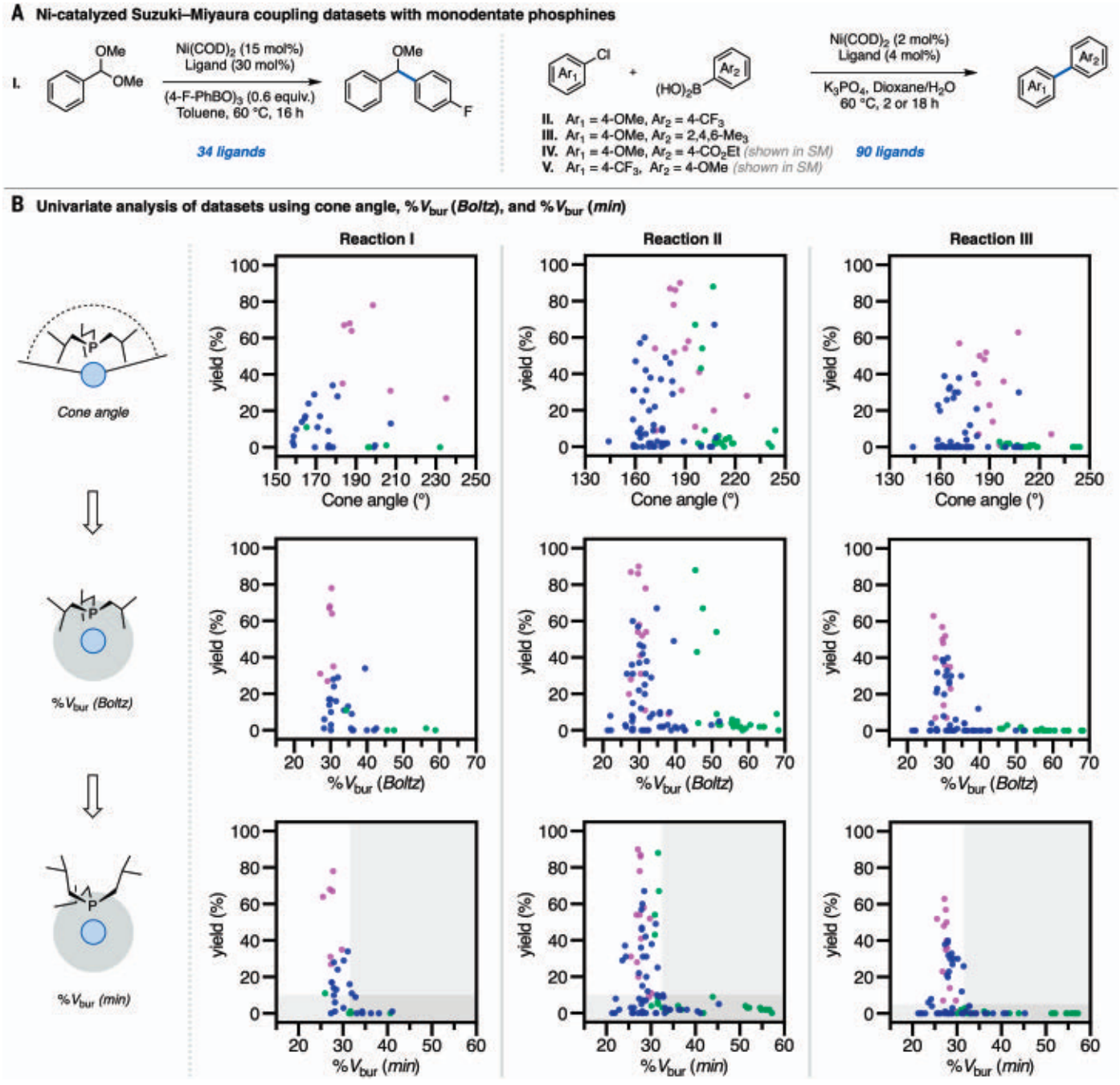



**Fig. 2. Investigation of phosphine steric parameters in Ni-catalysis datasets.** (**A**) Ni-catalyzed Suzuki-Miyaura datasets collected with monodentate phosphine ligands, including the DinoPhos ligands. Ar, aryl; Et, ethyl. (**B**) Analysis of datasets with the steric parameters cone angle, %$V_{bur}$ (Boltz), and %$V_{bur}$ (min). Pink dots represent DinoPhos-type ligands, green dots represent Buchwald-type (dialky)-*ortho*-biaryl ligands, and blue dots represent all others.

some of the largest cone angle values among ligands evaluated. The %$V_{bur}$ (Boltz) descriptor correctly grouped the DinoPhos-type ligands with other $L_2$Ni-forming phosphines; however, some flexible ligands [MeJohnPhos, P(*i*-Bu)$_3$, and PBn$_3$; Me, methyl] remained misclassified with $L_1$Ni-forming ligands. Similar to the catalytic reactions, %$V_{bur}$ (min) resolved these outliers and resulted in a sharp cutoff just below 32% between the $L_2$Ni and $L_1$Ni regimes. This value closely matched the reactivity thresholds observed in reactions I to V.

Given this finding, we hypothesized that we could predict the ligation state of new ligands, enabling a prescreening of structures that may be challenging to access synthetically. To test this, we conceptualized a new DinoPhos ligand, PteroPhos, which possesses two aryl groups with 2,4,6-triisopropylphenyl (TRIP) substituents at the 3,5 positions (Fig. 4A). Its computed cone angle is 235°, one of the largest among all monodentate phosphines studied. Despite the enormous size of this ligand, as ascertained by visual inspection, and its cone-angle value, its relatively low %$V_{bur}$ (min) of only 27.2% suggested that PteroPhos should form $L_2$Ni complexes and be effective in Ni-catalyzed SMCs. Indeed, upon preparation of this ligand, we found that it formed a $L_2$Ni(4-fluorobenzaldehyde) complex (Fig. 3B) and

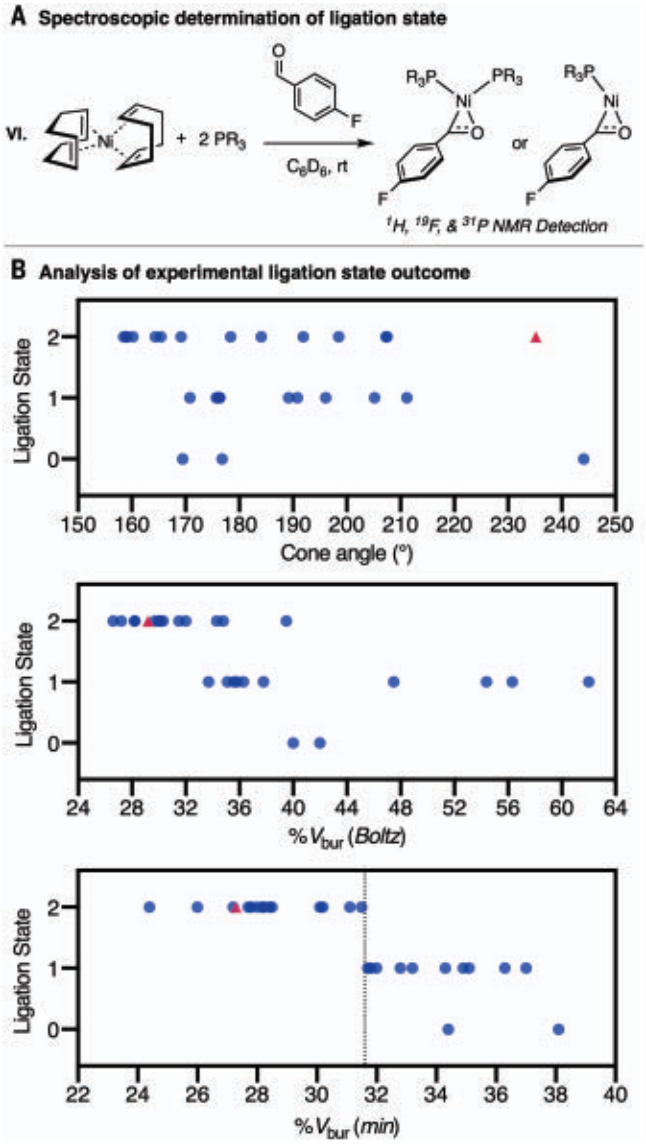



**Fig. 3. Ligation state studies.** (**A**) Reaction of $Ni(COD)_2$, 4-fluorobenzaldehyde, and two equivalents of phosphine for spectroscopic determination of ligation state using $^1H$, $^{19}F$, and $^{31}P$ NMR spectroscopy in $C_6D_6$ (25°C). R, variable substituent; rt, room temperature. (**B**) Ligation state experiments plotted against cone angle, $\%V_{bur}$ (Boltz), and $\%V_{bur}$ (min). The red triangle represents PteroPhos; blue dots represent all other phosphines studied. The (*t*-Bu)BrettPhos (nonligating) datapoint is beyond the *x*-axis range for the $\%V_{bur}$ (min) and $\%V_{bur}$ (Boltz) plots. See table S3 for conversion data for complexes. All ligands assigned as $L_1Ni$ or $L_2Ni$ had conversions >50% except for $L_1Ni$-forming $PhP(t\text{-}Bu)_2$ (39%) and $P(t\text{-}Bu)_3$ (17%). The vertical dotted line represents the location of the $\%V_{bur}$ (min) reactivity cliff.

was moderately active in Ni-catalyzed reactions I to V, consistent with the $\%V_{bur}$ (min) classification analysis.

We then sought to understand the structural importance of $\%V_{bur}$ (min) through crystallographic and computational studies (complexes **1** to **4**; Fig. 4A), to first rationalize why cone angle is not predictive of ligation state. Although our attempts at crystallizing $L_2Ni$(benzaldehyde) complexes bearing the DinoPhos ligands were not successful, we were able to obtain x-ray diffraction–quality crystals of $(DrewPhos)_2Ni$(2-naphthaldehyde) (**3**; Fig. 4A). As observed in the solid state, the two phosphines are positioned 116° cis to each other on the complex; because DrewPhos possesses a cone angle of 207°, a literal interpretation of cone angle as a description of phosphine size would suggest that there is 90° overlap of two solids, giving rise to considerable steric repulsion or distortion. Yet, when compared with the x-ray structure of $(PPh_3)_2Ni$(2-naphthaldehyde) (**1**; Fig. 4A), the two complexes exhibit very similar Ni–P bond lengths and P–Ni–P "bite angles," despite DrewPhos having a cone angle 48° larger than $PPh_3$. What cone angle does not capture is the nonuniform topology of the ligand structure distal to the metal. These less dense and more conformationally flexible regions can be arranged in a manner that accommodates a second ligand and/or substrate within the "cone," a recognized limitation that has led to the development of alternative steric descriptors (*12*, *15*, *17*, *35*). On this basis, cone angle is unsuccessful at classifying the ligation state and reactivity of the phosphines under study. Indeed, when compared with that of $L_1$ [$P(t\text{-}Bu)_3$]Ni(benzaldehyde) (**2**; Fig. 4A), the density functional theory (DFT)–optimized structure of $(PteroPhos)_2Ni$(benzaldehyde) (**4**; Fig. 4A) provides the most notable example of $L_2$ complexation being agnostic to substantial remote steric bulk.

Although this case study highlights the dependence of ligation state on steric bulk within the first coordination sphere, which is captured by the $\%V_{bur}$ descriptors, the ligands investigated in Fig. 4A had small variations between their $\%V_{bur}$ (min) and $\%V_{bur}$ (Boltz) values. To better understand the structural importance of $\%V_{bur}$ (min) in the classification of the catalytic (Fig. 2B) and ligation state experiments (Fig. 3B), the $L_2Ni$-forming ligand with the greatest difference between $\%V_{bur}$ (min) and $\%V_{bur}$ (Boltz) values, $P(i\text{-}Bu)_3$, was investigated. For this phosphine, the $\%V_{bur}$ values of the lowest-energy conformer and the Boltzmann-weighted average of all conformers (38.7 and 39.5%, respectively) are both substantially higher than those for other $L_2Ni$-forming ligands, and even higher than the $\%V_{bur}$ (Boltz) of $L_1Ni$-forming $P(t\text{-}Bu)_3$ (36.3%). However, an x-ray crystal structure of $[P(i\text{-}Bu)_3]_2Ni$(2-naphthaldehyde) (**5**; Fig. 4B) confirmed that two phosphines were bound to Ni, with $\%V_{bur}$ values determined to be 29.0 and 32.2%, nearly matching that of $\%V_{bur}$ (min) (31.1%). This analysis highlights that $\%V_{bur}$ (min) is a better representation of a phosphine when two equivalents are bound to a metal center. It suggests that the energetic benefit of coordinating a second phosphine to the metal outweighs the energetic cost of that phosphine assuming a conformation with a smaller buried volume below ~32%.

Finally, given the sharp catalytic and ligation state reactivity cliffs, we postulated that there would be a strong dependence on the thermodynamics of phosphine dissociation from $L_2Ni$(benzaldehyde) complexes with $\%V_{bur}$ (min). Out of all 28 phosphines studied in Fig. 3, only one [$Cy_2P(t\text{-}Bu)$] was found to have any spectroscopically observable equilibrium between $L_2Ni$ and $L_1Ni$ (1:4 $L_2$:$L_1$), with a $\%V_{bur}$ (min) value of 32.0% near the ligation state cutoff. Using $Cy_2P(t\text{-}Bu)$ as a reference, the free energy of ligand dissociation ($\Delta G_{dissoc}$) was calculated for 20 $L_2Ni$(benzaldehyde) complexes using DFT (Fig. 4C). [$\Delta G_{dissoc}$ was not found to be dependent on electronics in these

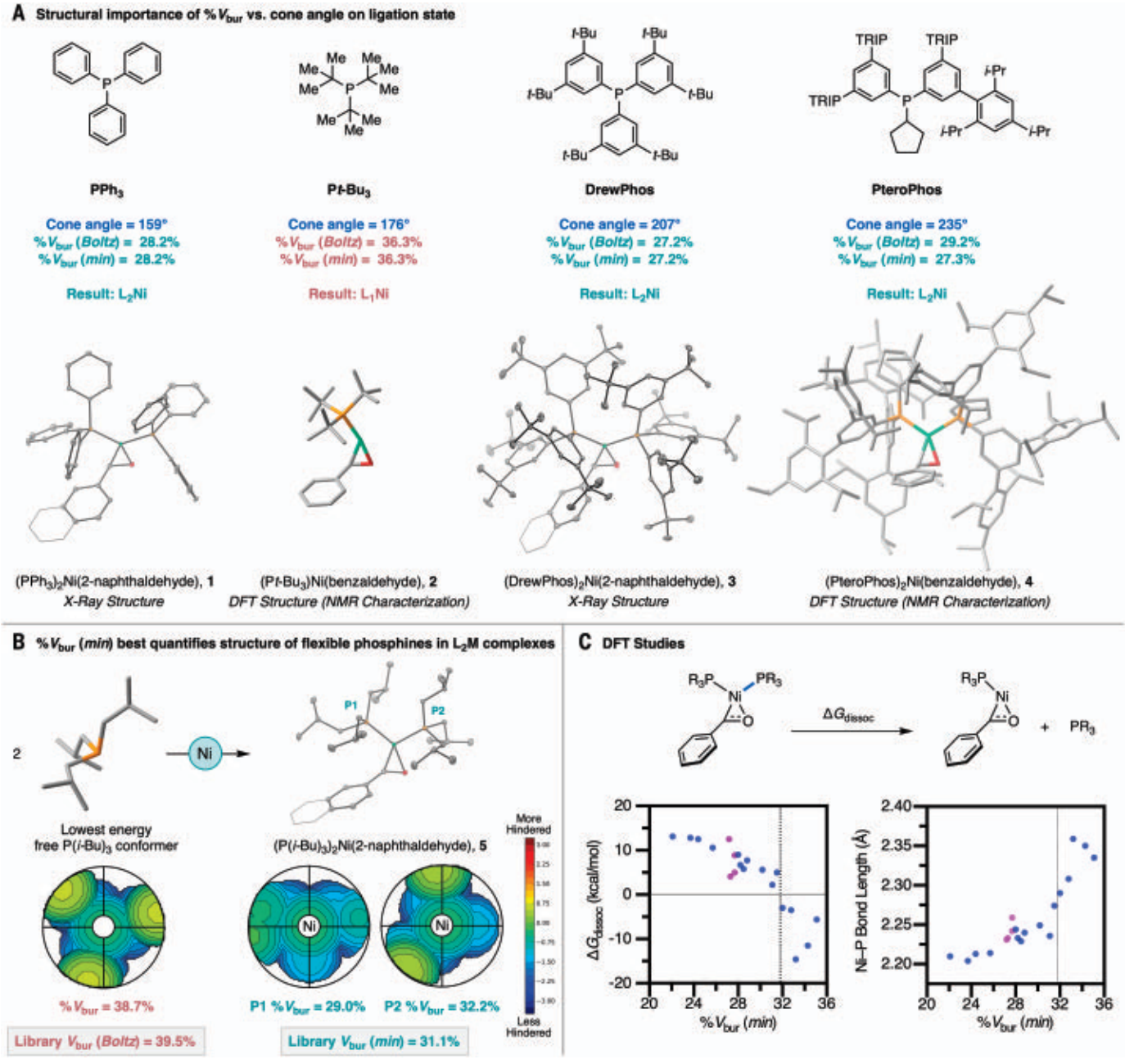



**Fig. 4. Mechanistic studies of %$V_{bur}$ (min) ligation state threshold.** (**A**) Ligands with large cone angles, but relatively small %$V_{bur}$ values, can form $L_2Ni$, supported by experimentally characterized solid-state structures [Oak Ridge thermal ellipsoid plot (ORTEP) with 30% thermal ellipsoids is shown; hydrogens are omitted for clarity] of $(PPh_3)_2Ni$(2-naphthaldehyde) and $(DrewPhos)_2Ni$(2-naphthaldehyde), along with DFT-simulated structures of [P(*t*-Bu)$_3$]Ni(benzaldehyde) and $(PteroPhos)_2Ni$(benzaldehyde) optimized at the B3LYP-D3/6-31G(d,p)[SDD] level of theory. Pr, propyl. (**B**) Formation of [P(*i*-Bu)$_3$]$_2$Ni(2-naphthaldehyde) (ORTEP is as above). The %$V_{bur}$ values of the ligands in the crystal structure are in agreement with the corresponding library value of %$V_{bur}$ (min). Steric heat maps of the phosphine ligands are shown, looking down the Ni–P bond. (**C**) DFT-calculated dissociation energies and bond lengths for $L_2Ni$(benzaldehyde) complexes. Ni–P bond length was calculated from the highlighted bond. DinoPhos-type ligands (including DrewPhos) are highlighted in pink. Calculations performed at M11-L/def2-TZVP|SMD(benzene)//B3LYP-D3/6-31G(d,p)[SDD]|SMD(benzene) level of theory.

studies; see supplementary materials (SM) for details.] A %$V_{bur}$ (min) cutoff of 32% cleanly separated the regions of positive and negative $\Delta G_{dissoc}$, with the remote steric bulk of the DinoPhos ligands and DrewPhos having minimal impacts on dissociation energy. Furthermore, the sharp downtick in the $\Delta G_{dissoc}$ values as the %$V_{bur}$ (min) of the phosphine approached 32% corresponded with a substantial uptick in the Ni–P bond length of the computed $L_2$ complex, a reflection of the increased steric pressure caused by filling Ni's first coordination sphere at this value. Together, these studies demonstrate the physical significance of the %$V_{bur}$ (min) descriptor and the 32% threshold value observed experimentally.

### Development of a threshold analysis algorithm

With the identification of sharp reactivity cliffs in Ni SMC datasets, we sought to formalize the analysis to aid in the automated discovery

of thresholds and consequent classification of ligands within a diverse array of reaction datasets. Although %$V_{bur}$ (min) was the most successful descriptor in classifying ligands in the Ni-catalyzed cross-coupling reactions, the workflow was designed to use all available descriptors and was implemented as follows: The user first defines the percent yield or selectivity for a "successful" reaction (either above "ligand-less" control reactivity or a single catalyst turnover), after which the data are parsed by a single-node decision tree algorithm

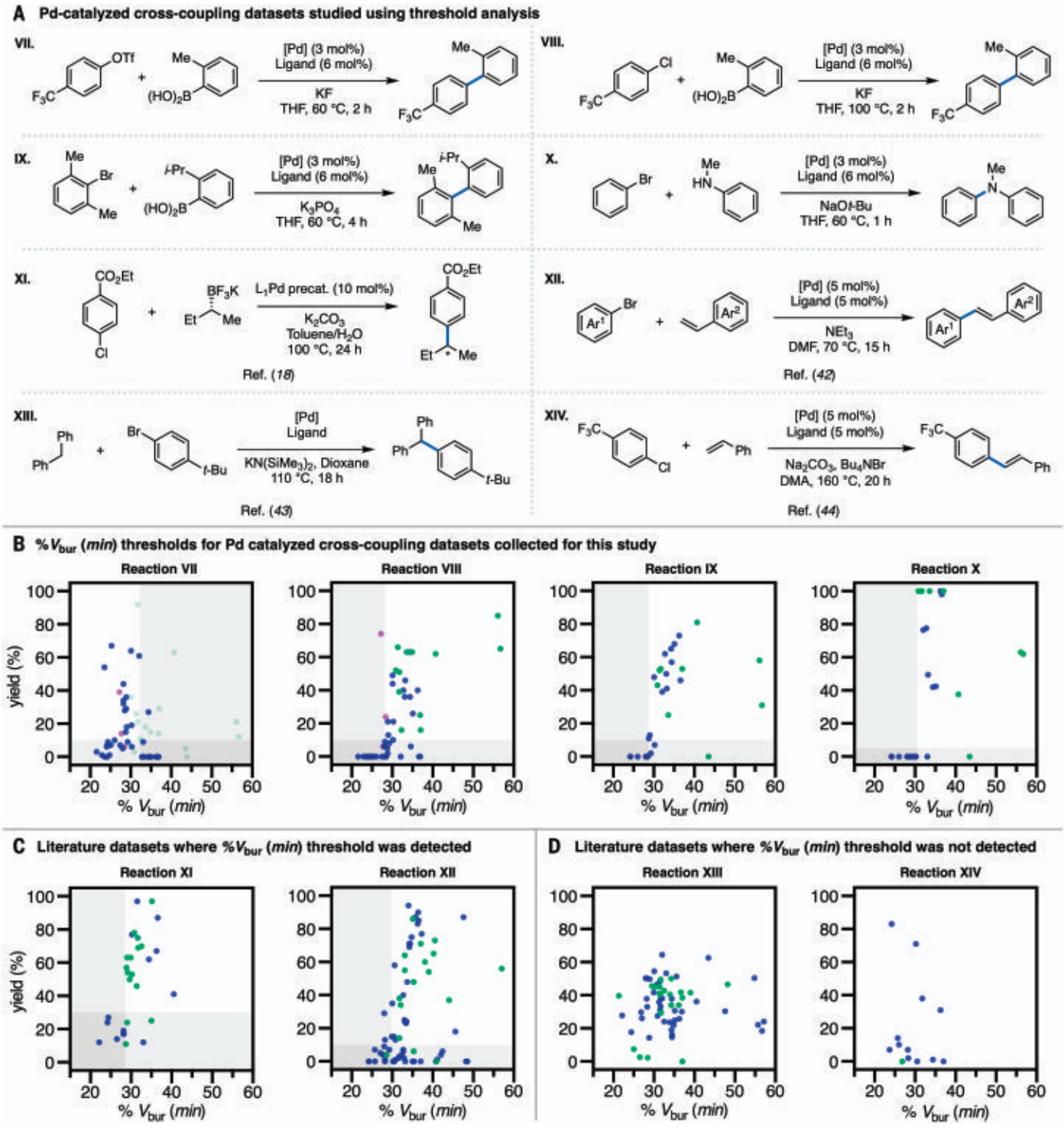



**Fig. 5. Threshold analysis for Pd-catalyzed cross-coupling reactions.** (**A**) Reaction schemes for the reactions analyzed. Reactions VII to X were collected for this study (see SM for exact reaction conditions), and reactions XI (*18*), XII (*42*), XIII (*43*), and XIV (*44*) were mined from literature sources. Ar, aromatic ring. (**B**) Threshold analysis of Pd datasets for reactions VII to X. (**C**) Threshold analysis of literature Pd datasets for reactions XI and XII. (**D**) Threshold analysis of literature Pd datasets for reactions XIII and XIV, where no %$V_{bur}$ (min) threshold is present. For all plots, pink dots represent DinoPhos-type ligands, green dots represent Buchwald-type ligands, and blue dots represent all others.

equipped with any descriptor sets provided to it. The algorithm then determines the location of the threshold (if present) for each descriptor, along with statistical measures to assess the quality of the binary classification (see SM for details). A consequence of this analysis is the dissection of datasets into four quadrants of a confusion matrix: (i) active ligands that meet the threshold ligation state criterion (true positives), (ii) inactive ligands that correctly do not meet the criterion (true negatives), (iii) active ligands that do not meet the criterion (false negatives), and (iv) inactive ligands that meet the criterion but fail for other reasons (false positives) (*36*). These groupings can then be analyzed further using statistical modeling or as targets for mechanistic interrogation (*37*).

Applying this workflow to the Ni datasets, we were able to evaluate the algorithm's ability to computationally identify and quantify the thresholds initially observed for these reactions (see SM and tables S11 and S12 for details). Most of the false positives in the classifications were ligands that were either electron-poor or contained catalyst-poisoning functionality and thus failed for reasons other than ligation state. As an example, P(4-$CF_3$Ph)$_3$ was found to form $L_2$Ni(4-fluorobenzaldehyde) complexes but was inactive in all of the catalytic reactions.

### Threshold analysis in Pd-catalyzed cross-coupling reactions

Considering the success of the classification tool for Ni-catalyzed cross-coupling reactions, we sought to determine if %$V_{bur}$ (min) could also be applied in identifying reactivity thresholds more broadly. In particular, we investigated various Pd-catalyzed cross-coupling reactions wherein the formation of $L_1$Pd versus $L_2$Pd species has been demonstrated to be a key factor in determining reaction outcome (*26*, *38*). In one case study, we probed a Pd-catalyzed SMC of aryl triflates, for which $L_2$Pd is proposed to be necessary for C–O bond oxidative addition (reaction VII; Fig. 5A) (*26*, *39*). The requirement for $L_2$Pd in catalysis also provided an opportunity to test the DinoPhos ligands for the first time in Pd catalysis. We generated a dataset using 56 phosphines and found a %$V_{bur}$ (min) reactivity threshold at 32.4% (reaction VII; Fig. 5B). The reactivity cliff mirrored those of the Ni systems in both the cutoff value and directionality of active and inactive regions. Further, DrewPhos and CyTyrannoPhos were both correctly classified as active ligands based on their %$V_{bur}$ (min) values below the threshold and performed similarly to "undecorated" $PPh_3$ and $CyPPh_2$, respectively. Spectroscopic studies corroborated these results, with the detection of $L_2$Pd(dba) by $^{31}$P NMR for DrewPhos and the DinoPhos ligands studied, including the less active CyTriceraPhos (dba, dibenzylideneacetone; figs. S97 to S101). A necessity of the analysis for reaction VII was to partition and classify Buchwald-type phosphines separately, regardless of %$V_{bur}$ (min) (see SM for details). This is likely a result of their ability to occupy two coordination sites, a design element of these ligands that allows them to mimic an $L_2$Pd species during catalysis through stabilizing Pd-arene interactions (*23*, *39*, *40*). Consistent with previous investigations by the Sigman lab into the chemoselectivity of aryl triflate SMCs (*39*), P(*o*-tolyl)$_3$ was an outlier in the reaction, despite likely favoring $L_1$Pd species under the reaction conditions (*41*).

In the next set of case studies, various reactions where $L_1$Pd species are implicated were evaluated with the classification workflow. We generated datasets for two Pd-catalyzed SMCs (reactions VIII and IX; Fig. 5A) with aryl halides, including one reaction with sterically hindered coupling partners, as well as a dataset for a Buchwald-Hartwig amination (reaction X; Fig. 5A). For the two SMCs, the classification tool revealed a threshold at about 29% $V_{bur}$ (min), with the active region occurring above that value, indicating that larger ligands—as defined by %$V_{bur}$ (min)—are required for effective catalysis (reactions VIII and IX; Fig. 5B). When comparing reactions VII and VIII**,** the opposite directionalities of the observed thresholds are consistent with the extensive literature precedent studying the effects of ligation state on the chemoselectivity of aryl triflate versus aryl chloride bond activation (*26*, *38*, *39*). Perhaps the most pronounced example is that of the Buchwald-Hartwig amination (reaction X; Fig. 5B) with a clear reactivity cliff between high-yielding reactions and those with 0% yield at and below 31% $V_{bur}$ (min). Two datasets mined from the literature were also evaluated: a stereospecific Pd-catalyzed SMC previously studied by Biscoe, Sigman, and co-workers (*18*) (reaction XI; Fig. 5A) and a Heck reaction reported by Hartwig and co-workers (*42*) (reaction XII; Fig. 5A). In both cases, %$V_{bur}$ (min) thresholds were observed and suggest that $L_1$Pd species are responsible for catalysis (reactions XI and XII; Fig. 5B).

We recognized that there are circumstances where ligand steric properties may not have as great of an impact on the reaction outcome, and thus a threshold would not be expected. In our literature-mining efforts, we identified two datasets where no phosphine steric threshold was found: a Pd-catalyzed $Csp^3$–H arylation reported by Dreher, Walsh, and co-workers (*43*) (reaction XIII; Fig. 5) and a Pd-catalyzed Heck cross-coupling studied by Zapf and Beller (*44*) (reaction XIV; Fig. 5). Possible explanations for these examples include high levels of ligandless background reactivity and/or catalytic cycles that are less sensitive to ligation state, possibly owing to the generation of catalytically active Pd nanoparticles (*45*, *46*). The ability to rapidly identify these scenarios in datasets can serve as a valuable mechanistic probe given the physical significance of %$V_{bur}$ (min) (vide supra).

### Summary and outlook

The threshold values and directionality found for each of the datasets studied in this work are summarized in Fig. 6. Comparing the Ni- and Pd-catalyzed reactions reveals several mechanistically interesting features about Ni and Pd catalysis, as well as phosphine steric effects more generally. First, the opposite threshold directionalities of the Ni- and Pd-catalyzed SMC reactions of aryl halides (reactions II to V, VIII, and IX) point to different ligation state requirements between the two metals for this reaction (*47*). Thus, the classification workflow provides a means to compare Ni and Pd head to head and elucidates the orthogonal ligand design principles often necessary for each metal (*19*, *48*).

Second, the location of the threshold for the Ni and Pd datasets occurs with some variation in value. However, for all of the reactions tested (with the exception of reaction X), an overlap area can be observed wherein ligands with %$V_{bur}$ (min) values between 29 and 32% work, independent of the directionality of the threshold. This may signify a region of %$V_{bur}$ (min) values in which $L_2$M is thermodynamically favored in the resting state but $L_1$M is also present in solution; within this region, the equilibrium between $L_2$M and $L_1$M would be affected by the temperature, solvent, and concentration of the reaction. This is known to be the case with phosphines such as $PCy_3$ (*38*), and the decrease in $\Delta G_{dissoc}$ between 29 and 32% (Fig. 4C) corroborates this hypothesis. Even in the data for Ni-catalyzed cross-couplings, the most active ligands generally fell within these limits (Fig. 2). This observation suggests that $L_1$ intermediates and/or transition states

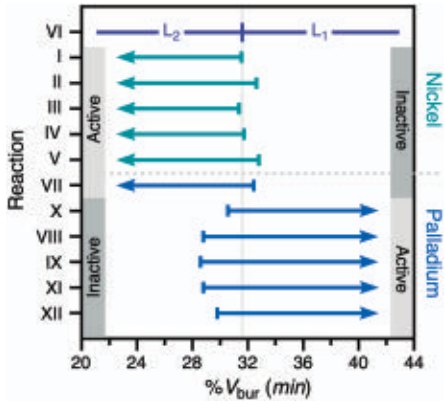



**Fig. 6. Summary of reactivity thresholds.** %$V_{bur}$ (min) reactivity threshold locations and directionalities for all Ni-catalyzed [reactions I to V (green)] and Pd-catalyzed [reactions VII to XII (blue)] reactions under study (excluding reactions XIII and VIV, where no threshold was found). The $L_1$ and $L_2$ regions defined by the spectroscopic ligation state studies are shown [reaction VI (purple)].

may be relevant in Ni catalysis (*49*, *50*) but that the ability to attain an $L_2$ resting state is necessary (*32*, *51*, *52*), as evidenced by the sharp reactivity cliffs.

We have developed a strategy for the binary classification of monodentate phosphine ligation state and reactivity in cross-coupling catalysis. In searching the feature space of structurally diverse monodentate phosphines, we identified $\%V_{bur}$ (min) as the descriptor capable of bifurcating datasets on the basis of catalyst ligation state. We envision that this tool should facilitate mechanistic studies of related organometallic reactions and enable reaction development through prediction of active and inactive and mono- and bis-ligating phosphines before synthesis. Though we recognize that $\%V_{bur}$ (min) will not capture reactivity trends across all phosphines, the ability to identify outliers (especially false negatives) in the analysis can motivate the development of new descriptors and targeted mechanistic study. Taken together, this study highlights how classification analysis can serve as an important mechanistic and predictive tool to facilitate the understanding of structure-reactivity relationships in catalysis.

## REFERENCES AND NOTES


1. E. V. Anslyn, D. A. Dougherty, *Modern Physical Organic Chemistry* (Univ. Science Books, 2006).
2. C. B. Santiago, J. Y. Guo, M. S. Sigman, *Chem. Sci.* **9**, 2398–2412 (2018).
3. L. P. Hammett, *J. Am. Chem. Soc.* **59**, 96–103 (1937).
4. P. R. Wells, *Chem. Rev.* **63**, 171–219 (1963).
5. N. A. Romero, D. A. Nicewicz, *Chem. Rev.* **116**, 10075–10166 (2016).
6. D. G. Brown, J. Boström, *J. Med. Chem.* **59**, 4443–4458 (2016).
7. S. Xu, E. H. Kim, A. Wei, E. I. Negishi, *Sci. Technol. Adv. Mater.* **15**, 044201 (2014).
8. J. F. Hartwig, *Organotransition Metal Chemistry: From Bonding to Catalysis* (Univ. Science Books, 2010).
9. C. C. C. Johansson Seechurn, M. O. Kitching, T. J. Colacot, V. Snieckus, *Angew. Chem. Int. Ed.* **51**, 5062–5085 (2012).
10. D. J. Durand, N. Fey, *Chem. Rev.* **119**, 6561–6594 (2019).
11. C. A. Tolman, *J. Am. Chem. Soc.* **92**, 2956–2965 (1970).
12. C. A. Tolman, *Chem. Rev.* **77**, 313–348 (1977).
13. D. White, B. C. Tavener, P. G. L. Leach, N. J. Coville, *J. Organomet. Chem.* **478**, 205–211 (1994).
14. K. C. Harper, E. N. Bess, M. S. Sigman, *Nat. Chem.* **4**, 366–374 (2012).
15. A. Poater *et al.*, *Eur. J. Inorg. Chem.* **2009**, 1759–1766 (2009).
16. H. Clavier, S. P. Nolan, *Chem. Commun.* **46**, 841–861 (2010).
17. L. Falivene *et al.*, *Nat. Chem.* **11**, 872–879 (2019).
18. S. Zhao *et al.*, *Science* **362**, 670–674 (2018).
19. K. Wu, A. G. Doyle, *Nat. Chem.* **9**, 779–784 (2017).
20. H. Y. Liu, K. Eriks, A. Prock, W. P. Giering, *Organometallics* **9**, 1758–1766 (1990).
21. T. Gensch *et al.*, ChemRxiv 12996665 [Preprint] (2021); http://doi.org/10.26434/chemrxiv.12996665.v1.
22. F. Barrios-Landeros, B. P. Carrow, J. F. Hartwig, *J. Am. Chem. Soc.* **131**, 8141–8154 (2009).
23. R. Martin, S. L. Buchwald, *Acc. Chem. Res.* **41**, 1461–1473 (2008).
24. A. A. Thomas, A. F. Zahrt, C. P. Delaney, S. E. Denmark, *J. Am. Chem. Soc.* **140**, 4401–4416 (2018).
25. G. C. Fu, *Acc. Chem. Res.* **41**, 1555–1564 (2008).
26. F. Schoenebeck, K. N. Houk, *J. Am. Chem. Soc.* **132**, 2496–2497 (2010).
27. T. Gensch *et al.*, ChemRxiv chemrxiv-2021-fgm7v [Preprint] (2021); http://doi.org/10.33774/chemrxiv-2021-fgm7v.
28. A. P. Cinderella, B. Vulovic, D. A. Watson, *J. Am. Chem. Soc.* **139**, 7741–7744 (2017).
29. Cone angle values used in the text are those for the Boltzmann weighted average of the conformational ensemble from the descriptor library. Minimum cone angles were also evaluated but for simplicity are shown in fig. S2.
30. A. J. Kendall, L. N. Zakharov, D. R. Tyler, *Inorg. Chem.* **55**, 3079–3090 (2016).
31. Buchwald-type phosphines are known to act as structurally bidentate-like ligands with Pd owing to stabilizing interactions between the arene and the metal. Although little is known as of now about the structural behavior of these ligands with Ni, this semibidentate binding may enable certain Ni-catalyzed reactions to occur despite only one phosphine bound; their mode of action in these case studies is the subject of ongoing investigation.
32. E. A. Standley, S. J. Smith, P. Müller, T. F. Jamison, *Organometallics* **33**, 2012–2018 (2014).
33. D. Walther, *J. Organomet. Chem.* **190**, 393–401 (1980).
34. M. Ohashi, H. Saijo, T. Arai, S. Ogoshi, *Organometallics* **29**, 6534–6540 (2010).
35. I. A. Guzei, M. Wendt, *Dalton Trans.* **2006**, 3991–3999 (2006).
36. Given that other ligand features can and will lead to poor catalyst performance, the presence of false positives should not be viewed as detrimental to the success and/or accuracy of the classifier.
37. Z. L. Niemeyer *et al.*, *J. Am. Chem. Soc.* **139**, 12943–12946 (2017).
38. A. F. Littke, C. Dai, G. C. Fu, *J. Am. Chem. Soc.* **122**, 4020–4028 (2000).
39. Z. L. Niemeyer, A. Milo, D. P. Hickey, M. S. Sigman, *Nat. Chem.* **8**, 610–617 (2016).
40. K. L. Billingsley, K. W. Anderson, S. L. Buchwald, *Angew. Chem. Int. Ed.* **45**, 3484–3488 (2006).
41. J. F. Hartwig, F. Paul, *J. Am. Chem. Soc.* **117**, 5373–5374 (1995).
42. J. P. Stambuli, S. R. Stauffer, K. H. Shaughnessy, J. F. Hartwig, *J. Am. Chem. Soc.* **123**, 2677–2678 (2001).
43. J. Zhang, A. Bellomo, A. D. Creamer, S. D. Dreher, P. J. Walsh, *J. Am. Chem. Soc.* **134**, 13765–13772 (2012).
44. A. Zapf, M. Beller, *Chemistry* **7**, 2908–2915 (2001).
45. S. S. Zalesskiy, V. P. Ananikov, *Organometallics* **31**, 2302–2309 (2012).
46. N. W. J. Scott *et al.*, *Chem. Sci.* **10**, 7898–7906 (2019).
47. S. Z. Tasker, E. A. Standley, T. F. Jamison, *Nature* **509**, 299–309 (2014).
48. C. M. Lavoie, M. Stradiotto, *ACS Catal.* **8**, 7228–7250 (2018).
49. H. Shirataki, M. Ohashi, S. Ogoshi, *Eur. J. Org. Chem.* **2019**, 1883–1887 (2019).
50. We recognize that there are some instances where Ni will be $L_1$ throughout the catalytic cycle, allowing ligands like $P(t\text{-Bu})_3$ to be used successfully; see (*49*).
51. N. Hazari, P. R. Melvin, M. M. Beromi, *Nat. Rev. Chem.* **1**, 0025 (2017).
52. P.-A. Payard, L. A. Perego, I. Ciofini, L. Grimaud, *ACS Catal.* **8**, 4812–4823 (2018).
53. E. Peters, SigmanGroup/threshold: Single parameter threshold (1.0.0). Zenodo (2021); http://doi.org/10.5281/zenodo.5227162.


## ACKNOWLEDGMENTS

We thank P. Jeffrey for assistance with x-ray structure determination and K. Wu for helpful discussions and directions. Additionally, we thank F. D. Toste for helpful discussions related to the preparation of the manuscript. **Funding:** NSF CCI Center for Computer Assisted Synthesis (CHE-1925607). **Author contributions:** T.G., M.S.S., and A.G.D. conceived the project. S.H.N.-S., S.R.S., J.E.B., and H.C.J. performed the experiments and analyzed the data. E.P., T.G., and S.R.S. developed the classification algorithm. M.S.S. and A.G.D. oversaw the project. S.H.N.-S., J.E.B., S.R.S., M.S.S., and A.G.D. wrote the manuscript with contributions from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** X-ray crystallographic data are available free of charge from the Cambridge Crystallographic Data Centre, under reference numbers 2069892 to 2069895. Coordinates of all computed structures are included in a separate zip file. All other experimental, computational, spectroscopic, and crystallographic data are included in the supplementary materials. The code used in this project is available at https://github.com/sigmangroup/threshold (*53*).

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj4213
Materials and Methods
Figs. S1 to S121
Tables S1 to S30
References (*54*–*95*)
Data S1

11 May 2021; accepted 2 September 2021
10.1126/science.abj4213

ZEOLITE SYNTHESIS

# A priori control of zeolite phase competition and intergrowth with high-throughput simulations

**Daniel Schwalbe-Koda[1], Soonhyoung Kwon[2], Cecilia Paris[3], Estefania Bello-Jurado[3], Zach Jensen[1], Elsa Olivetti[1], Tom Willhammar[4], Avelino Corma[3], Yuriy Román-Leshkov[2], Manuel Moliner[3]*, Rafael Gómez-Bombarelli[1]***

**Zeolites are versatile catalysts and molecular sieves with large topological diversity, but managing phase competition in zeolite synthesis is an empirical, labor-intensive task. In this work, we controlled phase selectivity in templated zeolite synthesis from first principles by combining high-throughput atomistic simulations, literature mining, human-computer interaction, synthesis, and characterization. Proposed binding metrics distilled from more than 586,000 zeolite-molecule simulations reproduced the extracted literature and rationalized framework competition in the design of organic structure-directing agents. Energetic, geometric, and electrostatic descriptors of template molecules were found to regulate synthetic accessibility windows and aluminum distributions in pure-phase zeolites. Furthermore, these parameters allowed us to realize an intergrowth zeolite through a single bi-selective template. The computation-first approach enables control of both zeolite synthesis and structure composition using a priori theoretical descriptors.**

Zeolites are nanoporous materials with a wide range of applications in industrial and sustainable catalysis and separations (*1*, *2*). Although topological diversity makes zeolites versatile, this diversity stems from phase competition between metastable polymorphs, which hinders the rational design of new synthesis routes for zeolites (*3*–*5*). Computer simulations can guide experimental work by predicting the affinity between organic structure-directing agents (OSDAs) and targeted topologies (*6*–*8*). Recent studies have

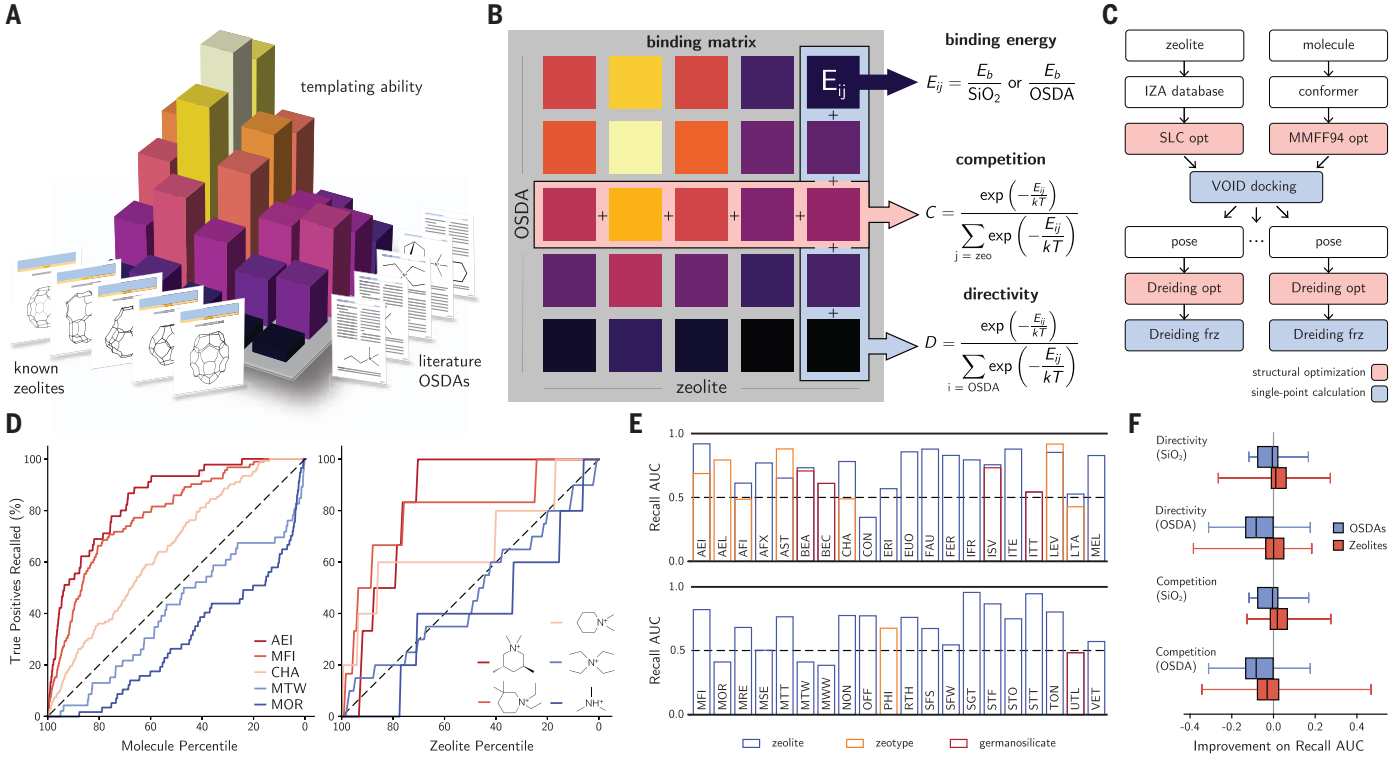



**Fig. 1. Computational method used to quantify phase competition in zeolites.** (**A**) OSDAs reported in the literature are docked into known zeolites, spanning a full matrix of binding energies. (**B**) Each element of the binding matrix is defined as the molecule-framework binding energy at the most favorable loading, normalized by the number of framework atoms or OSDA molecules (materials and methods). Along OSDAs, binding energies are ranked to determine how directive a molecule is toward a zeolite. Along zeolites, phase competition is quantified for a given OSDA. (**C**) Computational pipeline used to calculate binding energies for zeolite-OSDA pairs. (**D**) Explaining the literature with the templating energy. A recall curve is obtained by using the metric as a sorting algorithm for ranking OSDAs for a given zeolite (left) or zeolites for a given molecule (right). The dashed line indicates the expected recall curve for a random sorting algorithm. (**E**) Recall AUC for several zeolite frameworks according to their typical compositions. (**F**) Improvement of the recall AUC of OSDAs and zeolites using the proposed templating energy compared with other individual metrics. Including phase competition analysis improves the explanation of literature outcomes for ~70% of the zeolites in the literature.

adopted this strategy by simulating up to thousands of OSDAs for a few frameworks at a time (*8*–*10*), but they have been unable to predict whether the proposed OSDAs are more favorable toward another zeolite rather than the desired ones. Accordingly, relatively few experimental realizations have been reported from such studies, because computer-designed OSDA candidates often fail to crystallize the targeted structures. Additionally, design algorithms usually lead to molecules with low synthetic accessibility (*6*, *11*). Manual literature analysis can inform heuristics for zeolite synthesis and avoid expensive simulations (*12*), but datasets compiled from tens of papers are limited in their ability to explain phase selectivity in combinatorial host-guest pairings and have only been attempted for template-free synthesis routes. Finally, although several databases of experimental and hypothetical zeolite structures exist (*13*–*15*), few datasets containing OSDAs or OSDA-zeolite affinities are publicly available. A general approach to simultaneously rationalize phase competition of zeolites, retroactively validate synthesis routes from the literature, control the trade-off between OSDA selectivity and synthetic complexity, and tailor the heteroatom distribution of the targeted materials has not yet been realized. In this work, we simulated over half a million zeolite-OSDA pairs, proposed design principles that outperformed traditional binding energy metrics in reproducing synthesis outcomes from more than one thousand papers, and demonstrated the phase-selective synthesis of targeted zeolites from newly identified OSDAs. The computational approach allowed the synthesis of SSZ-39 (AEI) and SSZ-13 (CHA) zeolites under a wide range of conditions, relating broader synthetic accessibility windows to binding and geometric descriptors of templates. Furthermore, fine-tuning the charge distribution of OSDAs modulated the aluminum pairing in zeolites, as demonstrated for the CHA framework. Finally, the design toolkit also enabled us to synthesize an aluminosilicate CHA/AEI intergrowth from a single OSDA, showcasing opportunities for tuning catalytic properties by controlling zeolite phase competition. The integrated platform from this work is made available for the public (see the Data and materials availability statement) and is expected to markedly accelerate the design and optimization of zeolites.

## Phase competition and literature data

To capture phase competition with atomistic simulations, we postulated that a selective OSDA must exhibit both strong binding affinity toward the desired host and weak binding affinity toward all other frameworks (*16*, *17*). To retroactively validate this hypothesis, we obtained 549 OSDAs from the literature using

[1]Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [2]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]Instituto de Tecnología Química, Universitat Politècnica de València-Consejo Superior de Investigaciones Científicas, 46022 Valencia, Spain. [4]Department of Materials and Environmental Chemistry, Stockholm University, SE-106 91 Stockholm, Sweden.
***Corresponding author. Email: mmoliner@itq.upv.es (M.M.); rafagb@mit.edu (R.G.-B.)**

automated extraction tools (*18*, *19*). Then, we calculated the binding affinity of each OSDA toward 209 zeolite frameworks (fig. S1), generating a binding matrix (materials and methods and Fig. 1, A and B). More than 586,000 zeolite-OSDA poses across different frameworks, OSDAs, loadings, and initial conformations were generated using the computational pipeline shown in Fig. 1C (*20*, *21*). By selecting the pose that optimized the binding energy for each OSDA-zeolite pair, we obtained a binding matrix with ~112,400 entries. Phase selectivity was then quantified by comparing OSDAs and zeolites across the rows and columns of the binding matrix using two metrics: the directivity of an OSDA ($D$), or how close a molecule is to the best OSDA for a given framework, and the competitivity of a framework ($C$), or how close a zeolite is to the best host for the given OSDA (materials and methods and Fig. 1B). These quantities did not necessarily correlate with the conventional binding energies (fig. S2). As both the directivity and competitivity metrics were useful for optimizing zeolite-OSDA pairs, it was convenient to unify them under a single one-dimensional descriptor. We defined a quantity called templating energy ($E_T$) from the geometric mean of the Boltzmann average of each energy metric

$$E_T = -kT \log(C_{OSDA} D_{OSDA} C_{SiO2} D_{SiO2})^{1/4}$$

where $T$ is a temperature, $k$ is the Boltzmann constant, and $C$ and $D$ are the metrics of competition and directivity, respectively, normalized by either OSDA or $SiO_2$ (materials and methods). The templating energy quantifies the notion that a particular zeolite-OSDA pair is better over another if both the competition and directivity energies are favorable according to both normalizations (fig. S3).

The simulation data were then benchmarked on their ability to reproduce the 1122 zeolite-OSDA pairs extracted from the literature. We used the templating energy to rank the best OSDAs for a given zeolite and the best zeolites for a given OSDA. Because pairs not observed in the literature did not represent true negative data, only positive data points were used to assess the ranking scheme. Literature recall was quantified by plotting the cumulative number of positive OSDA-zeolite pairs recalled against the molecule/zeolite percentile (fig. S4). A perfect agreement between the binding metric and the literature data would lead to an area under the recall curve (AUC) equal to 1, and a random metric would have an AUC of 0.5. Figure 1D exemplifies the recall curve for some zeolites and OSDAs. For zeolites such as AEI or ZSM-5 (MFI), the AUC values were high, which suggests that the positive pairs were systematically ranked high based on the templating energy metric. In fact, 34 of the 40 most common zeolite frameworks in the literature display AUC values higher than 0.5 (Fig. 1E), which demonstrates that the ranking scheme reproduced past synthesis outcomes. Only OSDAs for a few frameworks, such as ZSM-12 (MTW) and mordenite (MOR), were not accurately recalled by the ranking scheme. These hosts are often synthesized through OSDA-free routes or with small amines and thus are expected to appear as products in the presence of unselective OSDAs, with their synthesis dominated by factors beyond host-guest interactions, such as gel composition or temperature. Analogously to zeolites, the recall curves of OSDAs demonstrated that certain molecules were more selective than others. Smaller OSDAs typically have lower selectivity toward large cages and pores, and more complex, highly decorated OSDAs tend to be more selective (*22*). In keeping, zeotype structures have recall AUC on average 10% lower than that of their zeolite counterparts (Fig. 1E), as AlPO-type structures are frequently synthesized from smaller, less selective OSDAs (*22*). Additionally, a lower agreement between the literature and the binding metrics for certain frameworks suggested that binding energy metrics alone were limited predictors of synthesis outcomes in the context of Al-rich zeolites or dual-OSDA design. This is the case for zeolite A (LTA), for instance, which is typically synthesized using tetramethylammonium to selectively direct the formation of sod cages in combination with another OSDA to direct the formation of the lta cage (*23*), or CIT-1 (CON), whose intersecting pores allow OSDAs to create molecular aggregates that direct crystallization (*24*).

Finally, we demonstrated that the templating energy metric outperformed plain binding energies in recalling positive data points from the literature (Fig. 1F). When selecting molecules for a given zeolite, the templating energy explained the literature better for ~70% of the structures, for which the average improvement in AUC was 0.06 with respect to the binding energy per framework atom, which is the traditional metric of OSDA-zeolite affinity (fig. S5). On the other hand, the best hosts for a given OSDA were not well predicted by the templating energy, as the directivity

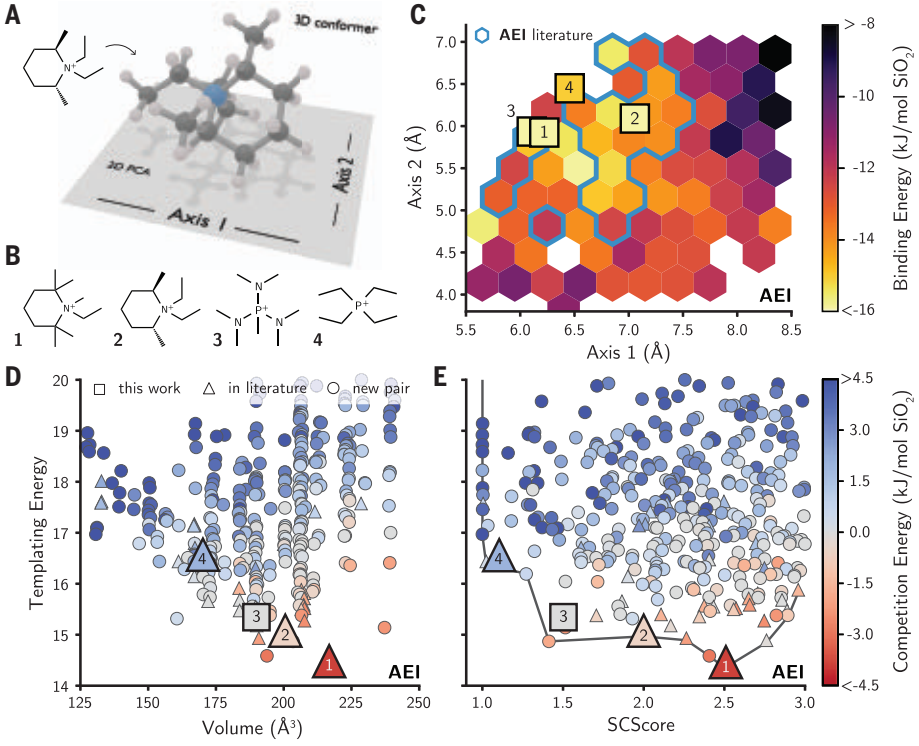



**Fig. 2. OSDA selectivity for AEI zeolite.** (**A**) Schematic on the shape representation of OSDAs. The three-dimensional conformation is converted into two axes by projecting the atomic coordinates into a two-dimensional space using PCA. (**B**) OSDA candidates for the synthesis of AEI. (**C**) Relationship between the shapes of OSDAs and their binding energies toward AEI. The color of each hexagon indicates the mean competition energy for all OSDAs within that area. Blue outlines indicate that at least one OSDA within that region is known to synthesize AEI with zeolite composition. (**D**) Templating energy of OSDAs against their volume. (**E**) Templating energy of OSDAs against their SCScore. The Pareto frontier is shown with a solid gray line.

component did not influence the structure-directing ability of an OSDA. Thus, we propose both a single aggregate parameter to design OSDAs that is effective for screening and independent energy metrics from which further insights could be derived.

### OSDA design through phase competition and shape analysis

OSDAs must satisfy other design targets beyond binding energies. Physical descriptors of the OSDA molecular structure, for example, are useful predictors of templating ability and complement phase competition metrics (*25*). Often, the strongest host-guest interactions occur for well-defined OSDA volumes or shapes, but their interplay is not necessarily bijective. To describe the shape of OSDAs with low-dimensional parameters, we performed a principal components analysis (PCA) of the three-dimensional distribution of atomic coordinates of ground-state conformers into a two-dimensional space (materials and methods and Fig. 2A). Additionally, we calculated the volume of each OSDA by using a voxel-based approach on the same conformers (materials and methods). The two PCA axes and the volume are useful parameters to describe the overall shape and size of a molecule and can be used as filters when proposing new OSDAs.

Designing molecules with targeted shapes, sizes, and selectivity can be performed in a purely computational manner, but other performance metrics of OSDAs, such as synthetic accessibility, are hard to evaluate algorithmically. By contrast, desiderata such as the intellectual novelty of molecules are more easily recognized by experts, whereas quantitative templating metrics are not. A human-computer partnership drastically accelerates the process of molecular design by combining chemical intuition with computationally selected leads (*26*, *27*). To navigate and visualize hundreds of thousands of zeolite-OSDA pairs and leverage shape similarity, chemical intuition, and domain expertise, we created an online database of zeolite-OSDA pairs containing all simulation outcomes of this work. This database, called Organic Structure-directing agent DataBase (OSDB), provides a graphical interface to query, select, and compare molecules as well as obtain information about prior art in the extracted literature and patents (fig. S6 and movie S1). Thus, all relevant design metrics were easily filtered and visualized in OSDB, allowing a computer-augmented design of OSDAs for targeted zeolites using interactive down-selection of candidates (movie S1). We illustrate this design process with several examples.

On the basis of the AUC analysis, our proposed metric showed high explanatory power for the small-pore zeolite AEI. The synthesis of the aluminosilicate SSZ-39 with AEI framework is typically carried out using alkyl-substituted piperidinium cations as OSDAs, such as the computationally designed *N*-ethyl-*N*-methyl-2,2,6,6-tetramethylpiperidinium (OSDA **1**) (*28*), or the commercial *N*,*N*-diethyl-2,6-dimethylpiperidinium (OSDA **2**) (*29*) (Fig. 2B). According to our simulations, both OSDAs had favorable shapes for the AEI zeolite, with the nearly 1:1 aspect ratio at which binding affinities toward this framework were stronger (Fig. 2C). In fact, most OSDAs that are able to synthesize AEI with zeolite composition appeared at the top left region of the shape space, suggesting that shape and templating ability were correlated for this zeolite (Fig. 2C). Furthermore, OSDAs **1** and **2** exhibit excellent templating energy and volume to fill the aei cage (200 to 225 Å$^3$) (Fig. 2D). Nevertheless, their synthetic complexity, quantified by the SCScore proposed by Coley *et al.* (*30*), was higher than that of many other known or candidate OSDAs (Fig. 2E),

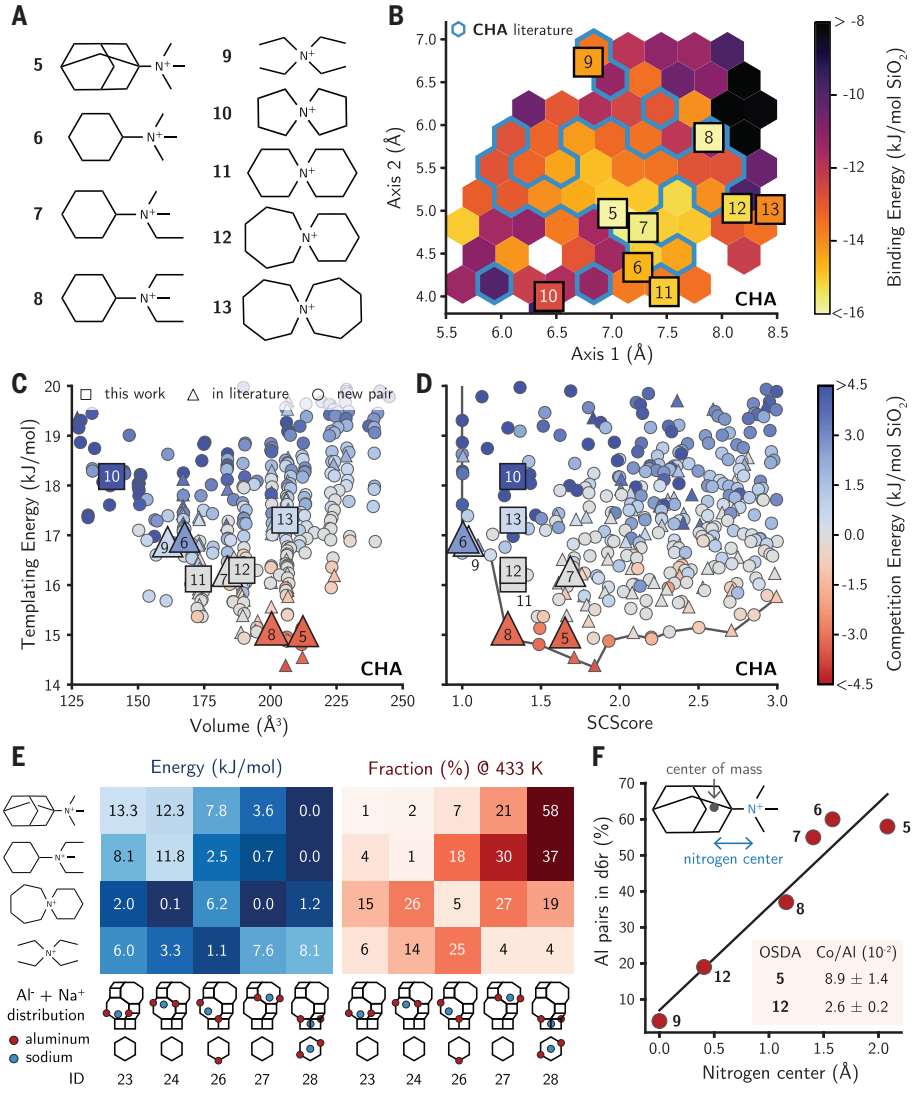



**Fig. 3. OSDA selectivity for CHA zeolite.** (**A**) OSDA candidates for the synthesis of CHA. (**B**) Relationship between the shapes of OSDAs and their binding energies toward CHA. The color of each hexagon indicates the mean competition energy for all OSDAs within that area. Blue outlines indicate that at least one OSDA within that region is known to synthesize CHA with zeolite composition. (**C**) Templating energy of OSDAs against their volume. (**D**) Templating energy of OSDAs against their SCScore. The Pareto frontier is shown with a solid gray line. (**E**) Influence of OSDA in the energy (blue) and fraction (red) of aluminum pairs in the cha cage in the presence of sodium. The predicted fraction of aluminum pairs is calculated at 433 K. (**F**) Relationship between the nitrogen center and the fraction of aluminum pairs in d6r units (distribution 28) for different OSDAs. Molecules with larger separation between the center of mass and the nitrogen center favor the pairing of aluminum atoms in the d6r units. The solid line is a linear regression of the data. (Inset) Co/Al ratio from cobalt titration experiments for CHA samples synthesized with OSDAs **5** and **12**.

suggesting higher preparation costs. In particular, the lower SCScore of OSDA **2** compared with OSDA **1** could be related to its usage as a commercial template for the industrial preparation of SSZ-39.

On the basis of the design strategies summarized by Fig. 2 and visualized in OSDB (movie S1), we proposed tris(dimethylamino)(methyl) phosphonium (OSDA **3**) as a candidate for the synthesis of SSZ-39. It not only displayed favorable volume (Fig. 2D) and phase competition metrics (fig. S7) toward the AEI framework, but also showed considerably lower synthetic complexity than piperidinium-type cations (Fig. 2E). Following this lead, SSZ-39 was first prepared using OSDA **3** under the following synthesis conditions: 1 $SiO_2$: 0.036 $Al_2O_3$: 0.3 OSDA **3**: 0.2 NaOH: 15 $H_2O$, using the high-silica FAU CBV720 as the Si/Al source, and with the crystallization carried out at 135°C for 7 days. The main physicochemical characteristics of this material are summarized in fig. S8, showcasing tetrahedrally coordinated Al species in framework positions as well as crystalline nature and textural properties comparable to an AEI zeolite synthesized with baseline commercial OSDAs.

We evaluated the selectivity of computationally designed OSDA **3** by comparing its ability to crystallize AEI against a known, also phosphorous-containing template for this framework, tetraethylphosphonium (OSDA **4**). Instead of using the high-silica FAU zeolite (Zeolyst, CBV720) as the Si and Al source, we used fumed silica as the Si source and the low-silica FAU (CBV500) as the Al source, thereby lowering the overall cost of the synthesis but increasing the difficulty in crystallizing the target zeolite. The synthesis of SSZ-39 was attempted under the following conditions: 1 $SiO_2$: 0.091 or 0.125 $Al_2O_3$: 0.2 OSDA **3** or **4**: 0.25 NaOH: 5 $H_2O$. Although OSDA **4** did not yield AEI in either of these conditions, OSDA **3** successfully crystallized SSZ-39 in both (fig. S9), attesting to the higher selectivity of the latter toward AEI. Although both OSDAs display similar shapes, OSDA **3** is closer to the ideal OSDA **1** than OSDA **4** (Fig. 2B), exhibiting more favorable templating and competition energies by ~1 kJ/mol and 5 kJ/mol $SiO_2$, respectively. These values suggested that the proposed computational metrics were effective in capturing the subtleties of template selectivity during the OSDA design process despite the fine resolution of the zeolite-OSDA interaction energies. Our binding metrics were also effective at singling out highly selective OSDAs that led to the discovery of frameworks such as MFI, ITQ-4 (IFR), and ITQ-7 (ISV) (fig. S10). These results suggested that the proposed tools could not only identify new OSDAs for known structures, but that they hold potential to detect archetypical templates for unrealized zeolite topologies.

As a second example, we analyzed the interplay between OSDA shape and binding metrics

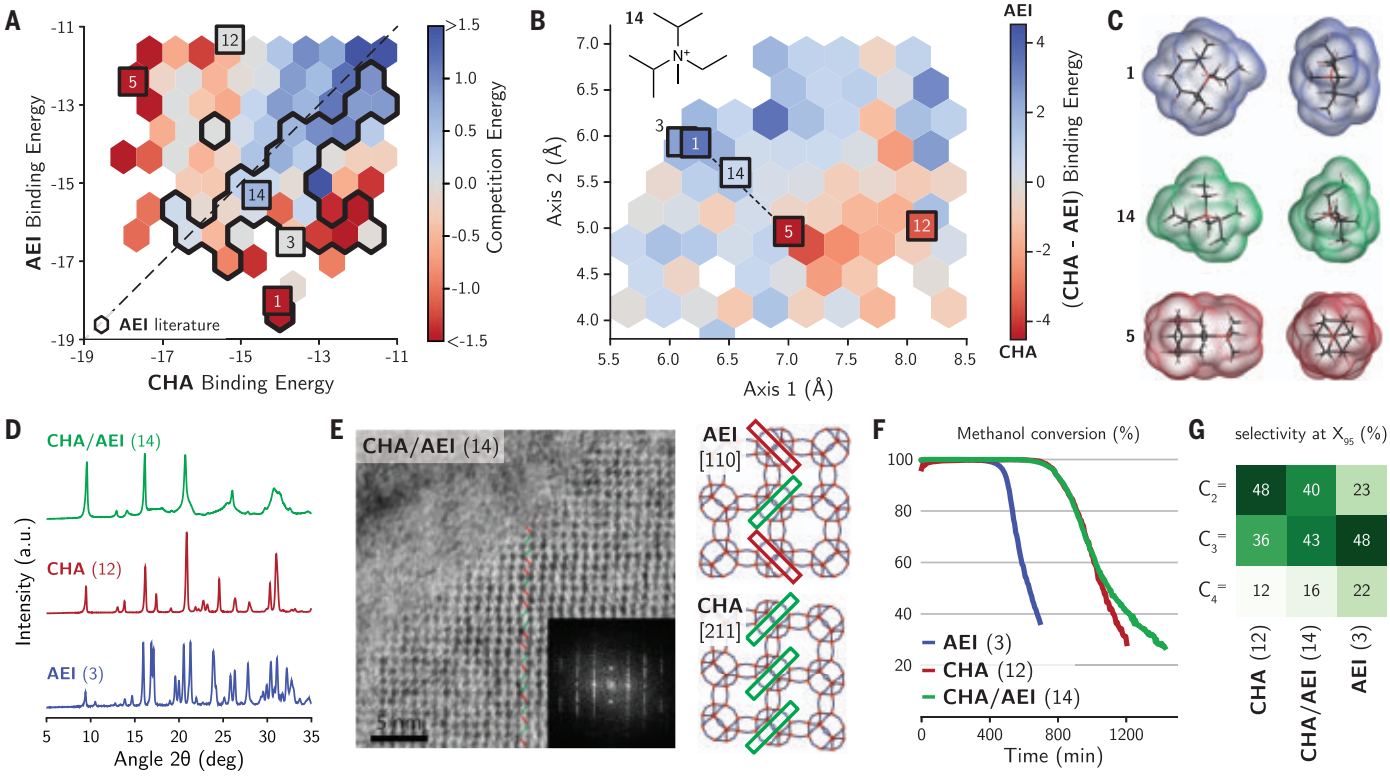



**Fig. 4. CHA/AEI intergrowth zeolite.** (**A**) Comparison between binding energies of OSDAs in CHA and AEI zeolites. The color of each hexagon indicates the mean competition energy for all OSDAs within that area. All energies are given in kilojoules per mole $SiO_2$. Black outlines indicate that at least one OSDA within that region is known to synthesize AEI with zeolite composition. (**B**) Relationship between the shapes of OSDAs and their binding energies toward CHA or AEI. Colors shifted toward red or blue indicate that the OSDA has a shape that favors the synthesis of CHA or AEI, respectively. OSDA **14**, the one that enables the CHA/AEI intergrowth, lies midway between the best OSDA for AEI (OSDA **1**) and the best OSDA for CHA (OSDA **5**). Energy difference is given in kilojoules per mole $SiO_2$. (**C**) DFT-calculated charge densities for OSDAs **1**, **5**, and **14**. OSDA **14** is between OSDAs **1** and **5** in terms of shape. (**D**) PXRD patterns of the as-prepared materials obtained with OSDAs **3** (AEI), **12** (CHA), and **14** (CHA/AEI). a.u., arbitrary units. (**E**) Integrated differential phase contrast–scanning transmission electron microscopy (iDPC-STEM) image of CHA/AEI as synthesized by OSDA **14** acquired from a sample sectioned using ultramicrotomy, and structural models for AEI and CHA showing the stacking patterns seen in the STEM image. (Inset) Fourier transform of the STEM image. (**F**) Comparison between methanol conversion profiles for CHA/AEI (OSDA **14**), CHA (OSDA **12**), and AEI (OSDA **3**). (**G**) Selectivity at $X_{95}$ for CHA (OSDA **12**), AEI (OSDA **3**), and CHA/AEI (OSDA **14**). The selectivity of CHA/AEI (OSDA **14**) is between that of CHA (OSDA **12**) and AEI (OSDA **3**).

in the experimental phase selectivity of CHA. Aluminosilicate CHA (SSZ-13) is typically synthesized using *N,N,N*-trimethyladamantammonium (TMAda, denoted here as OSDA **5**), a relatively expensive and complex OSDA (*31*) (Fig. 3A). Beyond TMAda, *N*-ethyl-*N,N*-dimethylcyclohexanaminium (OSDA **7**) has also been described for the synthesis of SSZ-13 as a simpler OSDA (*32*, *33*). Accordingly, our descriptors indicated that OSDAs **5** and **7** displayed similar shapes (Fig. 3B) and synthesis complexity (Fig. 3D), although OSDA **5** was closer to the ideal volume of 210 $Å^3$ than OSDA **7** (Fig. 3C).

To probe the influence of shape in the phase selectivity of zeolites, we compared the ability of OSDA **7** to synthesize CHA with that of *N,N,N*-trimethylcyclohexanaminium (OSDA **6**) (*34*) and *N,N*-diethyl-*N*-methylcyclohexanaminium (OSDA **8**) (*35*). OSDAs **6** to **8** had favorable binding and templating energies toward CHA (fig. S11) and were effective templates for this framework. Although they are chemically similar, their geometrical shapes were substantially different within the precision of our descriptors (Fig. 3B), and their binding strengths increased from OSDA **6** through **8**. A design approach considering only the binding metrics would predict OSDA **8** as the best template for CHA. To verify this hypothesis, we attempted the synthesis of SSZ-13 using these three templates under a variety of conditions, including using amorphous or high-silica crystalline Si and Al sources, different Si/Al and Na/Si ratios, and with and without the presence of seeds (fig. S12). As expected by its less favorable templating energy and poorer geometrical fit, OSDA **6** showed the narrowest synthesis windows among the three templates, being unable to fully crystallize CHA when high-silica FAU was used as a Si source under Na-free conditions (fig. S12A), or in most syntheses using amorphous sources (fig. S12, B and C). Notably, OSDA **7** outperformed OSDA **8** in more demanding synthesis conditions (fig. S12D), despite OSDA **7** having the weaker binding energy of the two. These results indicate that other effects may lower the experimental selectivity of OSDA **8** toward CHA compared with OSDA **7**. For example, the shape of OSDA **8** is further away from the ideal shape of OSDA **5** than OSDA **7**. When OSDA **8** was occluded in the chabazite cage, its shape experienced larger geometrical distortions to achieve the final, occluded conformation (fig. S13), which increased its free energy and reduced its ability to crystallize CHA. Thus, designing OSDAs with appropriate shapes and volumes may be as important as optimizing binding energies for improving phase selectivity in zeolite synthesis, particularly when molecules already exhibit favorable competition metrics toward the targeted frameworks.

### Modulating aluminum distributions through OSDA design

Although dispersion interactions typically provide framework selectivity (*22*), electrostatic interactions between the framework and structure-directing agents rule heteroatom distributions, influencing the stability and catalytic properties of the product zeolite (*36*–*38*). OSDAs with strong affinity and distinct charge distribution may overcome natural aluminum biases and crystallize desired frameworks with tailored aluminum distributions. To screen for charge distribution in OSDAs without using electronic structure calculations, we proposed a geometric descriptor by taking the distance between the quaternary nitrogen atom in a monocationic OSDA to the center of the mass of the molecule (fig. S14A). This descriptor allowed us to compare molecules with different charge distributions and templating abilities toward CHA (fig. S14C). Accordingly, the commercially available tetraethylammonium (OSDA **9**) and spiro-type molecules (OSDAs **10** to **13**) were molecules with good templating energy for the synthesis of SSZ-13 and have charge distributions that considerably differ from TMAda (Fig. 3A and fig. S14C). Although alkyl-substituted spiro-type OSDAs have been used for the synthesis of CHA before (*39*), alkyl groups increase the complexity of the organic synthesis, potentially adding purification and separation steps, decreasing the overall product yields, and increasing the costs. Furthermore, alkyl substituents result in uncontrolled diastereomeric mixtures, which can have a substantial influence on the zeolite nucleation and crystallization processes (*40*, *41*). On the other hand, nonsubstituted spiro-type molecules are more easily prepared, as reflected by their lower SCScores compared with that of OSDA **5** (Fig. 3D). In particular, the 6-azaspiro[5.6] dodecan-6-ium (OSDA **12**) had a near-ideal volume for CHA (Fig. 3C). Although its templating energy was higher than that of OSDA **5**, it was comparable to the commercial OSDA **7** and lay near the Pareto frontier between templating energy and SCScore for OSDAs in CHA (Fig. 3D). This fact implied that OSDA **12** offered one of the best trade-offs between templating ability and synthetic complexity among the candidates studied in this work in addition to its distinctive charge distribution. Furthermore, despite having been previously described as an OSDA unable to synthesize CHA in various conditions (*42*), our calculations showed that OSDA **12** had favorable competition metrics toward CHA (fig. S15).

On the basis of these predictions, the synthesis of CHA using OSDA **12** was first performed using the following synthesis gel composition: 1 $SiO_2$: 0.036 $Al_2O_3$: 0.2 OSDA **12**: 0.3 NaOH: 15 $H_2O$ at 150°C for 12 days, where the high-silica FAU CBV720 was used as the Si/Al source. The main physicochemical characteristics of this material are summarized in fig. S16, where the tetrahedrally coordinated Al species in framework positions are demonstrated along with a crystalline nature and textural properties similar to the baseline CHA. With this successful proof of concept, we compared the effect of templating for the broader family of spiro-type OSDAs **10** to **13** using our binding metrics and a wider range of synthesis conditions. Figure S17A shows that OSDAs **10** and **13** were unable to crystallize CHA in the synthesis conditions tested, as predicted by their unfavorable shapes, volumes, templating energies, and phase selectivity toward CHA (Fig. 3, B and C, and fig. S15). On the other hand, OSDAs **11** and **12** displayed more favorable phase selectivity toward CHA and successfully crystallized the framework (fig. S17A). In particular, a better pore-filling effect from OSDA **12**, as evidenced by its larger volume, improved its selectivity toward SSZ-13 and exhibited a larger synthesis window, despite the comparable templating energies between OSDAs **11** and **12**. Additionally, OSDA **12** was able to crystallize CHA in more demanding conditions, such as using amorphous or Al-rich crystalline sources (fig. S17, D and E).

We then performed density functional theory (DFT) calculations, which explicitly account for electrostatics (*21*), to analyze the distribution of aluminum sites in the cha cage as a function of the OSDA (materials and methods). Because sodium is typically present in the reaction medium, it was included explicitly in the simulations as a cotemplate. By computing the relative energies of 48 different Al pairs for each molecule (*43*) (fig. S18), we found that different OSDAs imprinted different aluminum distribution onto the material (Fig. 3E and fig. S19). Assuming that the Al sites distributed themselves in the framework at the thermodynamic equilibrium of OSDA-zeolite interactions, energy differences could be translated into ensemble probabilities for the Al distribution at a given temperature (433 K in Fig. 3E and fig. S19). Although kinetic factors beyond OSDAs can also bias aluminum distributions, this equilibrium assumption struck a balance between accuracy and computational cost and enabled us to compare how OSDAs influenced aluminum pairing quantitatively and predictively. For the particular case of CHA, the equilibrium distribution of aluminum sites was found to be highly sensitive to the OSDA spatial charge distribution, and the nitrogen center descriptor showed a linear relationship with the fraction of aluminum pairs expected in d6r units (Fig. 3F). OSDAs with the nitrogen center closer to the center of the cha cage, such as OSDAs **9** or **12**, were less

likely to form paired sites in the d6r units compared with the more polarized OSDAs **5** to **7**. Notably, OSDA **8** directed the formation of fewer paired aluminum sites in the d6r than OSDA **7** because of a larger distance between the nitrogen and the d6r unit imposed by the two ethyl groups. Although disentangling the effects of electrostatic and shape in the CHA (OSDA **8**) synthesis is not straightforward, the nearly equivalent charge distributions of OSDAs **5**, **6**, and **7** support the fact that the differences in phase selectivity between the three OSDAs were the result of templating effects rather than electrostatics.

Cotemplating of OSDA **12** and $Na^+$ was predicted to yield substantially fewer aluminum pairs in d6r units compared with cha cages synthesized with OSDA **5** and $Na^+$. To verify these predictions, we performed cobalt titration experiments to quantify Al pairing in CHA samples with similar physicochemical characteristics but synthesized with different OSDAs (materials and methods). Although cobalt titrations cannot fully characterize aluminum pairs because of uncertainties intrinsic to the technique, they are often useful for uncovering trends between samples (*44*). CHA crystals synthesized with OSDA **5** showed, on average, 3.4 ± 0.8 times as many $Co^{2+}$ ions per aluminum than counterparts synthesized by OSDA **12** (Fig. 3F, inset, and fig. S20). This result was in quantitative agreement with the 3.05 ratio predicted by DFT and demonstrates that the newly proposed OSDA **12** provided a different aluminum bias when synthesizing CHA. As paired aluminum sites in the eight-membered ring of the cha cage are more relevant for catalytic applications (*38*), OSDA **12** may enhance the catalytic properties of CHA zeolites. Hence, controlling the distribution of aluminum sites through OSDA design may be a pathway to improve the performance of zeolite catalysts.

Finally, to analyze the cotemplating effect of sodium as the inorganic structure-directing agent, we used OSDA **12** with potassium hydroxide in the synthesis media. In these experiments, CHA was absent, and the EAB framework was detected instead (fig. S17E). Notably, OSDA **12** had the same binding energy toward the main cavities of both zeolites (−183 kJ/mol OSDA). We surmise that because the $K^+$ ion is too large to occupy the d6r position as directed by $Na^+$ (*43*), the gme cage of the EAB zeolite is favored instead. Therefore, the final zeolite crystallized by OSDA **12** had cages that solvated the template equally well, but the cooperation between inorganic and organic SDAs determined the other building units that crystallized the framework around the OSDA. This fact suggests that quantifying phase competition beyond zeolite-OSDA interactions could move the field beyond expert-informed selection of inorganic conditions and toward a fully automated prediction of zeolite synthesis.

### OSDA design for intergrown zeolites

Intergrown zeolites offer distinctive properties compared with pure zeolite phases, such as different diffusional pathways and molecule confinements, or particular crystallographic environments to incorporate metallic active sites (*45*–*47*). Among the intergrown catalysts of industrial relevance, the silicoaluminophosphate (SAPO) form of CHA/AEI has shown distinct product selectivity and catalyst lifetime for the methanol-to-olefins (MTO) process compared with the pure crystalline CHA and AEI phases (*48*, *49*). However, SAPO-type catalysts have smaller hydrolytic and hydrothermal stability, as well as lower acidity compared with aluminosilicate zeolites. High-silica, intergrown small-pore zeolites (e.g., AFX/CHA) are often prepared using mixtures of OSDAs favorable for each individual zeolite (*50*, *51*). However, this dual OSDA approach requires empirically fine-tuning the synthesis parameters to avoid the crystallization of two independent zeolite phases, which substantially increases the complexity and cost of the process. To obtain a low-cost, aluminosilicate CHA/AEI intergrowth, we used our tools to design a single OSDA that crystallizes the targeted intergrown structure.

On the basis of the interplay between shape and binding shown in this work, we hypothesized that designing an OSDA for a targeted intergrowth required not only balancing phase competition, but also controlling the synthesis of structural motifs in the frameworks of interest (*52*). Figure 4A compares the binding energies of OSDAs toward AEI, CHA, and all other zeolites. Most of the OSDAs that synthesized AEI were more favorable toward this framework than toward CHA, as expected from our AUC analysis and traditional phase competition assumptions (*17*). The selectivity of OSDAs **1**, **3**, **5**, and **12** to their product phases was also demonstrated by their distance from the equal-energy diagonal and low competition energy. Notably, the simulation outcomes revealed a phase boundary in the OSDA shape landscape (Fig. 4B). Stronger binders toward AEI fell within the region of OSDAs with aspect ratio closer to 1:1, and longer molecules favored the CHA framework. Differences of binding energies of OSDAs toward CHA and AEI highlighted these two distinct domains, shown in Fig. 4B with red and blue colors, respectively. From these two principles and after data exploration in OSDB (movie S1), we selected *N*-ethyl-*N*-isopropyl-*N*-methylpropan-2-aminum (OSDA **14**) as a potential candidate to direct the crystallization of the CHA/AEI intergrowth. OSDA **14** had comparable binding energies to both pure phases (fig. S21), and its shape lay on the phase boundary between the two frameworks. In fact, OSDA **14** resembled an interpolation between the best OSDAs for AEI (OSDA **1**) and CHA (OSDA **5**) both in the shape space (Fig. 4B) and when the OSDAs are visualized with their charge densities (Fig. 4C).

On the basis of this theoretical lead, a CHA/AEI intergrowth was prepared under synthesis conditions similar to the ones successfully used in the previous examples: 1 $SiO_2$: 0.036 $Al_2O_3$: 0.3 OSDA **14**: 0.2 NaOH: 15 $H_2O$, with the crystallization carried out at 140°C for 5 days. The powder x-ray diffraction (PXRD) pattern of the as-prepared material showed broadened peaks at angles mainly characteristic of CHA (Fig. 4D). Additional smeared peaks also appeared at angles compatible with the AEI phase, confirming the presence of both phases in the solid. To rule out the presence of a physical mixture of crystalline particles, the as-prepared sample was imaged by scanning electron microscopy, revealing uniform particles with an average size of ~200 nm (fig. S22A). OSDA **14** molecules remained intact after the hydrothermal synthesis of the CHA/AEI (OSDA **14**) sample (table S1 and fig. S22B). The PXRD pattern of the calcined form of CHA/AEI (OSDA **14**) indicated that the sample was formed by a 50:50 CHA:AEI intergrowth as compared with DIFFaX simulations (figs. S23 and S24). High-resolution scanning transmission electron microscopy images confirmed the intergrowth structure (Fig. 4E), showing an alternating stacking sequence caused by the orientation of the d6r units in CHA and AEI frameworks. The intergrowth CHA/AEI material showed a Si/Al molar ratio of 11 (table S2) and a micropore volume of ~0.28 $cm^3/g$, similar to those of the pure AEI and CHA materials.

To illustrate the application of the intergrowth, the structure was tested as a catalyst for the MTO reaction at 350°C (Fig. 4F). The $X_{95}$ catalyst lifetime observed for the CHA/AEI (OSDA **14**) material was 788 min, which was similar to that of the CHA (OSDA **12**) catalyst reported in this work. We speculate that different active site distributions along cha and aei cavities influenced the overall diffusion pathways of the reaction-involved light olefins, thereby affecting the catalyst deactivation mechanisms. However, more fundamental characterization work should be done to elucidate this phenomenon. Regarding product selectivity, the CHA/AEI intergrowth showed the production of equivalent amounts of ethylene and propylene (~40 to 42%; table S3 and fig. S25) and ~16% of butene (table S3). These olefin selectivities were between the values achieved for the pure phases (Fig. 4G), as also shown by the $C_3^=/C_2^=$ and $C_4^=/C_2^=$ ratios (table S3), which further supports the claim that the product

selectivity for the MTO reaction was governed by the zeolite cages present in the final intergrown crystals.

## Conclusions

High-throughput computation allows for the control of phase competition in zeolite synthesis beyond Edisonian approaches that have dominated zeolite synthesis to date. Proposed metrics of phase competition, retroactively validated on comprehensive literature data, identified selective OSDAs through energetic, geometric, and electrostatic arguments. In three examples, competition was suppressed by designing OSDAs with favorable shapes, sizes, and binding metrics, achieving pure phases with wider synthesis windows. Reciprocally, it was possible to tailor intergrowths by designing a single OSDA that balanced phase competition both through shape and binding metrics. The ability of the method to recover literature outcomes from noncharged simulations suggested that product zeolites could often be predicted using only dispersion interactions. However, the final product is often dependent on cotemplating effects between inorganic and organic SDAs. Additionally, tuning OSDA charge and polarization modulates the aluminum distribution of the crystallized CHA zeolite, which can have important implications for catalysis. The computational results on templating effects can be further used to train data-driven selectivity models based on molecular structure descriptors, increasing the breadth of high-throughput screening undertakings. At the same time, developing more general electrostatic descriptors may also guide the investigation of aluminum distributions in zeolites beyond CHA. Finally, interactive visualization and analysis of said data through the OSDB portal is expected to empower domain experts to leverage these advances and accelerate the discovery cycle of zeolites.

## REFERENCES AND NOTES


1. W. Vermeiren, J.-P. Gilson, *Top. Catal.* **52**, 1131–1161 (2009).
2. Y. Li, L. Li, J. Yu, *Chem* **3**, 928–949 (2017).
3. M. Moliner, F. Rey, A. Corma, *Angew. Chem. Int. Ed.* **52**, 13880–13889 (2013).
4. P. Eliášová *et al*., *Chem. Soc. Rev.* **44**, 7177–7206 (2015).
5. E. M. Gallego *et al*., *Science* **355**, 1051–1054 (2017).
6. D. W. Lewis, D. J. Willock, C. R. A. Catlow, J. M. Thomas, G. J. Hutchings, *Nature* **382**, 604–606 (1996).
7. G. Sastre, A. Cantin, M. J. Diaz-Cabañas, A. Corma, *Chem. Mater.* **17**, 545–552 (2005).
8. J. E. Schmidt, M. W. Deem, M. E. Davis, *Angew. Chem. Int. Ed.* **53**, 8372–8374 (2014).
9. F. Daeyaert, F. Ye, M. W. Deem, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 3413–3418 (2019).
10. K. Muraoka, W. Chaikittisilp, T. Okubo, *Chem. Sci.* **11**, 8214–8223 (2020).
11. R. Pophale, F. Daeyaert, M. W. Deem, *J. Mater. Chem. A Mater. Energy Sustain.* **1**, 6750–6760 (2013).
12. K. Muraoka, Y. Sada, D. Miyazaki, W. Chaikittisilp, T. Okubo, *Nat. Commun.* **10**, 4459 (2019).
13. C. Baerlocher, L. B. McCusker, Database of Zeolite Structures (2021); www.iza-structure.org/databases/.
14. M. M. J. Treacy, I. Rivin, E. Balkovsky, K. H. Randall, M. D. Foster, *Microporous Mesoporous Mater.* **74**, 121–132 (2004).
15. M. W. Deem, R. Pophale, P. A. Cheeseman, D. J. Earl, *J. Phys. Chem. C* **113**, 21353–21360 (2009).
16. P. Wagner *et al*., *J. Am. Chem. Soc.* **122**, 263–273 (2000).
17. A. W. Burton, G. S. Lee, S. I. Zones, *Microporous Mesoporous Mater.* **90**, 129–144 (2006).
18. Z. Jensen *et al*., *ACS Cent. Sci.* **5**, 892–899 (2019).
19. Z. Jensen *et al*., *ACS Cent. Sci.* **7**, 858–867 (2021).
20. D. Schwalbe-Koda, R. Gomez-Bombarelli, *J. Phys. Chem. C* **125**, 3009–3017 (2021).
21. D. Schwalbe-Koda, R. Gómez-Bombarelli, *J. Chem. Phys.* **154**, 174109 (2021).
22. R. F. Lobo, S. I. Zones, M. E. Davis, *J. Incl. Phenom. Macrocycl. Chem.* **21**, 47–78 (1995).
23. A. Corma, F. Rey, J. Rius, M. J. Sabater, S. Valencia, *Nature* **431**, 287–290 (2004).
24. B. H. Toby, N. Khosrovani, C. B. Dartt, M. E. Davis, J. B. Parise, *Microporous Mesoporous Mater.* **39**, 77–89 (2000).
25. R. E. Boyett, A. P. Stevens, M. G. Ford, P. A. Cox, *Zeolites* **17**, 508–512 (1996).
26. R. Gómez-Bombarelli *et al*., *Nat. Mater.* **15**, 1120–1127 (2016).
27. M. Moliner, Y. Román-Leshkov, A. Corma, *Acc. Chem. Res.* **52**, 2971–2980 (2019).
28. J. E. Schmidt, M. W. Deem, C. Lew, T. M. Davis, *Top. Catal.* **58**, 410–415 (2015).
29. S. I. Zones, Y. Nakagawa, S. T. Evans, G. S. Lee, Zeolite SSZ-39, US Patent 5,958,370 (1999).
30. C. W. Coley, L. Rogers, W. H. Green, K. F. Jensen, *J. Chem. Inf. Model.* **58**, 252–261 (2018).
31. S. I. Zones, *J. Chem. Soc., Faraday Trans.* **87**, 3709–3716 (1991).
32. G. Cao, M. M. Mertens, A. S. Guram, H. Li, J. C. Yoder, Synthesis of chabazite-containing molecular sieves and their use in the conversion of oxygenates to olefins, US Patent 7,754,187 (2007).
33. X. Wang *et al*., *Chem. Commun.* **51**, 16920–16923 (2015).
34. T. M. Davis, S. Elomari, S. I. Zones, Method for preparing cha-type molecular sieves using colloidal aluminosilicate, US Patent 2014/0147378 (2014).
35. G. Cao, M. Mertens, A. S. Guram, H. Li, J. C. Yoder, Synthesis of chabazite-containing molecular sieves and their use in the conversion of oxygenates to olefins, US Patent 2008/0045767 (2008).
36. J. R. Di Iorio, C. T. Nimlos, R. Gounder, *ACS Catal.* **7**, 6663–6674 (2017).
37. E. M. Gallego *et al*., *Chemistry* **24**, 14631–14635 (2018).
38. F. Göltl, S. Bhandari, M. Mavrikakis, *ACS Catal.* **11**, 7719–7734 (2021).
39. Y. Nakagawa, G. S. Lee, T. V. Harris, L. T. Yuen, S. I. Zones, *Microporous Mesoporous Mater.* **22**, 69–85 (1998).
40. M. Dusselier *et al*., *Chem. Mater.* **27**, 2695–2702 (2015).
41. C. Kim, S.-J. Hwang, A. W. Burton, S. I. Zones, *Microporous Mesoporous Mater.* **116**, 227–232 (2008).
42. R. Millini, L. Carluccio, F. Frigerio, W. O'Neil Parker Jr., G. Bellussi, *Microporous Mesoporous Mater.* **24**, 199–211 (1998).
43. J. R. Di Iorio *et al*., *J. Am. Chem. Soc.* **142**, 4807–4819 (2020).
44. J. R. Di Iorio, R. Gounder, *Chem. Mater.* **28**, 2236–2247 (2016).
45. H. Kosslick *et al*., *J. Chem. Soc., Faraday Trans.* **90**, 2837–2844 (1994).
46. T. Willhammar *et al*., *Nat. Chem.* **4**, 188–194 (2012).
47. J. S. Bates, B. C. Bukowski, J. W. Harris, J. Greeley, R. Gounder, *ACS Catal.* **9**, 6146–6168 (2019).
48. L. Guo *et al*., *Ind. Eng. Chem. Res.* **57**, 10398–10402 (2018).
49. R. L. Smith *et al*., *Appl. Catal. A Gen.* **505**, 1–7 (2015).
50. Y. Naraki, K. Ariga, K. Nakamura, K. Okushita, T. Sano, *Microporous Mesoporous Mater.* **254**, 160–169 (2017).
51. G. Cao *et al*., Chabazite-containing molecular sieve, its synthesis and its use in the conversion of oxygenates to olefins, US Patent 7,094,389 B2 (2006).
52. D. Schwalbe-Koda, Z. Jensen, E. Olivetti, R. Gómez-Bombarelli, *Nat. Mater.* **18**, 1177–1181 (2019).
53. J. Towns *et al*., *Comput. Sci. Eng.* **16**, 62–74 (2014).
54. D. Schwalbe-Koda, learningmatter-mit/VOID: VOID 1.0.1., version 1.0.1, Zenodo (2021); https://doi.org/10.5281/zenodo.5260054.
55. D. Schwalbe-Koda, learningmatter-mit/gulpy: GULPy 1.0., version 1.0, Zenodo (2021); https://doi.org/10.5281/zenodo.5260056.
56. B. Blaiszik *et al*., *JOM* **68**, 2045–2052 (2016).
57. B. Blaiszik *et al*., *MRS Commun.* **9**, 1125–1133 (2019).
58. D. Schwalbe-Koda, R. Gómez-Bombarelli, Data for: Ab initio control of zeolite synthesis and intergrowth with high-throughput simulations, Data set, Materials Data Facility (2021); https://doi.org/10.18126/C5Z9-ZEJ7.
59. D. Schwalbe-Koda, learningmatter-mit/Zeolite-Phase-Competition: Article data for: Ab initio control of zeolite phase competition and intergrowth with high-throughput simulations, version 1.0, Zenodo (2021); https://doi.org/10.5281/zenodo.5272354.


## ACKNOWLEDGMENTS

The Electron Microscopy Service of the UPV is acknowledged for their help in sample characterization. J. Martínez-Triguero and C. Li are also acknowledged for helpful discussions on the MTO reaction, and B. Blaiszik is acknowledged for help with persistent data storage at the Materials Data Facility. **Funding:** D.S.-K. and R.G.-B. acknowledge the Energy Initiative (MITEI) and MIT International Science and Technology Initiatives (MISTI) Seed Funds. D.S.-K. was also funded by the MIT Energy Fellowship. C.P., E.B.-J., M.M., and A.C. acknowledge financial support by the Spanish government through the "Severo Ochoa" program (SEV-2016-0683, MINECO) and grant RTI2018-101033-B-I00 (MCIU/AEI/FEDER, UE). E.B.-J. acknowledges the Spanish government for an FPI scholarship (PRE2019-088360). Z.J., E.O., S.K., and Y.R.-L. acknowledge partial funding from Designing Materials to Revolutionize and Engineer our Future (DMREF) from the National Science Foundation (NSF); awards 1922311, 1922372, and 1922090; and the Office of Naval Research (ONR) under contract N00014-20-1-2280. S.K. was additionally funded by the Kwanjeong Educational Fellowship. Z.J. was also supported by the Department of Defense (DoD) through the National Defense Science Engineering Graduate (NDSEG) fellowship program. T.W. acknowledges financial support by the Swedish Research Council (grant no. 2019-05465). Computer calculations were executed at the Massachusetts Green High-Performance Computing Center with support from MIT Research Computing and at the Extreme Science and Engineering Discovery Environment (XSEDE) (*53*) Expanse through allocation TG-DMR200068. **Author contributions:** D.S.-K. designed and executed the simulations, computational analysis, and software development. C.P. carried out the design and synthesis of the OSDAs. S.K. and E.B.-J. performed the synthesis and characterization of zeolites. E.B.-J. executed the MTO experiments. T.W. performed the transmission electron microscopy analysis of the CHA/AEI intergrowth and the simulation of the PXRD patterns of different CHA/AEI intergrown phases. Z.J. and E.O. provided the literature data. R.G.-B. conceived the project and supervised the computational research. M.M., Y.R.-L., and A.C. supervised the experimental research. D.S.-K., M.M., and R.G.-B. wrote the first manuscript draft. All authors contributed to the analysis of the data and to the final manuscript. **Competing interests:** The authors are coinventors on two patent applications based on the materials synthesized in this work. **Data and materials availability:** The code used to automate the simulations of this work is available at Zenodo (*54*, *55*). Literature data are available in (*19*). Results of atomistic simulations are available at the Materials Data Facility (*56*, *57*) through the accession code available at (*58*). An interface to these data is available online at OSDB (http://zeodb.mit.edu/). The code and data used to produce all graphs in the manuscript are available at Zenodo (*59*). All other data needed to evaluate the conclusions of the paper are present in the paper or the supplementary materials.

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abh3350
Materials and Methods
Figs. S1 to S25
Tables S1 to S3
References (*60*–*80*)
Movie S1

5 March 2021; resubmitted 14 July 2021
Accepted 31 August 2021
Published online 16 September 2021
10.1126/science.abh3350

NEUROSCIENCE

# Modulation of tactile feedback for the execution of dexterous movement

James M. Conner, Andrew Bohannon†, Masakazu Igarashi†, James Taniguchi, Nicholas Baltar, Eiman Azim*

**Although dexterity relies on the constant transmission of sensory information, unchecked feedback can be disruptive. Yet how somatosensory feedback from the hands is regulated and whether this modulation influences movement remain unclear. We found that mouse tactile afferents recruit neurons in the brainstem cuneate nucleus, whose activity is modulated by distinct classes of local inhibitory neurons. Manipulation of these inhibitory circuits suppresses or enhances the transmission of tactile information, which affects manual behaviors. Top-down cortical pathways innervate cuneate in a complementary pattern, with somatosensory cortical neurons targeting the core tactile region of cuneate and a large rostral cortical population driving feed-forward inhibition of tactile transmission through an inhibitory shell. These findings identify a circuit basis for tactile feedback modulation that enables the effective execution of dexterous movement.**

Much of our interaction with the world occurs through movements of the hand. This "organ of considerable virtuosity" (*1*) achieves its impressive dexterity through dynamic interactions between motor output and sensory feedback (*2*, *3*). Not all feedback is treated equally, however. Signals can be disruptive to behavior when they are noisy, self-generated, or temporally delayed, implying circuit mechanisms for regulating the transmission of ascending information (*3*–*5*). Pathways responsible for feedback modulation throughout the nervous system have begun to be defined (*6*–*12*), revealing circuits that can regulate specific types of incoming signals to facilitate behavior. For somatosensation, injury of the pathways that carry ascending feedback from the limbs into the brain severely affects the smooth execution of dexterous behaviors across species—including rodents, cats, monkeys, and humans (*13*–*15*)—which highlights critical roles for afferent signals from the skin and muscles in the control of movement (*2*). Yet the functional organization of neural circuits that regulate the transmission of ascending somatosensory feedback and any influence this modulation has on dexterous motor output remain less clear.

The cuneate nucleus in the dorsal brainstem forms the major conduit for sensory signals from the hand ascending to the sensorimotor cortex (*16*, *17*). Cuneate neurons receive forelimb sensory signals directly from the dorsal root ganglia, as well as indirectly through postsynaptic dorsal column pathway neurons in the cervical spinal cord (Fig. 1A and fig. S1A) (*17*). The core, or clusters, region of the middle cuneate (hereafter referred to as Cu) receives tactile input mainly from the distal forelimbs and is thought to specialize in processing discriminative touch signals from the hand (*2*, *18*, *19*). Cuneolemniscal neurons in Cu receive afferent input and project to several subcortical targets, most predominately the ventral posterolateral nucleus (VPL) of the thalamus, which then conveys sensory information to primary somatosensory cortex (SSp) (Fig. 1A) (*16*, *19*). The input and output connectivity and tactile specialization of this clusters region appears to be conserved across mammalian species, although the precise organization of hand somatotopy can differ (*17*). Cu receives descending input from corticofugal projections, and cortical stimulation experiments, primarily performed in cats and monkeys, have provided evidence for both excitation and inhibition of cuneate neurons (*18*–*25*). These findings suggest a mechanism for top-down modulation, in which the same cortical circuits that receive forelimb sensory feedback are responsible for regulating the flow of this incoming peripheral information to the brainstem (*26*, *27*), potentially through the recruitment of local inhibitory neurons (*27*–*30*).

However, lack of circuit-specific access has limited efforts to define the organization of these inhibitory neurons, the relationship they might have to descending cortical pathways, and any impact their putative feedback modulation might have on dexterous forelimb movement. It thus remains unclear which neural circuits regulate the incessant flow of sensory signals to ensure that only the appropriate and salient information is used to affect ongoing behavior.

## Cuneate tactile and inhibitory inputs

We first localized the region of the middle cuneate that receives direct tactile input from the hand. Cutaneous afferents were targeted at their peripheral terminals by means of injection of cholera toxin B subunit (CTB) into the glabrous pad of the hand, revealing dense innervation of the ipsilateral Cu, which sends its most prominent output to contralateral VPL thalamus (Fig. 1B and fig. S1B) (*16*). Conversely, proprioceptive afferents were targeted by means of injection of CTB into either proximal or distal forelimb muscles, revealing minimal innervation of Cu but some targeting of other cuneate regions and dense projections to the neighboring ipsilateral external cuneate nucleus, which mainly projects to the cerebellum (Fig. 1B and fig. S1B) (*17*). We confirmed these findings by genetically restricting our tracing to proprioceptive afferents through conditional viral targeting of forelimb muscles in *Pv-Cre* mice (fig. S1C). These results establish that direct ascending cutaneous and proprioceptive afferents from the forelimb remain largely segregated at the level of the dorsal column nuclei complex in mice.

Sensory responses in Cu are subject to attenuation, which is potentially mediated by local inhibition (*21*, *27*, *28*, *30*). Yet the detailed organization of local circuits that can affect tactile transmission and any contribution they might have to movement are not well understood. Having localized the tactile recipient Cu region, we next sought to characterize the properties of cuneolemniscal neurons in Cu by recording from adult brainstem slices (Fig. 1C). Whole-cell recordings revealed that retrogradely labeled cuneolemniscal neurons had a relatively depolarized resting membrane potential (mean = −57.60 ± 1.22 mV SEM) (Fig. 1D) that was, on average, slightly below the action potential threshold (mean = −49.67 ± 1.78 mV SEM) (fig. S2A) (*31*). These neurons also received extensive spontaneous inhibitory input that was abolished by the combined application of strychnine and bicuculline (Fig. 1, E and F, and fig. S2B), which revealed γ-aminobutyric acid (GABA)–ergic and glycinergic inputs.

To identify the location of neurons that might provide this inhibition, we genetically restricted fluorophore expression to GABAergic and glycinergic neurons through conditional viral targeting of the Cu region of *VGAT-Cre* mice. We found inhibitory neurons within the core region of the Cu, where cuneolemniscal neurons reside, as well as throughout a shell region ventral to the Cu core, where cuneolemniscal neurons are absent (collectively referred to as Cu inhibitory neurons) (Fig. 1G). Visualization of a synaptically tagged fluorophore revealed that axon terminals that arise from these local inhibitory neurons project heavily into the Cu core region, where they provide dense synaptic input onto cuneolemniscal neurons (Fig. 1G). Genetic labeling of inhibitory subclasses revealed that glycinergic neurons are present throughout the Cu core and shell

Molecular Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, CA, USA.
*Corresponding author. Email: eazim@salk.edu
†These authors contributed equally to this work.

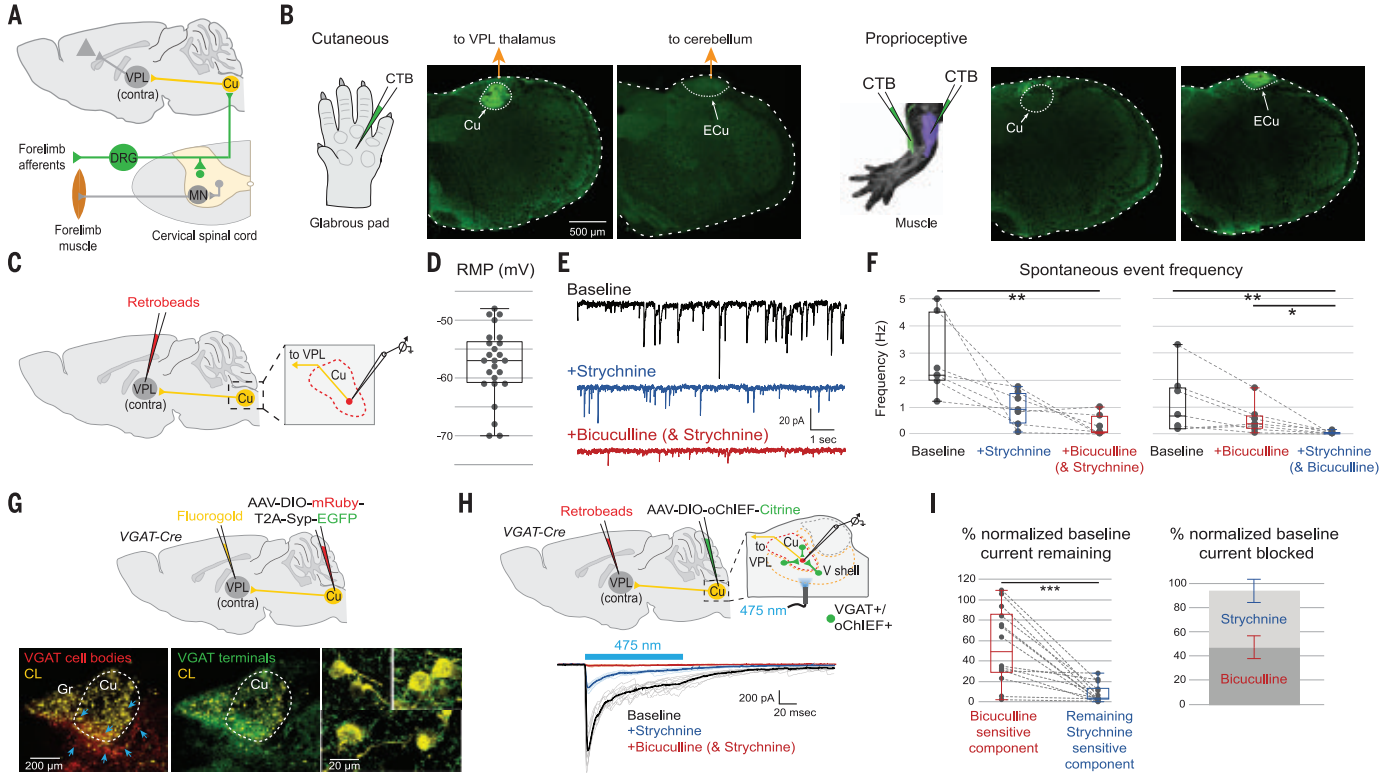



**Fig. 1. Cuneate tactile and inhibitory inputs.** (**A**) Cu receives direct input from forelimb afferents in the dorsal root ganglia (DRG) and indirect input from ascending projections from the spinal cord (not shown). Cuneolemniscal (CL) neurons project to VPL thalamus, as well as other targets (*17*). MN, motor neuron. (**B**) Labeling direct cutaneous (left) and proprioceptive (right) projections by means of CTB injection into peripheral end organs. Cutaneous afferents from the glabrous pad of the hand innervate the ipsilateral Cu but avoid the external cuneate (ECu), whose major target is the cerebellum (six mice) (*17*). Proprioceptive afferents from forelimb muscles (biceps and triceps brachii) rarely innervate Cu but target ipsilateral ECu (right) and other regions of the cuneate nucleus (five mice) (fig. S1, B and C). (**C**) Retrograde labeling of CL neurons for slice recording by retrobead injection into contralateral VPL. (**D**) Resting membrane potential (RMP) of CL neurons (25 neurons in 11 mice; current-clamp, 0 pA holding current; all box-and-whisker plots show median, 25th and 75th percentiles, and range). (**E**) Whole-cell recordings of a CL neuron (–70 mV) showing spontaneous events at baseline (black). Bath application of strychnine (blue) and bicuculline (red) progressively eliminated spontaneous inhibitory postsynaptic currents (IPSCs). (**F**) Sequential application of strychnine and bicuculline (left; seven neurons in four mice; $^{**}P = 0.0015$) or the reverse (right; seven neurons in three mice; $^{**}P = 0.0040$, $^{*}P = 0.0485$) decreased the frequency of spontaneous IPSCs (Friedman repeated-measures test with Dunn's multiple comparisons test). (**G**) Viral labeling of inhibitory cell bodies (red) and their synaptic terminals (green) in *VGAT-Cre* mice (two mice). Inhibitory neurons (left; red cells, arrows) reside in the core and ventral shell (V shell) regions of Cu and gracile (Gr) nucleus and project into the core of Cu and Gr (middle; green), where they contact retrogradely labeled CL neurons (right; yellow). (**H**) Slice recording from CL neurons (top). Selective optogenetic activation of Cu inhibitory neurons (viral oChIEF expression in *VGAT-Cre* mice) produced large IPSCs that were eliminated through application of strychnine and bicuculline. Faint lines represent single trials; dark lines represent mean. (**I**) Cumulative reduction in the amplitude of light-evoked IPSCs (left, normalized to baseline) with an approximately equal mix of GABA and glycine components (right, error bars indicate SEM) (14 neurons in seven mice; $^{***}P = 0.0001$; Wilcoxon two-tailed matched-pairs signed-rank test) (fig. S5C).

regions, whereas GABAergic neurons reside mostly in the ventral shell (fig. S3, A and B). Monosynaptic retrograde rabies tracing that originated specifically from cuneolemniscal neurons confirmed that both GABAergic and glycinergic neurons provide direct inputs (fig. S3, C and D).

To assess the functional connectivity of these circuits, we expressed the excitatory opsin oChIEF in Cu inhibitory neurons through conditional viral injection in *VGAT-Cre* mice and performed whole-cell recording from adult brainstem slices. Targeted inhibitory neurons showed robust responses to photoactivation (fig. S4, A to C). These inhibitory neurons also received extensive spontaneous inhibitory inputs that were largely GABAergic (fig. S4, D to F) (*30*), which suggests a means for local disinhibition. We next recorded postsynaptic responses from tagged cuneolemniscal neurons. Optogenetic activation of local Cu inhibitory neurons produced large inhibitory currents (Fig. 1H) with onset kinetics indicative of monosynaptic connectivity (fig. S5, A and B). Combined application of strychnine and bicuculline abolished light-evoked responses (Fig. 1, H and I), with each cuneolemniscal neuron showing a distinct mix of GABAergic and glycinergic inputs (Fig. 1I and fig. S5C), which confirmed functional connectivity between distinct classes of Cu inhibitory neurons and cuneolemniscal neurons.

### Local inhibitory modulation of ascending tactile feedback

Although these findings establish the anatomical and functional relationship between Cu inhibitory neurons and cuneolemniscal neurons, they do not demonstrate whether these circuits can modulate tactile signaling. We thus performed in vivo extracellular recordings from tactile-responsive neurons in the Cu of anesthetized mice while physical stimuli were applied to the glabrous pad of the ipsilateral hand by a rotating wheel (Fig. 2A), which allowed us to identify neurons that are robustly sensitive to tactile input (Fig. 2, B and C, and fig. S6).

We then investigated how the transmission of tactile signals from the ipsilateral hand is

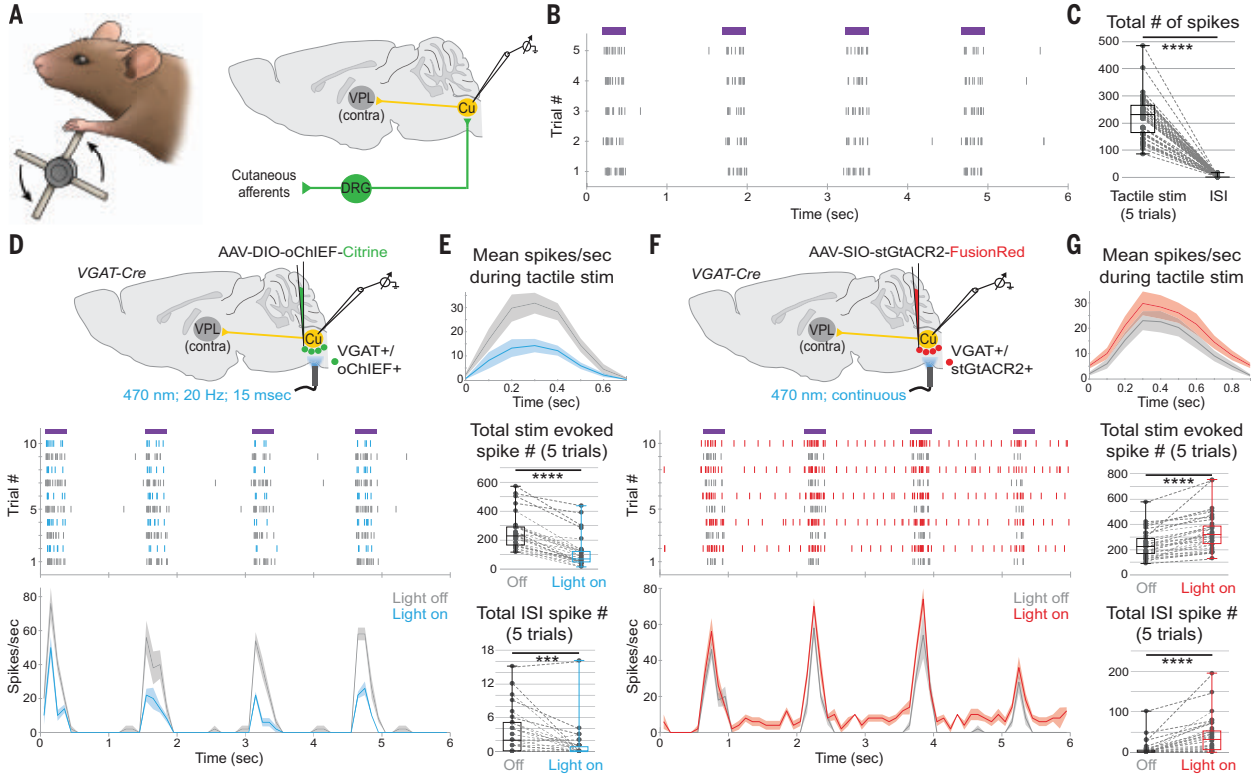



**Fig. 2. Local inhibitory circuits bidirectionally modulate tactile responses.** (**A**) In vivo extracellular recording in Cu of anesthetized mice with tactile stimuli applied to the ipsilateral glabrous pad of the hand. (**B**) Spike raster plots of tactile-evoked signals from an example Cu recording site across five trials with four stimuli per trial (20 total stimuli). Purple bars indicate periods when tactile stimulus was applied. (**C**) Total tactile-evoked spikes across all 20 stimuli for each recording site with minimal spontaneous activity during the ISIs (51 recordings across three mice; ****$P < 0.0001$; Wilcoxon two-tailed matched-pairs signed-rank test). (**D**) In vivo extracellular recording of tactile-responsive units in Cu in anesthetized mice with optogenetic activation of local Cu inhibitory neurons (top, viral oChIEF expression in *VGAT-Cre* mice). Spike raster plots (middle) and mean spike number histograms (bottom; 0.1 s bin) from an example recording site across 10 interleaved trials (five light off, gray; five light on, blue; each with four tactile stimuli, purple bars). Lines in lower histogram indicate mean; shaded areas represent SEM (fig. S6). (**E**) Mean number of spikes per second during a single tactile stimulus (top; 0.1 s bin) across all recordings (30 sites from three mice). Lines indicate mean; shaded areas represent SEM. Pairwise comparisons of the total number of spikes across five trials (20 tactile stimuli each for light off and light on) reveal suppression of tactile-evoked spikes (middle; ****$P < 0.0001$) and suppression of spontaneous activity during tactile ISI periods (bottom; ***$P = 0.0004$). (Wilcoxon two-tailed matched-pairs signed-rank test.) (**F**) In vivo extracellular recording with optogenetic inhibition of local Cu inhibitory neurons (top, viral stGtACR2 expression in *VGAT-Cre* mice). Spike raster plots from an example recording site [middle; as in (D); light off, gray; light on, red]. Lines in histogram (bottom) indicate mean; shaded areas represent SEM. (**G**) Quantification [as in (E)] across all recordings (34 sites from three mice) shows an increase in tactile-evoked responses in Cu neurons during photoinhibition (top). Pairwise comparisons reveal an increase in spikes evoked by tactile stimuli (middle; ****$P < 0.0001$) and an increase in spontaneous activity during tactile ISI periods (bottom; ****$P < 0.0001$) (Wilcoxon two-tailed matched-pairs signed-rank test).

affected by activation or inactivation of local Cu inhibitory circuits. First, conditional viral expression of oChIEF in *VGAT-Cre* mice was used to assess the impact of activating Cu inhibitory neurons on tactile-responsive neurons (Fig. 2D and fig. S6). During photostimulation, the number of tactile-evoked spikes was consistently suppressed (60.99 ± 3.71% SEM decrease in spike number), and this attenuation applied even to the rare spontaneous activity observed during tactile interstimulus interval (ISI) periods (56.93 ± 15.40% SEM decrease in spike number) (Fig. 2, D and E, and fig. S6). Next, we used conditional viral expression of the inhibitory opsin stGtACR2 in *VGAT-Cre* mice to evaluate the impact of inactivating Cu inhibitory neurons (Fig. 2F). Photoinhibition of Cu inhibitory neurons amplified the number of tactile-evoked spikes (50.46 ± 8.87% SEM increase in spike number) and elicited a large increase in ISI spiking activity (1,333.74 ± 354.81% SEM increase in spike number), which suggests that local inhibition prevents aberrant Cu neuronal firing and the transmission of spurious sensory information (Fig. 2, F and G). Reasoning that this modulation might regulate signals ascending through the medial lemniscus pathway that facilitate sensory-guided behaviors, we recorded from tactile-responsive neurons in VPL thalamus while performing the same optogenetic perturbations of Cu inhibitory neurons. Activation produced an equivalent attenuation of tactile responses in VPL thalamus as it did in Cu (61.13 ± 4.64% SEM decrease in spike number) (fig. S7, A, B, and E), whereas inhibition of Cu inhibitory neurons amplified responses in VPL thalamus, although to a more modest extent than in Cu (15.77 ± 5.32% SEM increase in tactile-evoked spikes; 517.22 ± 207.23% SEM increase in ISI spikes) (fig. S7, C to E).

### Disrupting cuneate modulation perturbs dexterous movements

The regulation of feedback is a fundamental aspect of coordinated behavior (*3*, *5*), and dexterous movements might be particularly dependent on the dynamic adjustment of sensory signaling (*2*). Fine movements are especially susceptible to the loss of somatosensory information (*1*, *13*–*15*), but lack of temporal and circuit specificity in lesion experiments leave any role for feedback modulation unclear.

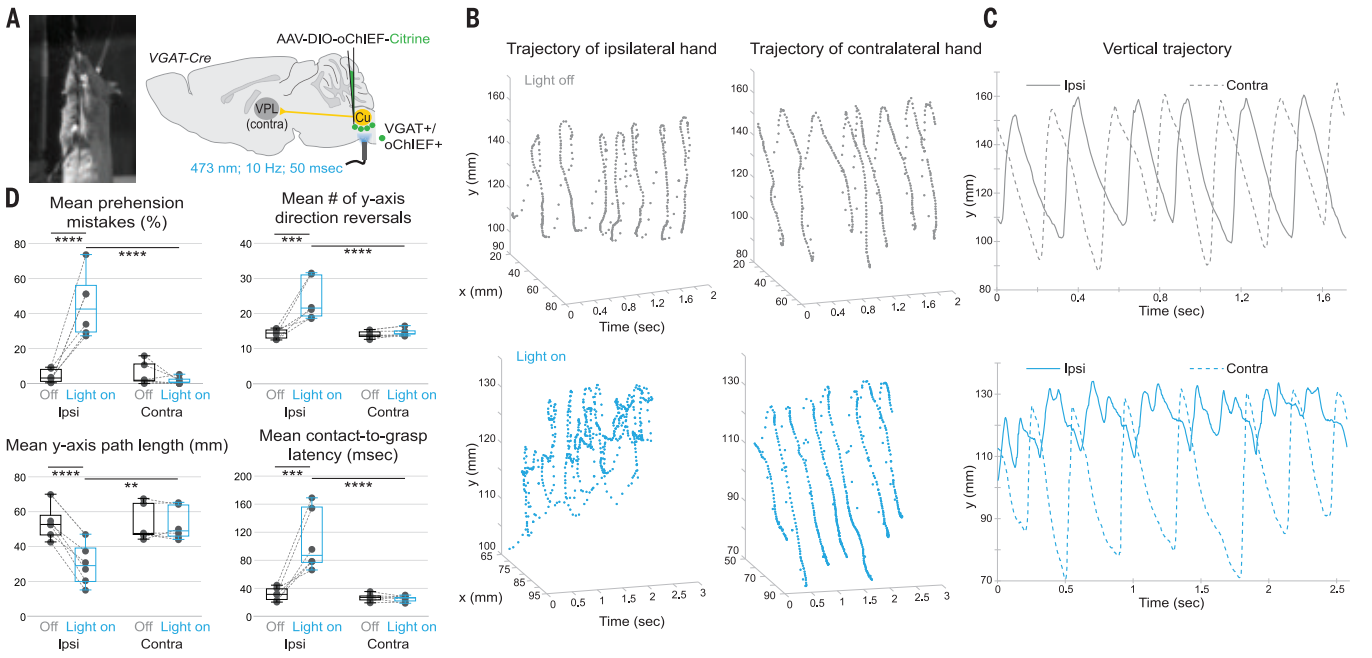



**Fig. 3. Perturbing cuneate inhibitory circuits disrupts dexterous movements.** (**A**) String-pulling task (left). Local Cu inhibitory neurons were targeted for optogenetic activation through viral oChIEF expression in *VGAT-Cre* mice (right). (**B**) Example trajectories of ipsilateral (left) and contralateral (right) hands in horizontal (*x*) and vertical (*y*) dimensions over time. Gray, light off; blue, photoactivation. (**C**) Example *y* trajectories of ipsilateral (solid line) and contralateral (dashed line) hands. Gray, light off; blue, photoactivation. (**D**) Quantification of prehension errors and kinematics across six mice. Prehension errors of the ipsilateral, but not contralateral, hand increased during photoactivation (top left; percentage of grasp attempts in which an error was made across trials; ****$P < 0.0001$), as did the mean number of ipsilateral vertical (*y*) direction reversals (top right; mean number of direction reversals per trial; ****$P < 0.0001$, ***$P = 0.0003$). The mean *y* path length traversed by the ipsilateral hand between direction reversals decreased during photoactivation (bottom left; mean absolute distance between the peak and trough of a given path segment across trials; ****$P < 0.0001$, **$P = 0.0016$). There was also an increase in the mean latency between first string contact and successful grasp during photoactivation on the ipsilateral side (bottom right; ****$P < 0.0001$, ***$P = 0.0002$). (Two-way repeated-measures analysis of variance with Sidak multiple comparisons test) (fig. S8).

Genetic access to Cu inhibitory circuits provided us the opportunity to answer a question that classical lesion studies cannot address: Does the regulation of tactile feedback as it ascends into the brain have any impact on dexterous motor control?

As a first assay, mice were trained to perform a string-pulling behavior that elicits the smooth alternation of left and right hands as the animal reaches, grasps, and pulls, mimicking many natural behaviors (Fig. 3A and movie S1) (*32*). This assay could help to distinguish two possible scenarios: Spinal tactile reflex circuits are sufficient for the smooth execution of string grasp and pull during this rhythmic behavior, or alternatively, ascending tactile signals need to be regulated appropriately in the cuneate for effective performance. To attenuate tactile signaling in the cuneate, we targeted viral delivery of oChIEF to Cu inhibitory neurons in *VGAT-Cre* mice (Fig. 3A). Photoactivation of these neurons affected the animals' ability to coordinate ipsilateral grasping movements with string contact (Fig. 3, B to D), which indicates that spinal tactile circuits alone are not sufficient for coordinating grasping in this task. Automated tracking revealed that these prehension mistakes affected limb kinematics because animals showed a greater number of hand direction reversals after grasping errors and a reduction in movement path length as they made successive attempts to correctly time the grasp (Fig. 3D and movies S1 and S2). Mice did not miss contact with the string as the ipsilateral hand moved upward to grasp, which indicates that the reach-targeting component of the behavior was not affected by Cu inhibitory neuron activation (fig. S8A). Rather, we found a large increase in the latency between initial string contact and a successful grasp (Fig. 3D), and once a grasp was successful and a pull attempted, there was a decrease in the displacement of the hands as the load was transferred from the ipsilateral to contralateral hand (fig. S8A). These findings point to a primary deficit in the accurate timing of the grasp, followed by insufficient grip as the affected hand is less able to execute the downward pull before the string is transferred to the other hand. These prehension errors and kinematic deficits did not appear in the contralateral limb or in control mice receiving photostimulation (Fig. 3, B to D, and fig. S8, A and B). Conversely, to evaluate how an aberrant increase in tactile signaling affects behavior, we expressed the inhibitory opsin stGtACR2 in Cu inhibitory neurons of *VGAT-Cre* mice. During photoinhibition, in a subset of trials, mice terminated or briefly paused string-pulling movements at light onset (six mice; mean 53.21% of trials ± 11.36% SEM) and, in the more extreme cases, dropped the string and gripped the ipsilateral hand (movie S3). The effects were variable, however, and in trials that continued, mice were no more likely to make string-targeting or prehension mistakes during photoinhibition than during control conditions. These findings suggest that eliciting aberrant transmission through the cuneate, perhaps akin to paresthesia, can disrupt behavior, although mice can be resilient to spurious sensory signals when performing a reward-driven task.

Although the string-pulling assay provides an ethologically relevant means to evaluate goal-directed limb movements, animals in this task are free to use any combination of tactile, proprioceptive, and visual feedback. We wanted to probe the tactile component of a dexterous movement more selectively. Human studies have shown that tactile acuity is surprisingly high during object manipulation when compared with psychometric detection thresholds (*33*), which suggests that

an assay of fine motor control could provide the most sensitive way to evaluate the impact of Cu feedback modulation. We therefore developed a quantitative tactile orienting assay for mice to isolate the contribution of salient cutaneous feedback to task execution (*33*). Head-fixed mice were trained to use their hand to turn a pedestal with an oriented texture cue to a defined target zone and to hold the pedestal in place (Fig. 4, A to D, and materials and methods). The pedestal was affixed to a motor-rotary encoder assembly, which enabled the pedestal texture to be set to a random starting orientation at the beginning of each trial (fig. S9A). During 4 to 6 weeks of training, the task conditions gradually became more difficult in a closed-loop fashion, depending on each animal's behavioral performance, until expert performance was achieved at the most stringent conditions (Fig. 4E, movie S4, and materials and methods).

To determine how disrupting the modulation of tactile feedback in the cuneate affected task performance, we expressed oChIEF in Cu inhibitory neurons (fig. S9B). Photoactivation did not affect overall task success in trained mice but resulted in an increase in the amount of time taken to reach the target, which appeared as a rightward shift in the cumulative distribution of successes over time (Fig. 4, F, G, and I, and fig. S9, B and C). To determine how these optogenetically induced deficits compare with normal performance when tactile information is more impoverished, we evaluated task performance using a smooth pedestal that lacked tactile ridges, which dissociated fine tactile cues from the cutaneous feedback used to detect contact with the pedestal. Unperturbed animals using a smooth pedestal also showed a rightward shift in the cumulative distribution of successes over time (Fig. 4H and fig. S9, B and C) that was nearly identical to those using a textured platform during Cu inhibitory neuron activation (Fig. 4I). In all conditions, mice exhibited equivalent peak angular velocities when making turns, which indicated that their ability to contact and rotate the pedestal was not affected (fig. S9B). Yet unperturbed mice using a textured platform were far more likely to achieve success rapidly, at the earliest stages of a trial, than were mice undergoing Cu inhibitory neuron activation or mice using a smooth platform (fig. S9C). Behavioral performance was not affected in control animals receiving photostimulation (fig. S9D). These findings suggest that when tactile signals are attenuated or salient tactile cues are removed, mice can still perform the task, but they lose a tactile

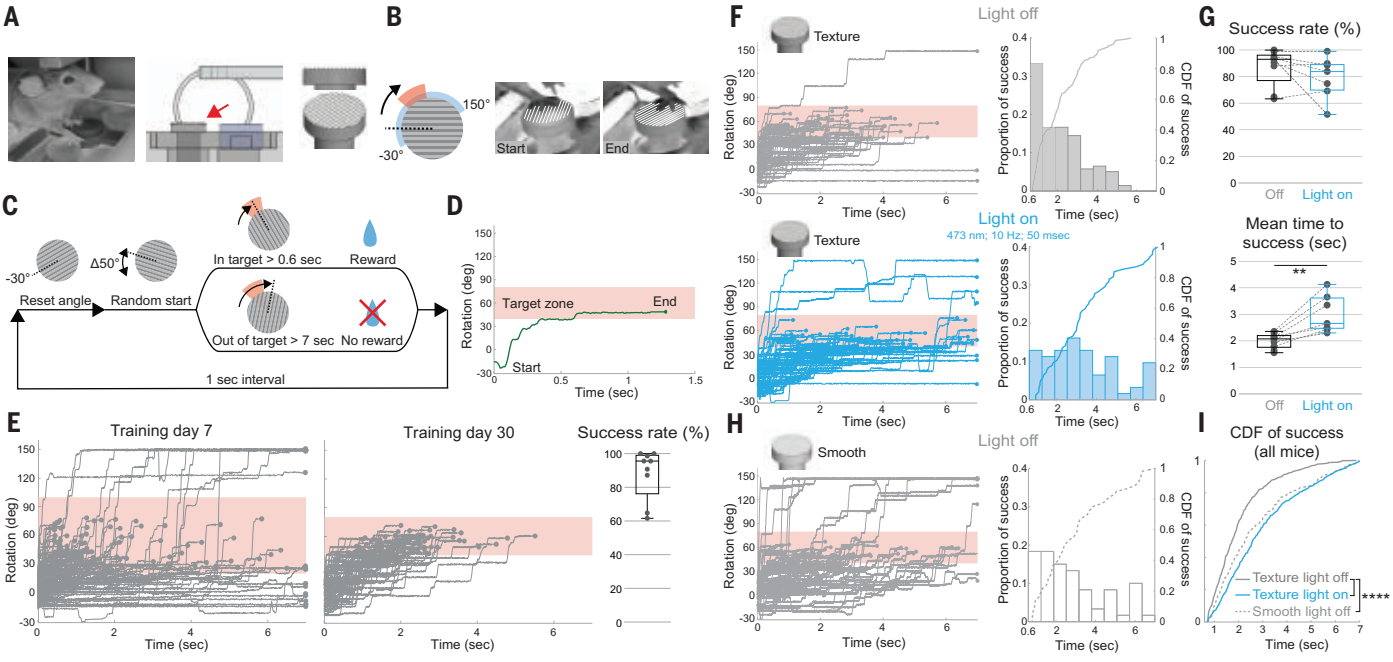



**Fig. 4. Activation of cuneate inhibitory circuits compromises tactile-dependent behavior.** (**A**) Head-fixed mouse performing the tactile orienting task (left). A pedestal (middle, red arrow) with parallel orientation ridges (right) was placed under the right hand. The pedestal was connected to a motor-rotary encoder assembly (fig. S9A). (**B**) Neutral angle of the ridges was defined as perpendicular to the body axis of the mouse (0°, dotted line), and the pedestal was passively mobile through a 180° range (blue, −30° to 150°). To receive water reward, the animal must turn the pedestal to align the ridges within a defined target range (red). (**C**) Task structure. After training, at the beginning of each trial, the pedestal was reset to −30° and then adjusted to a random starting orientation (0° ± 25°). The pedestal then became passively mobile, which indicated the start of the trial. Water reward was delivered if the orientation of the ridges stayed within the target zone (60° ± 20°) for >0.6 s. Otherwise, the trial was terminated without reward after 7 s. The next trial began after a 1-s interval, and trials continued until either 105 rewards were delivered or 160 trials were completed. (**D**) Example of angular trajectory of pedestal during a single trial. (**E**) Example angular trajectories during training. Animals began training with a small starting range and large target range (left). During training, the starting range became larger, and the target zone became smaller (middle). Mean success rate after training (nine mice) (right). (**F**) Ipsilateral Cu inhibitory neurons were targeted for optogenetic activation (as in Fig. 3A). Example angular trajectories (left; 70 trials) with light off (top; gray) and light on (bottom; blue) and plots (right) showing the proportion of successful trials at each time point (bars, 0.64 s bins) and the cumulative distribution function (CDF) for only the successful trials (lines). (**G**) During optogenetic activation (blue), the overall success rate was unaffected (top), but the mean elapsed time to achieve success increased (bottom; seven mice; $^{**}P = 0.0047$; two-way mixed-effects model with Geisser-Greenhouse correction and Sidak multiple comparisons test). (**H**) Example angular trajectories (70 trials), binned successes, and CDF for a mouse performing the task with a smooth surface and the light off. (**I**) CDF across all three conditions shows an equivalent drop in performance from control conditions (texture with light off, gray line, nine mice) when photoactivating Cu inhibitory neurons (blue line, seven mice) or with texture removed (dashed gray line, eight mice) ($^{****}P < 0.0001$; Kruskal-Wallis test with Dunn's multiple comparisons test) (fig. S9, B and C).

performance advantage and more frequently adopt incremental turn-and-wait strategies (Fig. 4, F to I, and fig. S9C).

### Top-down corticofugal control of cuneate circuits

In a final set of experiments, we investigated whether the modulatory circuits we identify can provide a basis for top-down control of tactile feedback. Ascending sensory transmission is attenuated before and during limb movement (*34*–*36*), potentially as a means to gate reafferent feedback caused by one's own movement or increase the signal-to-noise ratio of sensory information most relevant to task execution (*5*, *37*). The SSp innervates the cuneate nucleus, at least in part through collaterals of corticospinal projection neurons (*18*, *22*). Moreover, cortical activation can produce both excitatory and inhibitory effects in cat and monkey cuneate (*18*, *20*–*25*), which suggests that sensory cortex provides top-down regulation of its afferent inputs. We first set out to determine whether this modulation

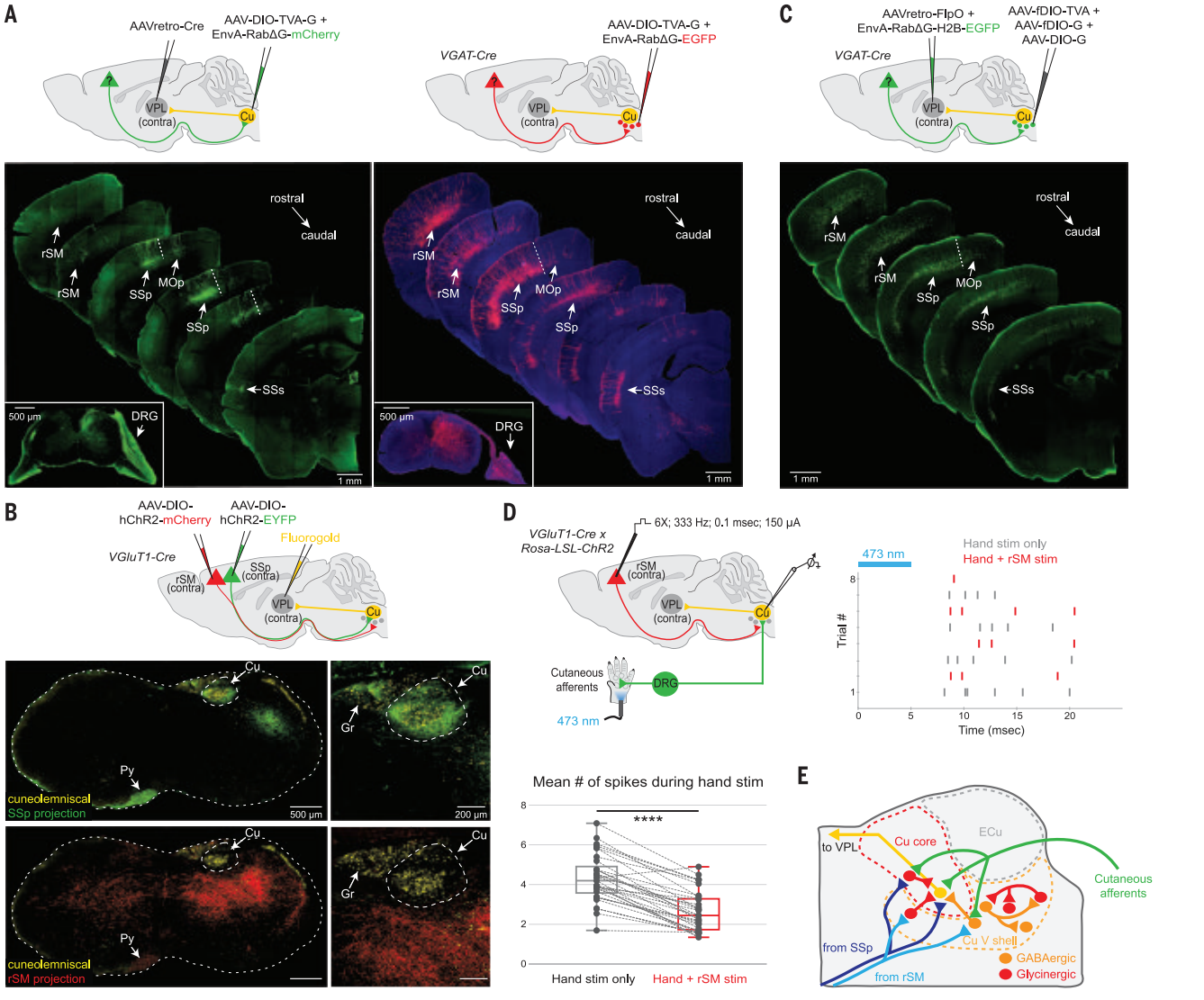



**Fig. 5. Distinct corticofugal pathways target cuneate circuits.** (**A**) (Left) Monosynaptic retrograde rabies tracing from cuneolemniscal (CL) neurons (three mice). Cortical inputs arise almost exclusively from contralateral SSp and SSs, with sparse or absent labeling in MOp or rSM. (Inset) Ipsilateral DRG neurons (C7 shown) were also labeled. (Right) Monosynaptic retrograde rabies tracing from Cu inhibitory neurons in *VGAT-Cre* mice (three mice). Cortical inputs arise from contralateral SSp and SSs but also include a large population of corticofugal neurons throughout rSM. (Inset) Ipsilateral DRG neurons (C7 shown) were also labeled. (Equivalent results with complementary viral approaches are provided in fig. S11.) (**B**) Dual viral anterograde tracing of corticofugal projections in the same *VGluT1-Cre* mice from contralateral SSp (green), innervating the core region of the Cu and Gr, and rSM (red), innervating more ventral brainstem, including the Cu ventral shell region. CL neurons retrogradely labeled with Fluorogold (yellow; two mice). Py, pyramidal tract. Viruses were switched across mice with no change in results (equivalent results with complementary viral approaches are provided in fig. S10). (**C**) (Left) Disynaptic retrograde rabies tracing (nuclear localized) from CL neurons, through presynaptic Cu inhibitory neurons, to corticofugal inputs (four mice). Cortical inputs arise from contralateral SSp, from SSs, and throughout rSM (control experiments are provided in fig. S11C). (**D**) In vivo extracellular recording of tactile-responsive neurons in Cu while stimulating rSM. Near-threshold tactile responses in Cu were elicited by photostimulating ChR2-expressing afferents in the pad of the ipsilateral hand of *VGluT1-Cre* × *Rosa-LSL-ChR2* mice (top left, 5 ms). Example spike raster plots (top right) from one recording site across interleaved trials show that tactile-evoked spikes in Cu are suppressed when rSM is electrically stimulated (red; six pulses, 333 Hz, 0.1-ms duration, 150 µA; first pulse 30 ms before tactile stimulation at time 0). Quantification across all recordings (bottom left; 33 recordings across three mice; 25-ms time window from hand stimulation; ****$P < 0.0001$; Wilcoxon two-tailed matched-pairs signed-rank test). (**E**) Schematic of local and long-distance Cu connections (fig. S12).

is exclusive to SSp or whether it might involve corticofugal neurons in other cortical regions. Using combinatorial genetic and viral tools, we broadly labeled cortical neurons that project to the cuneate, to cervical spinal cord, or to both. Corticospinal neurons throughout contralateral SSp send collateral projections to Cu, whereas corticospinal neurons in primary motor cortex (MOp) and secondary motor cortex mostly avoid the core region of Cu (fig. S10A). Targeted anterograde labeling of corticospinal neurons revealed that whereas SSp densely innervates the core region of Cu, sparse MOp projections are mostly found in ventral cuneate regions (fig. S10B) (*38*). We also found corticofugal neurons in a broad, contralateral anterior cortical region, which we refer to as rostral sensorimotor cortex (rSM), that project to the Cu region but do not appear to project to the spinal cord (fig. S10A).

To explore the synaptic connectivity of these corticofugal projection neurons more directly, we performed monosynaptic retrograde rabies tracing (Fig. 5A). First, we found that cuneolemniscal neurons are directly targeted by corticofugal neurons in SSp as well as from the cervical dorsal root ganglia. We also found cortical inputs arising from contralateral supplemental somatosensory cortex (SSs) but sparse or absent labeling in MOp or the more anterior rSM region (Fig. 5A, left). In addition to directly exciting cuneolemniscal neurons, some descending projections might also inhibit tactile transmission by recruiting the Cu inhibitory circuits that we identified. To test this idea, we restricted rabies tracing to Cu inhibitory neurons in *VGAT-Cre* mice (Fig. 5A, right). We found that cortical neurons in SSp and SSs also innervate these inhibitory circuits, providing a potential explanation for previous findings that sensory cortex can both excite and inhibit cuneate neurons. Cervical dorsal root ganglia neurons also target these inhibitory neurons, which suggests that sensory pathways can drive feed-forward inhibition of tactile circuits, potentially as a means for temporal and spatial modulation (*39*, *40*). We also found a much larger population of previously unidentified rSM neurons that directly innervate Cu inhibitory neurons but do not target cuneolemniscal neurons (Fig. 5A).

Supporting these findings, dual anterograde labeling of corticofugal neurons in SSp and rSM in *VGluT1-Cre* mice showed essentially nonoverlapping and complementary projection patterns; SSp heavily innervates the Cu core region, where cuneolemniscal neurons reside, whereas rSM innervates more ventral brainstem, including the Cu ventral shell region, where inhibitory neurons are located, but completely avoids the Cu core (Fig. 5B). Additional combinatorial retrograde viral targeting from the cervical spinal cord confirmed that SSp projections to the Cu core are collaterals of corticospinal projections (fig. S10C). rSM projections also descend through the pyramidal tract and decussate but, in contrast to SSp projections, instead innervate the Cu ventral shell region and do not send projections to forelimb regions of the spinal cord (fig. S10D). For a more comprehensive identification of regions that innervate Cu circuits, we used complementary rabies-tracing approaches with a nuclear-localized fluorophore and automated serial two-photon tomography to quantify neuronal populations throughout the brain that innervate either cuneolemniscal or Cu inhibitory neurons. This labeling confirmed our retrograde and anterograde anatomical findings and also identified several subcortical regions as candidate modulators of signaling in the dorsal column nuclei (tables S1 and S2 and fig. S11, A and B).

Corticofugal neurons in rSM project to a relatively large region of the brainstem (Fig. 5B and fig. S10D), which leaves open the possibility that the inhibitory neurons that they target (Fig. 5A, right) are not the same neurons that modulate tactile transmission within Cu. We therefore designed a disynaptic rabies-tracing approach to determine whether rSM corticofugal neurons innervate inhibitory neurons that modulate cuneolemniscal neurons. As a control, we initiated monosynaptic rabies tracing selectively from cuneolemniscal neurons that target VPL thalamus and found the expected inputs from SSp, cervical dorsal root ganglia, and local Cu inhibitory neurons (fig. S11C). In a separate group of mice, we again initiated monosynaptic tracing from cuneolemniscal neurons while introducing expression of supplemental rabies virus glycoprotein (G) selectively in Cu inhibitory neurons in *VGAT-Cre* mice, which thereby enabled rabies virus spread only through the inhibitory circuits that regulate the cuneolemniscal starter population. Using this approach, we found broad labeling in rSM as well as SSp, which demonstrates a feed-forward inhibitory link from these corticofugal populations to cuneolemniscal neurons (Fig. 5C).

We wanted to determine whether this rSM corticocuneate projection can indeed inhibit tactile feedback within Cu. Seeking greater temporal control over tactile stimulation, and motivated by work in cats that suggests that descending modulation of tactile feedback is most apparent when inputs are not saturating (*41*), we developed an approach in which we could more flexibly adjust tactile-stimulus timing and strength. Genetic expression of the excitatory opsin ChR2 in sensory afferents of *VGluT1-Cre* mice allowed us to activate cutaneous afferents by delivering brief light pulses to the glabrous pad of the hand. By adjusting the intensity of photostimulation, we could evoke near-threshold responses in Cu neurons (Fig. 5D). When the contralateral rSM was electrically stimulated just before hand afferent photoactivation, there was a consistent suppression of tactile-evoked spikes in Cu (38.56 ± 2.97% SEM reduction in spike number) (Fig. 5D). Together, these results demonstrate that rSM corticofugal neurons can elicit inhibitory modulation of tactile feedback and reveal a newly defined, top-down cortical pathway for regulating ascending somatosensory information (Fig. 5E and fig. S12).

## Discussion

Focusing on tactile signaling in the cuneate nucleus, we have established the functional connectivity of local GABAergic and glycinergic circuits that target cuneolemniscal neurons and enable bidirectional modulation of somatosensory feedback as it enters the brain. Perturbing the activity of these inhibitory circuits can suppress or enhance tactile responses in the cuneate, and a shift in the balance of ascending tactile transmission can severely affect the execution of behaviors that rely on sensory feedback from the hand. Distinct top-down neocortical circuits show complementary projection patterns that together provide circuit mechanisms for excitation or feed-forward inhibition of cuneolemniscal transmission to the thalamus (Fig. 5E and fig. S12). These results uncover previously uncharacterized anatomical and functional circuit architecture for the adjustment of tactile feedback that is critical for the execution of dexterous movements and provide insight into general circuit mechanisms, analogous to those identified for other sensory pathways (*6*–*12*, *42*), that can attenuate disruptive feedback to facilitate successful behavior.

Suppression of tactile signals in the cuneate would be beneficial for several reasons. Center-surround inhibition could sharpen receptive fields to augment resolution and acuity (*26*, *39*, *40*). In addition, feedback caused by movement should be attenuated to counteract disruptive feedback delays and to distinguish expected self-generated reafference from unexpected exafferent signals (*5*, *35*, *37*, *42*, *43*). Conversely, amplification of tactile feedback could in principle improve discrimination and select the inputs most relevant to a behavior (*5*, *24*, *26*, *37*). Moreover, excitation could convey predictions of upcoming events during movement, which would prepare sensory circuits to process impending feedback through anticipatory modulation (*2*, *13*, *18*). Given the utility of bidirectional modulation, attenuation and augmentation are likely to occur simultaneously, as is seen with the coincident suppression of cutaneous and enhancement of proprioceptive feedback in the spinal cord during wrist movement in monkeys (*10*).

The circuits described here could provide the means to regulate distinct channels of

tactile information as they ascend into the brain. Most of the corticofugal neurons we identified target inhibitory cuneate neurons, which suggests that widespread and perhaps somewhat indiscriminate inhibition is desirable, or at least necessary, for coordinated behavior (*36*, *42*). Indeed, broad suppression of somatosensory feedback and an increase in psychophysical detection thresholds before and during movement are well documented in several mammalian species (*34*–*36*, *43*). One might speculate that preparatory activity in corticofugal neurons, including those across the large region of rostral cortex that we describe, could mediate this movement-related suppression. By contrast, direct innervation of cuneolemniscal neurons was found almost exclusively from excitatory neurons in SSp, the cortical region that serves as the main recipient of cuneolemniscal signals and can directly influence motor output (*44*). Cortical connectivity to the cuneate might be somatotopically aligned (*17*, *19*, *27*), which suggests that in the face of broad suppression, specific cortical recipients might select and augment the feedback they receive to facilitate ongoing and future movements according to behavioral context (*13*, *24*, *26*, *45*, *46*). This phenomenon of global inhibition alongside selective excitation has been observed for odor representation in rat olfactory cortex (*47*) and may represent a more general strategy for processing sensory stimuli. The behavioral implications of cortical inhibition and excitation of cuneate signaling remain difficult to define, in part because the corticofugal pathways we describe collateralize to many targets as they descend through the pyramidal tract (*22*). Reliable approaches for perturbing specific axon collaterals while leaving other targets unaffected are needed to address whether broad attenuation punctuated by selective activation characterizes tactile processing in the cuneate during movement. Moreover, top-down modulation of cuneate might have widespread effects. Cuneate neurons target many regions, including the cerebellum, pontine nucleus, inferior olive, red nucleus, and superior colliculus (*16*, *17*). Whether these distinct cuneate outputs are differentially modulated remains unknown, but this scenario could provide added flexibility, which would allow cortical and subcortical areas to receive different versions of the same peripheral signals.

This cuneate circuit organization is consistent with a model in which top-down pathways convey sensory predictions of bottom-up sensory signals (*48*). One formulation of this predictive processing framework is that prediction error neurons sit at the interface of ascending and descending pathways and come in two flavors: positive prediction error neurons that signal unexpected inputs and negative prediction error neurons that respond to the absence of a predicted input (*49*). A putative circuit implementation takes the form of a positive prediction error neuron receiving bottom-up excitation and top-down feed-forward inhibition, whereas a negative prediction error neuron receives the reverse; when excitation exceeds inhibition in either class of neuron, a corresponding error signal is sent to update an internal model and modify future predictions (*49*). These types of error predictions could emerge at any, or every, layer of the sensorimotor hierarchy. At the first layer of tactile processing in the brain, we found bottom-up sensory and top-down cortical pathways forming all of the requisite excitatory and feed-forward inhibitory connections (Fig. 5E and fig. S12). A challenge will be to establish the resolution needed to determine whether each of the circuit elements match up in the appropriate combinations. Alongside advancing technologies for cuneate recording in behaving animals (*45*, *50*), these approaches should help to resolve whether the connectivity and recruitment of neurons in the dorsal column pathway are reflective of hierarchical predictive processing.

## REFERENCES AND NOTES


1. V. B. Mountcastle, *The Sensory Hand: Neural Mechanisms of Somatic Sensation* (Harvard Univ. Press, 2005).
2. R. S. Johansson, J. R. Flanagan, *Nat. Rev. Neurosci.* **10**, 345–359 (2009).
3. S. H. Scott, *Trends Neurosci.* **39**, 512–526 (2016).
4. R. Shadmehr, M. A. Smith, J. W. Krakauer, *Annu. Rev. Neurosci.* **33**, 89–108 (2010).
5. E. Azim, K. Seki, *Curr. Opin. Physiol.* **8**, 177–187 (2019).
6. A. M. Sillito, H. E. Jones, G. L. Gerstein, D. C. West, *Nature* **369**, 479–482 (1994).
7. C. D. Gilbert, M. Sigman, *Neuron* **54**, 677–696 (2007).
8. S. Lee, G. E. Carvell, D. J. Simons, *Nat. Neurosci.* **11**, 1430–1438 (2008).
9. A. J. Fink *et al.*, *Nature* **509**, 43–48 (2014).
10. J. Confais, G. Kim, S. Tomatsu, T. Takei, K. Seki, *J. Neurosci.* **37**, 2612–2626 (2017).
11. Y. Liu *et al.*, *Nature* **561**, 547–550 (2018).
12. D. M. Schneider, J. Sundararajan, R. Mooney, *Nature* **561**, 391–395 (2018).
13. P. D. Wall, *Brain* **93**, 505–524 (1970).
14. D. S. Glendinning, B. Y. Cooper, C. J. Vierck Jr., C. M. Leonard, *Somatosens. Mot. Res.* **9**, 61–73 (1992).
15. M. Ballermann, J. McKenna, I. Q. Whishaw, *Brain Res. Bull.* **54**, 237–242 (2001).
16. K. J. Berkley, R. J. Budell, A. Blomqvist, M. Bull, *Brain Res.* **396**, 199–225 (1986).
17. A. J. Loutit, R. M. Vickery, J. R. Potas, *J. Comp. Neurol.* **529**, 187–220 (2021).
18. H. G. Kuypers, J. D. Tuerk, *J. Anat.* **98**, 143–162 (1964).
19. S. Cheema, B. L. Whitsel, A. Rustioni, *Somatosens. Res.* **1**, 169–205 (1983).
20. S. J. Jabbur, A. L. Towe, *Science* **132**, 547–548 (1960).
21. P. Andersen, J. C. Eccles, T. Oshima, R. F. Schmidt, *J. Neurophysiol.* **27**, 1096–1116 (1964).
22. A. Rustioni, N. L. Hayes, *Exp. Brain Res.* **43**, 237–245 (1981).
23. J. D. Cole, G. Gordon, *Exp. Brain Res.* **90**, 384–392 (1992).
24. A. Canedo, J. Mariño, J. Aguilar, *Neuroscience* **97**, 317–334 (2000).
25. F. Harris, S. J. Jabbur, R. W. Morse, A. L. Towe, *Nature* **208**, 1215–1216 (1965).
26. A. Canedo, *Prog. Neurobiol.* **51**, 287–335 (1997).
27. J. Aguilar, C. Rivadulla, C. Soto, A. Canedo, *J. Neurophysiol.* **89**, 3328–3339 (2003).
28. A. Rustioni, D. E. Schmechel, S. Cheema, D. Fitzpatrick, *Somatosens. Res.* **1**, 329–357 (1984).
29. A. Popratiloff, J. G. Valtschanoff, A. Rustioni, R. J. Weinberg, *Brain Res.* **706**, 308–312 (1996).
30. J. H. Lue, Y. F. Jiang-Shieh, J. Y. Shieh, E. A. Ling, C. Y. Wen, *Neurosci. Res.* **27**, 123–132 (1997).
31. F. Bengtsson, R. Brasselet, R. S. Johansson, A. Arleo, H. Jörntell, *PLOS ONE* **8**, e56630 (2013).
32. A. A. Blackwell, M. T. Banovetz, I. Q. Qandeel, I. Q. Whishaw, D. G. Wallace, *Behav. Brain Res.* **345**, 49–58 (2018).
33. J. A. Pruszynski, J. R. Flanagan, R. S. Johansson, *eLife* **7**, e31200 (2018).
34. C. Ghez, M. Pisa, *Brain Res.* **40**, 145–155 (1972).
35. J. D. Coulter, *J. Neurophysiol.* **37**, 831–845 (1974).
36. J. K. Chapin, D. J. Woodward, *Exp. Neurol.* **72**, 164–178 (1981).
37. C. E. Chapman, *Can. J. Physiol. Pharmacol.* **72**, 558–570 (1994).
38. H. C. Shin, J. K. Chapin, *Brain Res. Bull.* **22**, 245–252 (1989).
39. C. Soto, J. Aguilar, F. Martín-Cora, C. Rivadulla, A. Canedo, *Eur. J. Neurosci.* **19**, 3006–3016 (2004).
40. C. L. Witham, S. N. Baker, *J. Neurophysiol.* **106**, 2764–2775 (2011).
41. A. L. Towe, S. J. Jabbur, *J. Neurophysiol.* **24**, 488–498 (1961).
42. A. J. McComas, *J. Neurophysiol.* **116**, 41–50 (2016).
43. S. J. Blakemore, C. D. Frith, D. M. Wolpert, *J. Cogn. Neurosci.* **11**, 551–559 (1999).
44. S. K. Karadimas *et al.*, *Nat. Neurosci.* **23**, 75–84 (2020).
45. C. Versteeg, J. M. Rosenow, S. J. Bensmaia, L. E. Miller, *J. Neurophysiol.* **126**, 693–706 (2021).
46. W. Omlor *et al.*, *Nat. Commun.* **10**, 4812 (2019).
47. C. Poo, J. S. Isaacson, *Neuron* **62**, 850–861 (2009).
48. R. A. Adams, S. Shipp, K. J. Friston, *Brain Struct. Funct.* **218**, 611–643 (2013).
49. G. B. Keller, T. D. Mrsic-Flogel, *Neuron* **100**, 424–435 (2018).
50. A. K. Suresh *et al.*, *J. Neurophysiol.* **118**, 3271–3281 (2017).
51. E. Azim, azimlabsalk/Tactile-Orienting-Task: Tactile orienting assay for mice, Zenodo (2021); doi:10.5281/zenodo.5218877.


## ACKNOWLEDGMENTS

We are grateful to P. Nguyen and G. Salmun (Salk Institute) for assistance with mouse husbandry, histology, behavioral assays, and laboratory operations; H. Gao and A. Thanawalla (Salk Institute) for assistance with string-pull experiments; B. Lim (University of California, San Diego) for the AAV-DIO-mRuby-T2A-Syp-EGFP plasmid; T. Margrie (Sainsbury Wellcome Centre) for the AAV-fDIO-N2cG-H2B-GFP plasmid; A. Cetin (Allen Institute for Brain Science) for the EnvA-RabΔG-H2B-EGFP plasmid; A. Hantman (Janelia Research Campus) for the *VGluT1-Cre*, *GAD2-Flp*, and *GlyT2-Flp* mouse lines; F. Cardone, M. Stambaugh, J. Sandubrae (Qualcomm Institute, University of California, San Diego), and D. Butler (Salk Institute) for assistance developing the tactile orienting assay; and D. Ramirez (Whole Brain Microscopy Facility, University of Texas Southwestern Medical Center) for assistance with serial two-photon tomography and quantification. We thank S. Aoki, S. Bensmaia, A. Fink, K. W. Huang, D. Jabaudon, K. Seki, J. Tuthill, M. Yartsev, and members of the Azim laboratory for valuable discussion and comments on the manuscript. **Funding:** This work was supported by a Salk Pioneer Fund Postdoctoral Scholar Award (to A.B.); Uehara Memorial Foundation Postdoctoral Fellowship (to M.I.); and National Institutes of Health (R00NS088193, DP2NS105555, R01NS111479, and U19NS112959), the Searle Scholars Program, The Pew Charitable Trusts, and the McKnight Foundation (to E.A.). **Author contributions:** Conceptualization: J.M.C. and E.A.; experimental design: J.M.C., A.B., M.I., and E.A.; viral injections, anatomical experiments, and in vivo electrophysiological recordings: J.M.C.; slice electrophysiology: A.B.; string-pulling behavior: N.B. and J.T.; development and implementation of tactile orienting assay: M.I., A.B., and E.A.; data analysis: J.M.C., A.B., M.I., J.T., N.B., and E.A.; writing and figures: E.A.; manuscript review and editing: J.M.C., A.B., M.I., J.T., N.B., and E.A. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. Design and code for the tactile orienting assay are available at (*51*).

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abh1123
Materials and Methods
Figs. S1 to S12
References (*52*–*77*)
Tables S1 and S2
MDAR Reproducibility Checklist
Movies S1 to S4

19 February 2021; accepted 19 August 2021
10.1126/science.abh1123

HEALTH ECONOMICS

# Heuristics in the delivery room

Manasvini Singh*

**Clinical decisions made in the delivery setting are often made under high pressure and great uncertainty, and have serious consequences for mother and baby. Theories of decision-making suggest that individuals in such settings may resort to using heuristics, or simplified decision rules, to aid complex decision-making. This study investigates whether physicians' delivery mode decisions (i.e., when to perform a vaginal versus a cesarean delivery) are influenced by such a heuristic. Electronic health records spanning 86,000 deliveries suggest that if the prior patient had complications in one delivery mode, the physician will be more likely to switch to the other—and likely inappropriate—delivery mode for the subsequent patient, regardless of patient indications. There is evidence that this heuristic has small, suboptimal effects on patient health.**

In 2014, the United States was awarded a grim distinction: It achieved the second-highest maternal mortality rate out of 31 OECD countries (*1*). In fact, even as most countries (157 of 183 countries in a WHO report) (*2*) are improving maternal health indicators, the rate of maternal mortality and morbidity in the US has been steadily increasing in the past few decades. Pregnancy-related mortality increased 140% from 1987 to 2017 (*3*), and severe maternal morbidity increased 200% from 1993 to 2014 (*4*). Given these alarming statistics, it is critical to understand how physicians make delivery decisions and identify ways to improve patient health outcomes.

The delivery setting is a particularly high-stakes context to study decision-making. Physicians' delivery decisions can severely and irrevocably affect the health of two patients (mother and baby), there is a lack of high-quality evidence to guide the central physician decision (i.e., when to perform a cesarean or a vaginal delivery), and physicians experience considerable cognitive and time pressures. Theories of decision-making suggest that these features, taken together, create an environment where individuals may be more likely to rely on heuristics, or simplified decision rules, to aid complex decision-making.

In this study, I investigate whether labor and delivery complications that occurred in a physician's prior patient affect the physician's delivery decisions for their next patient in predictable ways. There is no clinical reason why the delivery decisions for two separate patients—linked only by the accidental chance of being seen consecutively by the physician—should be causally related to each other. However, when faced with the complex decision of deciding whether the current patient is suited for a vaginal or a cesarean delivery, physicians may instead be influenced—suboptimally—by the outcome of the decision they made for their prior patient. Specifically, I test whether complications in the prior patient's delivery mode (whether a cesarean or vaginal delivery) make the physician more likely to switch to the other delivery mode for their next patient, and the effect of this heuristic on the patient for whom it is used.

Several lines of research in decision-making predict that this behavior may occur. This behavior can be, for example, loosely predicted by a "win-stay/lose-shift" heuristic, according to which the decision-maker either switches strategies if the last outcome was a "loss" or continues with the same strategy if the last outcome was a "win." The use of this heuristic has been observed in animals (*5*), lay individuals (*6*, *7*), and experts in domains such as medicine. For instance, physicians appear to "shift" strategies after a "loss": Incurring a complication due to a drug reduces usage of the drug (*8*, *9*); having malpractice concerns or missed diagnoses subsequently increases clinical testing (*10*–*12*). Of course, these patterns can be also predicted by other closely related theories of heuristics and decision-making, such as availability (*13*, *14*), salience (*15*), recency (*16*), and attribute substitution (*17*, *18*). Although identifying the precise mechanism of decision-making is outside the scope of this study, of particular interest here is the use of this general type of heuristic—switching strategies after a "loss"—by experts, and its real-world consequences.

In the delivery setting examined in this study, there are three reasons why the use of simplified heuristics by physicians is especially concerning. First, the serious and long-term effects of inappropriate delivery mode choices on the health of mother and child are well documented (*19*–*21*), making the use of heuristics especially risky. Several global campaigns have even been launched to make delivery decisions more evidence-based (*22*, *23*). Second, switching delivery modes after a complication (i.e., "loss") does not offer any guaranteed benefits, making it a potentially flawed

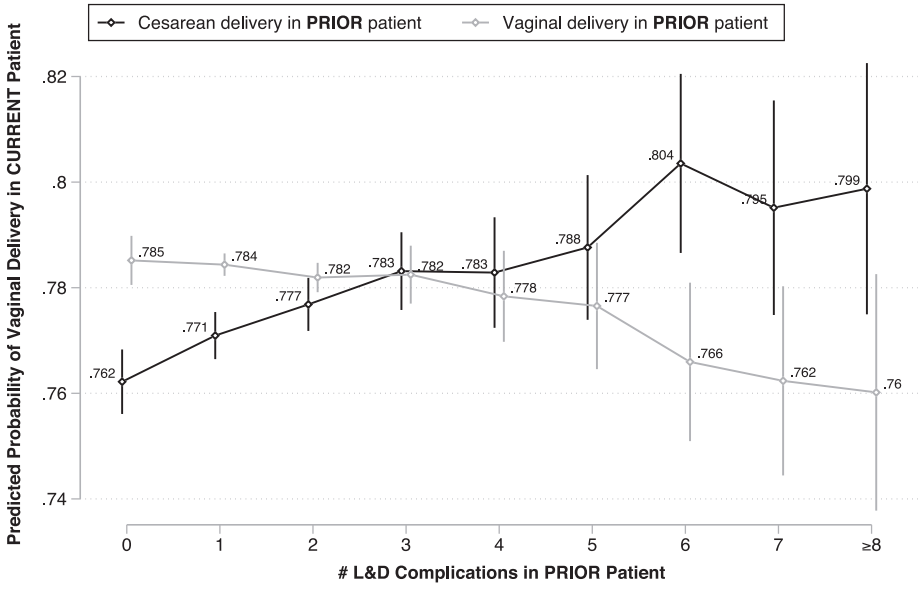



**Fig. 1. Relationship between number of labor and delivery complications in the prior patient and the delivery mode chosen for the current patient.** This figure plots estimates from table S6, column 2. Estimates are predicted probabilities of the vaginal delivery mode on the current patient (excluding mothers with suspected pre-scheduled cesareans). Delivery mode is coded as vaginal delivery = 1 and cesarean = 0. Probabilities of cesareans can be calculated by subtracting estimates from 1. Confidence intervals represent statistical significance at the 5% level. Example interpretation: When the prior patient has a cesarean delivery (black line) with six complications, the current patient has an 80.4% chance of delivering vaginally, or equivalently, a 19.6% chance of cesarean delivery.

College of Social and Behavioral Sciences, University of Massachusetts, Amherst, MA, USA.
*Corresponding author. Email: manasvinisin@umass.edu

rule. For example, increasing testing after a missed diagnosis always decreases the chances of another missed diagnosis. By contrast, there is no evidence that switching delivery modes after a complication avoids further complications, especially because labor and delivery complications occur similarly in both delivery modes (fig. S1). Finally, patients are usually more averse to having inappropriate procedures performed on them than they are to receiving unnecessary tests, which suggests that we should hold greater reservations about the use of heuristics in this setting.

### Data and analytic strategy

Labor and delivery (L&D) is an attractive setting to study heuristic use not just because of its ramifications for patient health, but also because of its tractability. In this setting, physicians have only two mutually exclusive and exhaustive choices: They must deliver the baby either vaginally or by cesarean section.

I used the universe of inpatient electronic health record (EHR) data across 21 years (January 2000 through September 2020) at the obstetric ward of two academic hospitals (one large and urban, the other small and suburban). These EHR data are managed by a clinical data warehouse that maintains data standards for research, patient safety, and quality improvement. These data are richly detailed and offer several advantages over most clinical administrative datasets widely available or used, such as providing time stamps as patients move through hospital wings (allowing me to queue them in the order seen by the physician) and identifying the physician who actually delivered the baby (*24*). Overall, the data cover 86,345 deliveries by 231 physicians; physicians performed an average of 389.8 deliveries (SD = 691.5).

The primary dependent variable is the physician's current patient's delivery mode: vaginal or cesarean. The two primary independent variables are (i) the physician's prior patient's delivery mode and (ii) complications during the prior patient's delivery. Some examples of complications are obstructed labor, postpartum hemorrhage, fetal distress, perineal laceration, umbilical cord complications, and obstetric trauma (table S1). Delivery mode and L&D complication are identified using ICD (International Classification of Diseases) procedure and diagnoses codes. See (*24*) for a discussion of the implications of these coding choices on hypotheses regarding delivery circumstances such as emergency cesareans.

Control variables for patient characteristics are age, race, ethnicity, insurance status, number of newborns delivered during encounter, and Elixhauser readmission and mortality scores, which are two validated measures of patient severity based on weighted combinations of 29 different comorbidities (*25*). As

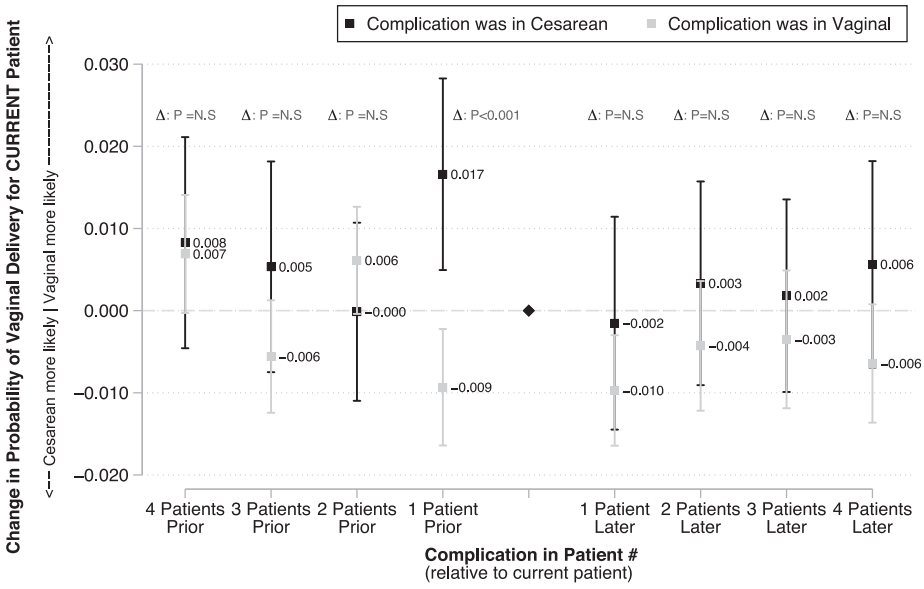



**Fig. 2. Effect of complications in up to four patients prior, and up to four patients later, on the delivery mode chosen for the current patient.** This figure plots the effect of complications in (i) prior and later cesareans (black lines), and (ii) prior and later vaginal deliveries (gray lines), on the delivery mode chosen for the current patient (coded as 1 = vaginal delivery, 0 = cesarean delivery). The gray dashed line represents a null effect of complications on the probability of vaginal delivery for the current patient. The *P* value of the difference (represented by Δ) between the two effects is shown for all patients. Confidence intervals represent statistical significance at the 5% level. Example interpretation: A complication in a cesarean delivery one patient prior increases the current patient's likelihood of a vaginal delivery by 1.7 percentage points, and a complication in a vaginal delivery one patient prior decreases her likelihood of a vaginal delivery (equivalently, increases her likelihood of a cesarean) by 0.9 percentage points. The *P* value of the difference between the two effects (Δ) is <0.001, whereas for all other prior and later patients, this difference is statistically not significant (N.S).

it is not possible to definitively identify prescheduled deliveries using procedure codes, I also included an indicator variable for suspected pre-scheduled deliveries by using suggestive diagnosis codes and free-form admit notes (*24*). Table 1 summarizes patient characteristics and L&D complications in cesarean and vaginal deliveries, and table S2 has a more detailed breakdown of the Elixhauser comorbidities. Complications are slightly more common in cesareans (mean of 1.44 in cesareans and 1.39 in vaginal deliveries), perhaps because cesareans are performed on more comorbid and older patients. However, it is critical to note that any differences in patients and complications between vaginal and cesarean deliveries are inconsequential to this study, because all comparisons are made within (not between) delivery modes. Standard errors in all analyses are clustered at the physician level.

Identification of causal estimates depends on one central assumption: that the occurrence of complications in the prior patient is uncorrelated with the delivery mode of the subsequent patient for reasons other than physician judgment, conditional on observed covariates. Neither the order of patients nor the order of delivery modes need be entirely random for identification. I found support for the validity of this assumption in four ways. First, I showed that there is no difference in observed characteristics between patients who arrive after complications in the prior delivery and those who arrive after no complications in the prior delivery (table S3). Second, in all analyses, I limited all comparisons to within a delivery mode. For example, I compared the delivery mode a physician chooses after experiencing a complication in a prior cesarean to the delivery mode the same physician chooses after experiencing no complications in a prior cesarean. Third, in all analyses, I used a host of fixed effects—physician, hospital, year, month, hour of day, and day of week—to control for differences across physicians, hospitals, and time (e.g., elective cesareans are more likely in the morning, some physicians treat more difficult patients, etc.). Finally, I performed a factor analysis on patient characteristics and found no serial correlation in factors between consecutive patients (*24*) (tables S4 and S5), which suggests that patient order is independent of

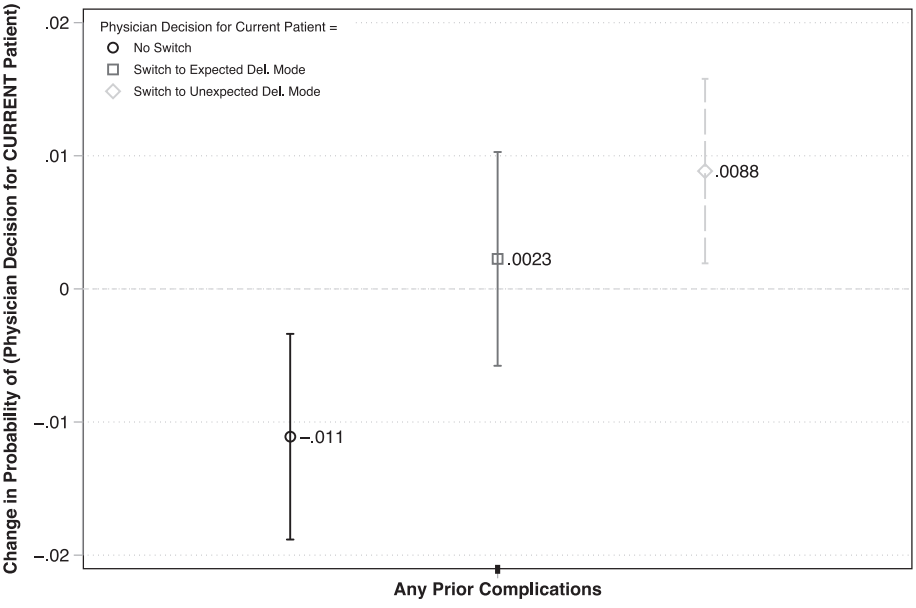



**Fig. 3. Effect of complications in the prior patient's delivery on the "unexpectedness" of the delivery mode chosen for the next patient.** Estimates represent changes in the probability of the Physician Decision for Current Patient with an increase of 1 (or more) prior complications. Physician Decision for Current Patient is coded as 0 = "No switch in delivery mode" from prior to current patient (black line), 1 = "Switch to Expected delivery mode" (dark gray line), or 2 = "Switch to Unexpected delivery mode" (light gray line). Confidence intervals represent statistical significance at the 5% level (*24*). Estimates are disaggregated by prior delivery mode in fig. S2. Example interpretation: When there are complications in a prior delivery, the physician is 0.88 percentage points significantly more likely to switch to the Unexpected delivery mode from the prior patient to the current patient (light gray line).

patient characteristics. Overall, this approach is supported by the literature as well: Other quasi-experimental studies have also exploited the random order of deliveries (*26*), and it is unlikely that physicians select patients on the basis of perceived delivery complication risk because even sophisticated algorithms can only predict about 50% of such complications (*27*).

This study consists of four analyses, where I tested whether complications cause physicians to make predictable deviations in subsequent delivery mode choices (Analysis 1), the extent to which these choices can be identified as "inappropriate" (Analysis 2), these decisions' effects on patient outcomes (Analysis 3), and whether these deviations can plausibly be thought of as a physician learning strategy (Analysis 4). These analyses were then followed by 16 theory-driven robustness checks and 2560 specification-driven sensitivity tests. In all analyses, a physician's prior patient is referred to as patient ($i$ – 1), and the same physician's subsequent patient (i.e., the "current patient") is referred to as patient $i$.

In Analysis 1, I examine whether L&D complications in the prior patient's delivery make the physician more likely to "switch" to the other delivery mode for the next patient—a behavior I refer to as "physician switching of delivery modes." This term refers exclusively to the act of using a different delivery mode for the next patient than was used for the prior patient, not the act of switching delivery modes within a single patient's delivery (such as when a physician performs an emergency cesarean during an attempted vaginal delivery). It is critical to clarify that when I identify an occurrence of a physician switching delivery modes from patient ($i$ – 1) to patient $i$ in the data, I am not suggesting that it is a conscious decision made by the physician. It is simply a term used to capture when a physician has observably "shifted" their strategy from the prior patient to the subsequent one, possibly as a consequence of experiencing a prior "loss" (i.e., a labor complication). Two stylized examples may be illuminating here. First, imagine that a patient's vaginal delivery incurs an L&D complication. The physician's next patient now arrives for a vaginal delivery. I hypothesize that because of complications in the physician's prior vaginal delivery, the physician will have a lower threshold (and therefore higher inclination) for deploying an emergency c-section during this next patient's vaginal delivery, even if a cesarean is not clinically indicated for that patient. As a result, this next patient will be more likely to deliver via cesarean section.

Now, imagine the converse example. The physician's prior patient has a cesarean that incurs complications. Assume the next patient comes in for a vaginal delivery. I now hypothesize that because of complications in the physician's prior cesarean delivery, the physician has a higher threshold (and therefore lower inclination) for deploying an emergency c-section during the vaginal delivery, even if a cesarean is clinically indicated for that patient. As a result, this next patient will be more likely to deliver vaginally.

In both examples, the final delivery mode used for this next patient will register in the data as a "switch" in delivery modes from the physician's prior patient. Of course, it is impossible to observe the specific mechanics of this process using these data, such as the patient's clinically appropriate and/or original delivery plan. However, if physicians are consistently using this heuristic even on the margin, we should still observe that delivery mode "switches" are more common after complications than after no complications. Thus, intuitively speaking, Analysis 1 (*24*) examines whether a physician's patterns of delivery mode choices deviate predictably in tandem with prior complications—similar to looking at a streak of coin flips and determining whether the draw of "heads" can be predicted from previous coin flips (which, if there is no heuristic use, should not be possible). Specifically, in table S6, I estimate the within-physician probability of switching delivery modes after experiencing complications in the prior patient, using a variety of specifications: a continuous, categorical (0, 1, 2+), and binary measure of prior complications in a linear probability model (columns 1, 3, and 4), a logistic regression (column 5), and excluding patients with suspected pre-scheduled deliveries (column 2).

Estimates are in line with my hypotheses and robust to all specifications (table S6). Depending on the specification, complications in a prior cesarean delivery make the physician 0.6 to 2.4 percentage points more likely to switch to a vaginal delivery for the next patient (representing an increase of up to 3.4% in the probability of a vaginal delivery, $P = 0.01$). Conversely, complications in a prior vaginal delivery make the physician 0.1 to 1.1 percentage points more likely to switch to a cesarean for the next patient (representing an increase of up to 3.6% in the probability of a cesarean delivery, $P = 0.045$). The difference between the two effects (i.e., the coefficient on the interaction) is significant at the 5% level. In sensitivity analyses described later, I show that these estimates are not sensitive to covariate combinations and are not driven by patients with an unusual number of complications.

Figure 1 plots the results from column 2 of table S6, as excluding patients with suspected

pre-scheduled deliveries should theoretically yield the strongest evidence of heuristic use (although including a control for such deliveries using the whole sample gives similar results with a higher $R^2$; see column 1). The figure clearly shows that the probability of switching delivery modes increases as the number of complications in the prior delivery increases. Interestingly—although comparisons between delivery modes should be taken with a grain of salt—we see that when there are 0 prior complications, physicians are more likely to "stay" with the delivery mode they used for the prior patient. For example, when a prior vaginal delivery has no complications, physicians have a 78.5% chance of performing a vaginal delivery again on the next patient, but when a prior cesarean has no complications, this likelihood decreases to 76.2% ($\beta_2$ in table S6, $P < 0.001$).

I then examine whether patients other than the prior patient similarly affect the current patient's delivery mode to assess how long the effects of complications on physician decisions persist. In the same model, I regress the current patient's delivery mode on an interaction between complications and delivery mode for up to four patients prior and up to four patients after (*24*). Figure 2 plots the marginal effects from this model. The effects of the immediately preceding patient stand out. Specifically, the prior patient is the only case in which (i) complications in the prior patient's cesarean increase the current patient's likelihood of delivering vaginally (by 1.7 percentage points, an increase of 2.4% in the probability of vaginal delivery, $P = 0.005$); (ii) complications in the prior patient's vaginal delivery increase the current patient's likelihood of having a cesarean (by 0.9 percentage points, an increase of 2.9% in the probability of a cesarean delivery, $P = 0.01$), and (iii) the difference between these two effects is significant ($P < 0.001$).

Having established in Analysis 1 that complications in the prior patient's delivery make the physician more likely to switch to the other delivery mode on the next patient (i.e., on patient *i*), Analysis 2 now examines whether the delivery mode the physician switched to for patient *i* is inappropriate for patient *i*. Because the appropriateness of delivery modes cannot be objectively measured, I instead test whether prior complications make the physician more likely to switch to a delivery mode that diverges from what is predicted—or "Expected"—for that patient on the basis of observable characteristics. I classify delivery modes as "Expected" (i.e., potentially appropriate) and "Unexpected" (i.e., potentially inappropriate) using the predicted delivery mode from a logistic regression of delivery mode on observable patient characteristics, which accurately classifies 70% of deliveries in this dataset. If the predicted delivery mode for a patient is the same as her observed delivery mode, the observed delivery mode is classified as the Expected choice, and Unexpected otherwise. I run an ordered multinomial logistic regression where the dependent variable—Physician Decision for Current Patient—has three categories: (i) no switch in delivery modes from prior to current patient, (ii) switch to the Expected delivery mode from prior to current patient, or (iii) switch to the Unexpected delivery mode from prior to current patient. The independent variables are prior patient's delivery mode interacted with a binary measure of prior complications (0/1+). [See (*24*) for a further explanation of variable construction and regression models.] Analysis 2 sidesteps some econometric concerns regarding Analysis 1, namely the potential for spurious state dependence (*28*) or biased selection procedures (*29*) from regressing $Y_i$ on $Y_{(i-1)}$. These concerns are most likely unfounded, most prominently because I never focus on the effect of prior delivery mode on current delivery mode; rather, I report the effect of prior complications on current delivery mode, conditional on prior delivery mode. Regardless, to alleviate these concerns, Analysis 2 examines the same question using a different specification that does not include a lagged outcome as a regressor.

My hypothesis suggests that prior complications should have no effect on "switching to the Expected delivery mode" (because a delivery mode that is Expected is driven solely by patient characteristics) but a positive effect on "switching to the Unexpected delivery mode" (because a delivery mode that is Unexpected is driven by factors other than patient characteristics, such as prior complications). The results presented in Fig. 3 support these hypotheses: Complications in the prior delivery mode reduce the likelihood that the physician does not switch from (or equivalently, that the physician "stays" with) the prior delivery mode by 1.1 percentage points ($P = 0.005$), have a null effect on the physician switching to the Expected delivery mode ($P = 0.58$), and increase the likelihood that the physician switches to the Unexpected delivery mode by 0.09 percentage points ($P = 0.012$) for the current patient. Figure S2 decomposes these effects by prior patient's delivery mode to show that physicians may react more strongly to complications in a prior cesarean than to those in a prior vaginal delivery. Theoretically, this makes sense: Performing a cesarean has a larger component of discretion than a vaginal delivery, which may thus induce a stronger response to complications in a cesarean. However, this imbalance may also be because my measure of delivery mode appropriateness is crude and/or differentially predictive for vaginal versus cesarean deliveries. Regardless, consistent with heuristic use, not only do complications in the prior patient cause

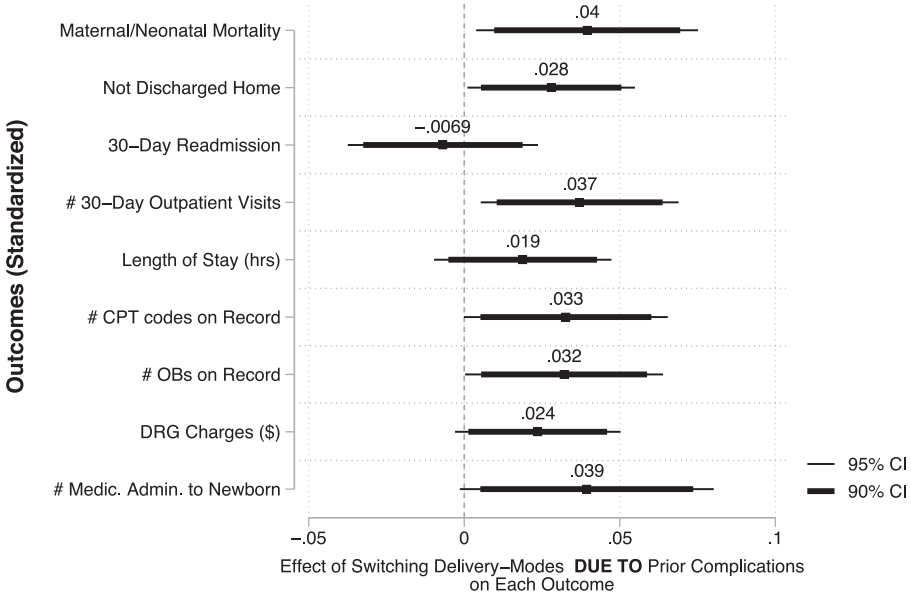



**Fig. 4. Effect of switching delivery modes due to prior patient's complications on current patient health outcomes and service use.** Each estimate plots the difference between (i) the effect of switching delivery modes on the outcome when the prior patient has had complications and (ii) the effect of switching delivery modes on the outcome when the prior patient has had no complications, thereby isolating the effect of switching delivery modes due to prior complications on the outcome. Confidence intervals represent statistical significance at the 10% and 5% levels (*24*). Example interpretation: Switching delivery modes after complications is associated with the current patient's likelihood of death increasing by 0.04 standard deviations more than when switching delivery modes after no complications.

physicians to switch delivery modes, they also cause physicians to deviate from the delivery mode choice predicted for the subsequent patient. Even without further evidence of patient harm, such deviations in delivery mode choices should be concerning given the long-term harm that inappropriate obstetric choices cause mother and baby (*19*–*21*).

The results from Analyses 1 and 2 thus far demonstrate that physician decision-making may be compromised. In Analysis 3 (*24*), I examine whether inappropriate switching for the current patient $i$—due to complications in the prior patient ($i$ − 1)—has downstream health consequences for patient $i$. Although prior complications may directly affect patient health through use of inappropriate care, they may also indirectly affect it if they spur physicians to act more cautiously and increase resource use. Thus, controlling for current patient's delivery mode, I regress an array of standardized current patient outcomes and measures of resource use on (i) an indicator for physician switching delivery modes from the prior patient to the current patient, (ii) an indicator for complications in the prior delivery, and (iii) an interaction between (i) and (ii). Any factors that cause a physician to switch delivery modes for reasons other than prior complications—such as random chance, specific sequences of patients, or relative proportion of cesareans to vaginals—should be similarly distributed between patients seen after complications and patients seen after no complications. Thus, the coefficient on the interaction—which is the estimate plotted for each (standardized) outcome in Fig. 4 and shown in column 4 of table S7—isolates the effect of switching due to prior complications by subtracting out these extraneous factors. Note that this estimate is not comparing whether current patient $i$ had worse patient outcomes than the prior patient ($i$ − 1); instead, it compares health outcomes between different current patients, some who were seen after prior complications and others who were not.

Figure 4 and table S7 show that relative to switching delivery modes after no prior complications, switching delivery modes after complications is associated with significantly worse health outcomes (e.g., a 0.04 SD increase in maternal and neonatal mortality, a 0.03 SD decrease in home discharge, and a 0.04 SD increase in the number of 30-day outpatient visits). Measures of service use also show modest increases, such as number of Current Procedural Terminology (CPT) procedure codes on record, diagnosis-related group (DRG) charges, number of medications administered to the newborn, and number of obstetricians on the encounter record (which may serve as a proxy for number of consultations). Although the effect on each individual health outcome is small and often only marginally significant, switching due to prior complications results in a 0.24 SD increase in the composite sum of all nine standardized outcomes ($P < 0.001$; table S7). Results are robust to the logit specification for binary outcomes (table S8). I also disaggregate effects by current patient's delivery mode (table S9) to investigate which delivery mode is driving these results; it appears both are, to some extent. Overall, switching delivery modes as a result of prior complications appears to have downstream and observable negative effects on patient health. However, results from Analysis 3 should be interpreted with caution; some of the outcomes examined are rare, and the estimates plotted in Fig. 4 are averaged across vaginal and cesarean deliveries. As such, care should be taken when extrapolating these estimates to other settings.

One could argue that this heuristic may still be a physician learning strategy—for instance, if physicians are trading off short-term harm to the patient for long-term improvements to their own performance. In Analysis 4, I perform two sub-analyses showing that the switching responses outlined thus far are not consistent with even basic models of physician learning (*24*). First, I show that, contrary to expectation, the more experienced the physician is (in terms of both years practiced and volume of patients seen), the more likely they are to switch delivery modes after complications ($P$ = 0.04 to 0.2; table S10). Second, I show that the physicians who tend to switch more after complications experience similar or worse trends in patient outcomes over time (e.g., increasing risk of death over time, $P$ = 0.01 to 0.1; no significant difference in 30-day readmission over time; table S11) than those who tend to switch less. All in all, patterns of physician decisions show predictable deviations after complications in ways that harm patients in the short and long run, as well as contradict even basic predictions from a learning model (e.g., effects persist for only one patient; physicians are more likely to switch delivery mode as they accumulate experience). Thus, switching behavior may be more likely to be evidence of physician use of a flawed heuristic than an optimal learning strategy.

I performed 16 theory-driven robustness checks (table S12, RC1 to RC16, most notably testing effects of potential fatigue, length of shifts, outliers in number of complications, severity of complications, etc.) and 2560 different regressions to create three specification curves (figs. S3 to S5) to strengthen the validity of these results (*24*). Not a single estimate in these robustness checks or specification curves compromises the main results, suggesting remarkable robustness of these estimates to alternative hypotheses and covariate balance.

**Table 1. Summary statistics: Vaginal versus cesarean deliveries.** Values in parentheses are standard deviations.

| | Mode of delivery: Cesarean | Mode of delivery: Vaginal | Total |
|---|---|---|---|
| Elixhauser In-Hospital Mortality Score (mean) | 2.00 (6.42) | 1.59 (5.19) | 1.71 (5.60) |
| Elixhauser Readmission Score (mean) | 4.93 (10.47) | 3.61 (8.29) | 4.01 (9.03) |
| Uninsured (%) | 1.56 (12.41) | 1.84 (13.43) | 1.75 (13.13) |
| Public insurance (%) | 50.06 (50.00) | 54.28 (49.82) | 53.00 (49.91) |
| Number of births during delivery | 1.05 (0.22) | 1.01 (0.10) | 1.02 (0.15) |
| Age (years) | 30.22 (6.15) | 28.31 (6.00) | 28.89 (6.11) |
| White (%) | 15.49 (36.18) | 18.09 (38.49) | 17.30 (37.83) |
| Hispanic (%) | 2.50 (15.62) | 2.57 (15.83) | 2.55 (15.76) |
| Black (%) | 71.17 (45.30) | 68.95 (46.27) | 69.62 (45.99) |
| 0 L&D complications (%) | 26.55 (44.16) | 22.46 (41.73) | 23.70 (42.53) |
| 1 L&D complication (%) | 33.60 (47.24) | 38.25 (48.60) | 36.84 (48.24) |
| ≥2 L&D complications (%) | 39.85 (48.96) | 39.28 (48.84) | 39.45 (48.88) |
| Number of L&D complications (mean) | 1.44 (1.32) | 1.39 (1.15) | 1.40 (1.21) |
| Number of observations | 26,151 | 60,195 | 86,346 |

## Discussion

It is perhaps not surprising that physicians—often making complex decisions for others under great uncertainty, high stakes, and increasing time, cognitive, and litigation pressures—fall prey to the very human tendency to rely on heuristics. My study shows that after a patient complication in one delivery mode, physicians are more likely to use the other delivery mode on their next patient. This strategy appears to have small but significantly negative effects on patient health outcomes and increases resource use.

This study should be interpreted in light of its limitations. First, it uses data from only two academic hospitals in the same state, which may limit generalizability of the results. However, the fact that the use of this heuristic has been documented in other settings and with other populations lends credence to these results. Second, despite the obvious advantages of this dataset, it still does not provide variables that would allow more precise identification of the physician's delivery decision (e.g., the patient's original delivery plan). However, this added difficulty should manifest as noise, thereby only making it harder to identify an effect—an assumption I bolster by including a range of sensitivity and robustness checks. Finally, although I present evidence for why this switching behavior is unlikely to be evidence of physician learning, I recognize that the optimality of this heuristic cannot be established using retrospective observational data. That said, controlled experimental designs that allow for more precise measurement of such heuristics will have their own issues (such as the inability to randomize delivery complications, changes in physicians' behavior under observation, etc.).

It is well-known that the US has poor maternal morbidity and mortality indicators. Not having conducted a cross-country examination of physician behavior, I cannot speak to why the US lags behind its peers. However, two features of the American delivery setting may make the use of heuristics—such as the switching heuristic identified here—more pronounced in the US. First, the delivery process in other countries often relies heavily on midwives, while in the US it is usually physician-led. Midwives are overall less likely than physicians to use obstetric interventions (such as cesarean sections) during the delivery (*30*–*32*) and may be less prone to switching delivery modes after complications. Second, malpractice litigation plays an outsize role in US obstetric services (*33*, *34*), which may make US physicians more reactive to delivery complications than physicians in other countries. Of course, this discussion is purely speculative and warrants further, rigorous cross-country study.

In terms of policy implications, the results of this study highlight the potential for algorithms to support and improve physician decision-making in this setting. The use of decision aids in medicine has increased exponentially in recent decades (*35*, *36*), and their role in medicine may be grouped into three broad categories: (i) to directly advise physician decisions while allowing some level of physician autonomy (*37*), for example, by identifying patients who benefit most from hospital discharge at a given hour; (ii) to provide information in ways that assist physician learning (e.g., by summarizing past performance without noise, algorithms may help physicians learn the causal link between their decisions and patient outcomes) (*38*, *39*); and (iii) to discreetly nudge physician decisions, such as by requiring a brief justification of high-stakes decisions (*40*). However, decision aids that are not mandatory or salient in their benefits may be ignored or misused (*41*), and interfering with the physician's ability to respond to clinical cues in real time (by adding administrative hurdles) could adversely affect patient health. Thus, decision-support systems may be able to improve the quality of clinical decision-making, but care must be taken in their implementation.

Finally, as an implication for research, this study highlights the importance of the physician as an individual—one with beliefs, preferences, and experiences rather than simply as a neutral provider of care—in influencing patient outcomes. Despite a plethora of data on patients (e.g., administrative data on sociodemographic factors, surveys on patient health beliefs, geocoded data on patient interactions with health care services), little is known about the physicians who dispense the care. The lack of such data is an added impediment in the study of physician decision-making, which already requires detailed and difficult-to-access datasets (such as the EHR used here). In this study, for instance, it would have been beneficial to understand which physicians are most likely to use this heuristic and why, to allow resources to be directed where they are most needed. Yet there are no (known) data that could allow such a research exploration. Thus, to improve the quality of care, relevant data should be carefully collected on the providers of care, in much the same way that data are collected on judges, police, and other arbiters of important societal decisions.

## REFERENCES AND NOTES


1. M. F. MacDorman, E. Declercq, H. Cabral, C. Morton, *Obstet. Gynecol.* **128**, 447–455 (2016).
2. World Health Organization, *Trends in Maternal Mortality: 1990 to 2013: Estimates by WHO, UNICEF, UNFPA, the World Bank and the United Nations Population Division: Executive Summary* (2014).
3. Pregnancy Mortality Surveillance System; www.cdc.gov/reproductivehealth/maternal-mortality/pregnancy-mortality-surveillance-system.htm.
4. Severe Maternal Morbidity in the United States; www.cdc.gov/reproductivehealth/maternalinfanthealth/severematernalmorbidity.html.
5. B. M. Slotnick, H. M. Katz, *Science* **185**, 796–798 (1974).
6. F. A. Matsen, M. A. Nowak, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 18053–18057 (2004).
7. M. R. Busse, D. G. Pope, J. C. Pope, J. Silva-Risso, *Q. J. Econ.* **130**, 371–414 (2015).
8. N. K. Choudhry *et al*., *BMJ* **332**, 141–145 (2006).
9. N. L. Keating, A. J. O'Malley, J.-P. Onnela, B. E. Landon, *BMJ Open* **7**, e014239 (2017).
10. P. Croskerry, *Acad. Emerg. Med.* **9**, 1184–1204 (2002).
11. D. P. Ly, *JAMA Intern. Med.* **179**, 1140–1141 (2019).
12. E. D. Katz, *Clin. Pract. Cases Emerg. Med.* **3**, 329–332 (2019).
13. A. Tversky, D. Kahneman, *Cognit. Psychol.* **5**, 207–232 (1973).
14. A. Tversky, D. Kahneman, *Science* **185**, 1124–1131 (1974).
15. P. Bordalo, N. Gennaioli, A. Shleifer, *Q. J. Econ.* **127**, 1243–1285 (2012).
16. D. Fudenberg, D. K. Levine, *Proc. Natl. Acad. Sci. U.S.A.* **111** (suppl. 3), 10826–10829 (2014).
17. D. Kahneman, S. Frederick, in *Heuristics and Biases: The Psychology of Intuitive Judgment*, T. Gilovich, D. Griffin, D. Kahneman, Eds. (Cambridge Univ. Press, 2002), pp. 49–81.
18. K. Haggag, D. G. Pope, K. B. Bryant-Lees, M. W. Bos, *Rev. Econ. Stud.* **86**, 2136–2183 (2019).
19. J. M. Belizán, F. Althabe, M. L. Cafferata, *Epidemiology* **18**, 485–486 (2007).
20. J. Sandall *et al*., *Lancet* **392**, 1349–1357 (2018).
21. M. E. Hannah *et al*., *Lancet* **356**, 1375–1383 (2000).
22. World Health Organization, *WHO Recommendations: Non-Clinical Interventions to Reduce Unnecessary Caesarean Sections* (2018).
23. Bill and Melinda Gates Foundation Global Grand Challenges, "Improving Access to and Use of Safe and Appropriate Cesarean Section"; https://gcgh.grandchallenges.org/challenge/improving-access-and-use-safe-and-appropriate-cesarean-section.
24. See supplementary materials.
25. B. J. Moore, S. White, R. Washington, N. Coenen, A. Elixhauser, *Med. Care* **55**, 698–705 (2017).
26. E. Johnson, M. M. Rehavi, D. C. Chan, D. Carusi, *A Doctor Will See You Now: Physician-Patient Relationships and Clinical Decisions* (National Bureau of Economic Research, 2016).
27. G. J. Escobar *et al*., *Am. J. Obstet. Gynecol.* **224**, 137–147.e7 (2020).
28. B. E. Honoré, E. Kyriazidou, *Econometrica* **68**, 839–874 (2000).
29. J. B. Miller, A. Sanjurjo, *Econometrica* **86**, 2019–2047 (2018).
30. P. A. Janssen, E. M. Ryan, D. J. Etches, M. C. Klein, B. Reime, *Birth* **34**, 140–147 (2007).
31. R. P. Newhouse *et al*., *Nurs. Econ.* **29**, 230–250 (2011).
32. J. Sandall, H. Soltani, S. Gates, A. Shennan, D. Devane, *Cochrane Database Syst. Rev.* CD004667 (2016).
33. P. Zwecker, L. Azoulay, H. A. Abenhaim, *Am. J. Perinatol.* **28**, 277–284 (2011).
34. L. M. Glaser, F. A. Alvi, M. P. Milad, *Am. J. Obstet. Gynecol.* **217**, 340.e1–340.e6 (2017).
35. R. T. Sutton *et al*., *NPJ Digit. Med.* **3**, 17 (2020).
36. Z. Obermeyer, B. Powers, C. Vogeli, S. Mullainathan, *Science* **366**, 447–453 (2019).
37. B. J. Dietvorst, J. P. Simmons, C. Massey, *Manage. Sci.* **64**, 1155–1170 (2018).
38. J. Choi, G. Hecht, I. D. Tafkov, K. L. Towry, *Account. Rev.* **95**, 153–172 (2020).
39. C. D. Ittner, D. F. Larcker, in *Controlling Strategy: Management, Accounting, and Performance Measurement*, C. S. Chapman, Ed. (Oxford Univ. Press, 2005), pp. 86–105.
40. D. Meeker *et al*., *JAMA* **315**, 562–570 (2016).
41. S. N. Weingart *et al*., *Arch. Intern. Med.* **163**, 2625–2631 (2003).


## ACKNOWLEDGMENTS

I am incredibly grateful for detailed feedback and general assistance from D. Howard, I. McCarthy, J. Hockenberry, A. Venkataramani, Z. Obermeyer, J. Skinner, K. Towry, J. Zureich, C. Fleming, S. Shah, R. Hawthorne, several anonymous reviewers, and various seminar participants. **Funding:** Supported by the Emory Laney PDS Program. **Author contributions:** I am solely responsible for obtaining and cleaning the data, designing this study, performing the analyses, and writing all versions of this paper. **Competing interests:** I have no conflicts of interest to report. **Data and materials availability:** All code pertaining to this study will be publicly available. Data include protected health information (PHI) and are therefore not publicly available, but are available for a nominal fee to Emory University/Healthcare members and their non-Emory collaborators.

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abc9818
Materials and Methods
Figs. S1 to S5
Tables S1 to S12

24 May 2020; resubmitted 19 March 2021
Accepted 18 August 2021
10.1126/science.abc9818

REPORTS

EXOPLANETS

# A compositional link between rocky exoplanets and their host stars

Vardan Adibekyan[1,2]*, Caroline Dorn[3], Sérgio G. Sousa[1], Nuno C. Santos[1,2], Bertram Bitsch[4], Garik Israelian[5,6], Christoph Mordasini[7], Susana C. C. Barros[1,2], Elisa Delgado Mena[1], Olivier D. S. Demangeon[1,2], João P. Faria[1,2], Pedro Figueira[8,1], Artur A. Hakobyan[9], Mahmoudreza Oshagh[5,6], Bárbara M. T. B. Soares[1,2], Masanobu Kunitomo[10], Yoichi Takeda[11,12], Emiliano Jofré[13,14,15], Romina Petrucci[14,15], Eder Martioli[16,17]

**Stars and planets both form by accreting material from a surrounding disk. Because they grow from the same material, theory predicts that there should be a relationship between their compositions. In this study, we search for a compositional link between rocky exoplanets and their host stars. We estimate the iron-mass fraction of rocky exoplanets from their masses and radii and compare it with the compositions of their host stars, which we assume reflect the compositions of the protoplanetary disks. We find a correlation (but not a 1:1 relationship) between these two quantities, with a slope of >4, which we interpret as being attributable to planet formation processes. Super-Earths and super-Mercuries appear to be distinct populations with differing compositions, implying differences in their formation processes.**

Characterizing the interiors of rocky exoplanets requires measurements of both their mass and radius. These parameters are usually provided by a combination of two techniques: planetary radius is determined via transit observations, and planetary mass via radial velocity (RV) measurements. The derivation of these parameters relies on properties of the host stars, so it is limited by the precision with which those are measured. Exoplanets orbiting stars that are similar to the Sun [i.e., F, G, or K (FGK) spectral types] benefit from the precise characterization that is possible for their host stars (*1*).

We consider a set of 32 low-mass exoplanets (planet mass, $M$, of <10 Earth masses, $M_\oplus$) orbiting 27 FGK stars with uncertainties in both planet mass and radius of <30% (*2*). Above ~4 $M_\oplus$, the distribution of these planets on a mass-radius diagram shows an apparent gap between two populations (Fig. 1). This gap has been interpreted as the separation between rocky exoplanets and gas-rich mini-Neptunes (*3*–*7*). Our goal is to investigate how the properties of the rocky planets depend on the composition of their host stars. To this end, we discard the planets above the radius gap in Fig. 1, leaving 22 planets without substantial water or gas-rich envelopes, that is, rocky planets.

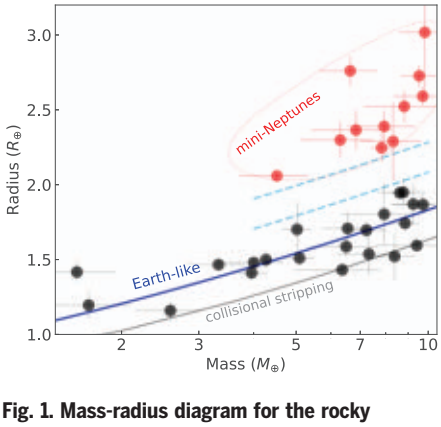



**Fig. 1. Mass-radius diagram for the rocky planets in our sample.** All planets in our sample have masses below 10 $M_\oplus$ and mass and radius uncertainties of <30%. The radii of the planets are given in Earth radii, $R_\oplus$. The light-blue dashed curves, drawn by eye, indicate the location of the radius gap that separates the mini-Neptunes with gaseous envelopes (red circles) from the planets without gaseous envelopes (black circles). The solid blue curve shows the mass-radius relationship expected for Earth-like composition (32% Fe + 68% $MgSiO_3$) (*18*). The gray solid curve marks the minimum planetary radius predicted by a collisional stripping (giant impact) model (*23*). All error bars show one standard deviation.

The abundances of elements in the atmospheres of main-sequence stars reflect their bulk composition (except the lightest elements) within a few percent (*8*), which meteorite measurements have shown to be valid for the Sun (*9*). Theory predicts that the Fe/Si and Mg/Si abundance ratios in stars and planets remain very similar during the planet formation process (*10*–*12*). The atmospheric abundances of refractory elements (such as Mg, Si, and Fe) of solar-type stars are therefore considered a proxy of the composition of the initial protoplanetary disk (*13*, *14*).

We analyzed spectra of 21 host stars of the selected planets (HD 80653 does not have an available spectrum) and measured their atmospheric chemical compositions (*2*). We determined the abundances of Mg, Si, and Fe in the host stars, which are the major rock-forming elements (*15*). We then inserted these abundances into a simple stoichiometric model (*16*, *17*) to estimate the iron-to-silicate mass fraction ($f_{\rm iron}^{\rm star}$) of protoplanetary disks. For the protosolar disk, this model predicts $f_{\rm iron}^{\rm Sun}$ of 33.2 ± 1.7%, which is consistent with the iron content of Earth (~32%), Venus (~32%), and Mars (~30%) but differs from that of Mercury (~70%) (*2*).

At a given mass, the rocky planets in our sample have a dispersion in radius around the curve expected for Earth-like composition (Fig. 1). To investigate whether the densities of these planets are linked to the primordial composition of the protoplanetary disk, Fig. 2 shows the planet density, ρ, normalized to that expected for an Earth-like composition $\rho_{\rm Earth\text{-}like}$ (*18*) as a function of $f_{\rm iron}^{\rm star}$. The normalization accounts for planets with different masses having different densities even with the same composition, owing to compression. We find a relationship between these two quantities, indicating that the final planetary density is a function of the composition of the protoplanetary disk. We performed an orthogonal distance regression (ODR) analysis to quantify the relationship and then used $t$ test statistics to assess its significance, finding that the observed correlation is statistically significant, with a $P$ value of ~$7 \times 10^{-6}$ (*2*).

Because the normalized density is calculated from the observed properties of the

[1]Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, Centro de Astrofísica da Universidade do Porto, 4150-762 Porto, Portugal. [2]Departamento de Física e Astronomia, Faculdade de Ciências, Universidade do Porto, 4169-007 Porto, Portugal. [3]Institute for Computational Science, University of Zurich, CH-8057 Zurich, Switzerland. [4]Max-Planck-Institut für Astronomie, 69117 Heidelberg, Germany. [5]Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain. [6]Departamento de Astrofísica, Universidad de La Laguna, E-38206 La Laguna, Tenerife, Spain. [7]Physikalisches Institut, University of Bern, 3012 Bern, Switzerland. [8]European Southern Observatory, Alonso de Córdova 3107, Vitacura, Región Metropolitana, Chile. [9]Center for Cosmology and Astrophysics, Alikhanian National Science Laboratory, 0036 Yerevan, Armenia. [10]Department of Physics, School of Medicine, Kurume University, Fukuoka 830-0011, Japan. [11]National Astronomical Observatory, Mitaka, Tokyo 181-8588, Japan. [12]Sōgō kenkyū daigakuin daigaku, The Graduate University for Advanced Studies, Mitaka, Tokyo 181-8588, Japan. [13]Instituto de Astronomía, Universidad Nacional Autónoma de México, Ciudad Universitaria, Ciudad de México, C.P. 04510, México. [14]Observatorio Astronómico de Córdoba, Universidad Nacional de Córdoba, X5000BGR Córdoba, Argentina. [15]Consejo Nacional de Investigaciones Científicas y Técnicas, C1425FQB Buenos Aires, Argentina. [16]Institut d'Astrophysique de Paris, UMR7095 Centre national de la recherche scientifique, Sorbonne Université, 75014 Paris, France. [17]Laboratório Nacional de Astrofísica, Itajubá MG 37504-364, Brazil.
*Corresponding author. Email: vadibekyan@astro.up.pt

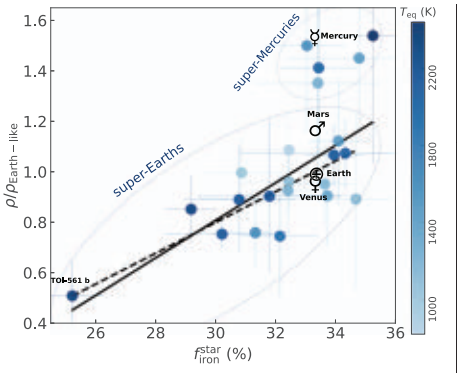



**Fig. 2. Densities of rocky planets as a function of iron fraction.** The measured density of each planet was normalized by the expected density of an Earth-like composition at that mass (*18*) (Fig. 1). The iron-mass fraction was calculated from the elemental abundances of the host star. The rocky planets in the Solar System are indicated with their respective symbols in black; these all have the same iron fraction because it was derived from the abundances of the Sun, not measured directly from the planets. The black solid and dashed lines show the correlation for the full sample and for the sample after excluding the five potential super-Mercuries, respectively. The Solar System planets are not included in the linear regressions. The exoplanet symbols (blue circles) are color coded by the equilibrium temperature of the planets to search for possible trends. No color gradient is visible. All error bars show one standard deviation.

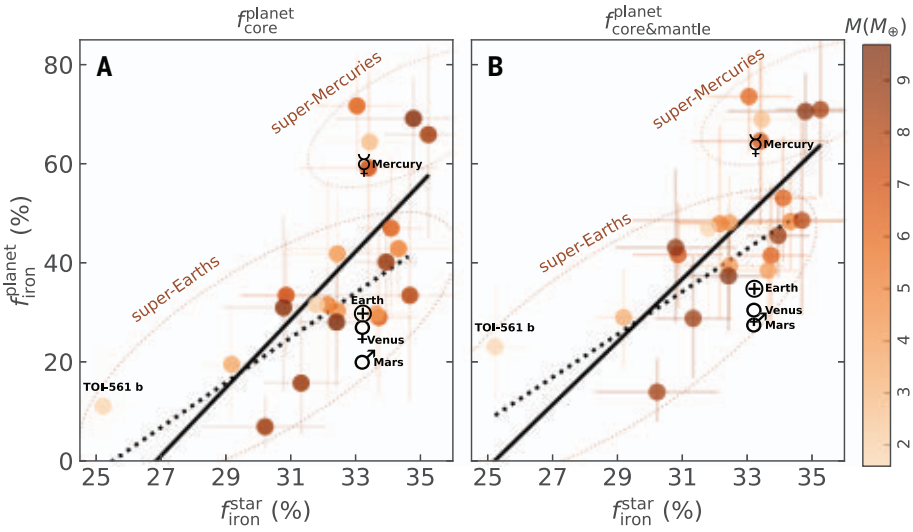



**Fig. 3. Iron contents of rocky planets.** The iron-mass fraction of the protoplanetary disk, estimated from the host star abundances ($f_{iron}^{star}$), is plotted as a function of the iron-mass fraction inferred from the planet's mass and radius ($f_{iron}^{planet}$), as in Fig. 2. The $f_{iron}^{planet}$ values are shown for two assumptions: (**A**) All iron resides in the core only or (**B**) iron is present in both the mantle and core. The color bar indicates the mass of the planets. Symbols and line styles are the same as in Fig. 2. The error bars of $f_{iron}^{star}$ show one standard deviation. The error bars of $f_{iron}^{planet}$ cover the interval between the 16th and the 84th percentiles.

planets, whereas $f_{iron}^{star}$ is inferred from the host star composition, this trend provides observational evidence for a compositional link between rocky exoplanets and their host stars. Rocky planets therefore preserve information about the composition of the protoplanetary disk in which they formed.

Although Fig. 2 shows a correlation, it is evident that for a given $f_{iron}^{star}$, rocky planets can have a range of densities. The observed scatter is compatible with the average uncertainty of $\rho/\rho_{Earth\text{-}like}$. We nevertheless tested whether part of this dispersion could have an astrophysical origin. The size, and therefore density, of planets could be influenced by the flux of high-energy photons that the planets receive from their host stars (*6*). Extreme irradiation from the host star can lead to atmospheric escape from low-mass planets, with a substantial impact on the evolution of their bulk composition (*19*). However, we found no correlation between the normalized density of the planets and their equilibrium temperature, $T_{eq}$ (Fig. 2). Assuming that these planets have maintained a constant $T_{eq}$ since their formation, this lack of correlation suggests that for rocky planets, $T_{eq}$ does not have a dominant impact on their radius. Alternatively, a radial oxidation gradient in the protoplanetary disks might lead to a correlation between orbital distance and composition and therefore also to a correlation between orbital distance and planet density (*20*). However, we also found no correlation between $\rho/\rho_{Earth\text{-}like}$ and the orbital distance of the planets (fig. S1).

We compared the iron-mass fraction of the planets with the iron-mass fraction of the protoplanetary disk, estimated from the host star composition. Using planet interior models (*18*), we calculated the possible iron-mass fraction of each planet ($f_{iron}^{planet}$) using only their mass and radius, without incorporating any constraints from the host star composition. We considered two scenarios: (i) iron is present only in the core or (ii) iron is present both in the core and mantle of each planet. Figure 3 shows the resulting relationship between $f_{iron}^{planet}$ and $f_{iron}^{star}$, which indicates a correlation between those quantities for both scenarios. We again applied ODR and a $t$ test, finding that the correlation is statistically significant ($P$ value of ~$2 \times 10^{-5}$). However, the planets span a wider range of iron-mass fractions than their host stars. The overall distribution of the core-mass fraction (which can be related to the iron-mass fraction) of rocky planets has been shown to be wider than the overall distribution expected from the exoplanet host stars' composition (*21*).

It has been suggested that the iron fraction in planets can be increased relative to the protostellar value if the planets formed close to rocklines (regions where refractory material condenses or sublimates) (*22*). In this model, the enhancement in iron, however, is not high enough to explain the amount of iron in Mercury. This effect could explain the generally higher values of $f_{iron}^{planet}$ compared with $f_{iron}^{star}$. The trend we find in Fig. 3 suggests that if this effect operates, it must depend on the stellar iron-mass fraction, such that stars with higher iron fraction have a larger rockline effect.

We identify a group of five planets (K2-38 b, K2-106 b, K2-229 b, Kepler-107 c, and Kepler-406 b) in Fig. 3 with a higher iron content than the rest of the planets. These five planets appear to be higher-mass analogs of Mercury, so we refer to them as super-Mercuries, a term previously proposed (*23*) by analogy with super-Earth, meaning planets with Earth-like compositions but higher masses. Several mechanisms of planet formation and evolution have been proposed to produce high-density and high-$f_{iron}^{planet}$ super-Mercury planets (*24*). The five super-Mercuries we identify have a wide range of masses, unlike the concentration around ~$5\ M_\oplus$ predicted by simulations of giant impacts (*23*). We suggest that a giant impact alone is not responsible for the high density of super-Mercuries. Planet formation simulations that incorporate collisions are unable to produce the highest-density

super-Mercuries (*25*). There is a possible gap in the $f_{\mathrm{iron}}^{\mathrm{planet}}$ and $\rho_{\mathrm{Earth\text{-}like}}$ plane (Fig. 3) between super-Mercuries and super-Earths; we expect a continuous distribution from collisional stripping, given its stochastic nature.

Although we find only five super-Mercuries, they all orbit stars with high $f_{\mathrm{iron}}^{\mathrm{star}}$ (meaning an overabundance of Fe relative to Mg and Si) and high iron abundance (table S3), which is a proxy for the overall content of heavy elements in stars. The first trend may suggest that the mechanism responsible for the overabundance of iron in these planets is related to the composition of the protoplanetary disk. The second trend could imply a more efficient planet formation, leading to a formation of multiple planets and resulting in frequent collisions. We suggest that both iron enrichment (*22*) and collisional mantle stripping (*25*) may need to be invoked to produce an iron enrichment in the general planet population and explain the presence of super-Mercuries.

Because the super-Mercuries may have an unusual origin and/or evolution, we investigated whether our findings still apply if we exclude them from our sample (*2*). We find that the $f_{\mathrm{iron}}^{\mathrm{planet}}$-$f_{\mathrm{iron}}^{\mathrm{star}}$ correlation remains significant ($P$ value of $<10^{-4}$) for just the sample of super-Earths (Fig. 3) but shifts the slope to lower values (tables S6 and S7). The difference in the slopes is, however, within the uncertainties: 6.3 ± 1.2 against 4.3 ± 0.8 for the case of $f_{\mathrm{core\&mantle}}^{\mathrm{planet}}$. In addition, we found that the Fe/(Mg + Si) abundance ratio estimated for these planets correlates with the Fe/(Mg + Si) ratio of their host stars (see the supplementary text section in the supplementary materials). Our results show a non–one-to-one relationship with a slope greater than 1 (supplementary text).

All but one of the stars in our sample are members of the Milky Way's thin disk. The exception is the ultrashort-period planet TOI-561 b, which orbits a metal-poor star in the thick disk; it was previously found to have an unusually low density (*26*). Figures 2 and 3 show that the low density of this planet is consistent with the general trend and dispersion we find for the entire sample. Theory predicts that planets orbiting around metal-poor thick disk and halo stars should have low iron-mass fractions (*17*).

Several previous studies have sought links between the composition of low-mass planets and their host stars. However, they were based either on single planetary systems (*27*, *28*), on a small sample of planets (*16*, *21*, *24*), or on a statistical comparison of the overall populations (*21*). Owing to these limitations, none of those studies found a strong correlation.

The stellar abundances of major rock-forming elements, such as Fe, Mg, and Si, are commonly used to infer the bulk compositions of rocky planets (*13*, *14*), including those in different stellar populations of the Milky Way (*17*, *29*). Our results provide support for the assumptions made in those studies. The observed correlation we find between $f_{\mathrm{iron}}^{\mathrm{planet}}$ and $f_{\mathrm{iron}}^{\mathrm{star}}$ has a slope steeper than 1 ($f_{\mathrm{iron}}^{\mathrm{planet}}$ is larger than $f_{\mathrm{iron}}^{\mathrm{star}}$); we interpret this as indicating that the composition of the protoplanetary disk (which varies with time and location) influences the resulting composition of planets in a nonlinear fashion.

## REFERENCES AND NOTES


1. S. G. Sousa *et al*., *Astron. Astrophys.* **487**, 373–381 (2008).
2. Materials and methods are available as supplementary materials.
3. B. J. Fulton *et al*., *Astron. J.* **154**, 109 (2017).
4. J. E. Owen, Y. Wu, *Astrophys. J.* **847**, 29 (2017).
5. S. Ginzburg, H. E. Schlichting, R. Sari, *Mon. Not. R. Astron. Soc.* **476**, 759–765 (2018).
6. S. Jin, C. Mordasini, *Astrophys. J.* **853**, 163 (2018).
7. J. Venturini, O. M. Guilera, J. Haldemann, M. P. Ronco, C. Mordasini, *Astron. Astrophys.* **643**, L1 (2020).
8. A. Dotter, C. Conroy, P. Cargile, M. Asplund, *Astrophys. J.* **840**, 99 (2017).
9. M. Asplund, N. Grevesse, A. J. Sauval, P. Scott, *Annu. Rev. Astron. Astrophys.* **47**, 481–522 (2009).
10. J. C. Bond, D. P. O'Brien, D. S. Lauretta, *Astrophys. J.* **715**, 1050–1070 (2010).
11. A. Thiabaud *et al*., *Astron. Astrophys.* **562**, A27 (2014).
12. A. Bonsor *et al*., *Mon. Not. R. Astron. Soc.* **503**, 1877–1883 (2021).
13. C. Dorn *et al*., *Astron. Astrophys.* **577**, A83 (2015).
14. C. T. Unterborn, E. E. Dismukes, W. R. Panero, *Astrophys. J.* **819**, 32 (2016).
15. K. Lodders, *Astrophys. J.* **591**, 1220–1247 (2003).
16. N. C. Santos *et al*., *Astron. Astrophys.* **580**, L13 (2015).
17. N. C. Santos *et al*., *Astron. Astrophys.* **608**, A94 (2017).
18. C. Dorn *et al*., *Astron. Astrophys.* **597**, A37 (2017).
19. J. E. Owen, *Annu. Rev. Earth Planet. Sci.* **47**, 67–90 (2019).
20. D. C. Rubie *et al*., *Icarus* **248**, 89–108 (2015).
21. M. Plotnykov, D. Valencia, *Mon. Not. R. Astron. Soc.* **499**, 932–947 (2020).
22. A. Aguichine, O. Mousis, B. Devouard, T. Ronnet, *Astrophys. J.* **901**, 97 (2020).
23. R. A. Marcus, D. Sasselov, L. Hernquist, S. T. Stewart, *Astrophys. J.* **712**, L73–L76 (2010).
24. J. G. Schulze *et al*., *Planet. Sci. J.* **2**, 113 (2021).
25. J. Scora, D. Valencia, A. Morbidelli, S. Jacobson, *Mon. Not. R. Astron. Soc.* **493**, 4910–4924 (2020).
26. G. Lacedelli *et al*., *Mon. Not. R. Astron. Soc.* **501**, 4148–4166 (2021).
27. J. Lillo-Box *et al*., *Astron. Astrophys.* **640**, A48 (2020).
28. A. Mortier *et al*., *Mon. Not. R. Astron. Soc.* **499**, 5004–5021 (2020).
29. B. Bitsch, C. Battistini, *Astron. Astrophys.* **633**, A10 (2020).
30. V. Adibekyan, abg8794_spectra_DataS1.zip, Figshare (2021); https://doi.org/10.6084/m9.figshare.16451166.v1.


## ACKNOWLEDGMENTS


V.A. thanks A. Tajitsu for reducing the High Dispersion Spectrograph spectra. Based on data collected at the Subaru Telescope (program ID S16B0178S; principal investigator: V.A.), which is operated by the National Astronomical Observatory of Japan. We are honored and grateful for the opportunity of observing the Universe from Maunakea, which has great cultural, historical, and natural significance in Hawaii. This research is also partly based on observations (program ID GN-2017A-FT-20; principal investigator: E.J.) obtained at the international Gemini Observatory, a program of NSF's NOIRLab, which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the NSF on behalf of the Gemini Observatory partnership: the NSF (United States), National Research Council (Canada), Agencia Nacional de Investigación y Desarrollo (Chile), Ministerio de Ciencia, Tecnología e Innovación (Argentina), Ministério da Ciência, Tecnologia, Inovações e Comunicações (Brazil), and Korea Astronomy and Space Science Institute (Republic of Korea). This research is also based on observations collected at the European Southern Observatory (ESO) under ESO programs listed in table S1. This work has made use of data from the European Space Agency (ESA) mission (Gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC). Funding for DPAC has been provided by national institutions, specifically the institutions participating in the Gaia Multilateral Agreement. This research used the facilities of the Italian Center for Astronomical Archive (IA2) operated by INAF at the Astronomical Observatory of Trieste; the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with NASA under the Exoplanet Exploration Program; and the SIMBAD database, operated at CDS, Strasbourg, France. **Funding:** V.A., S.G.S., N.C.S., S.C.C.B., E.D.M., O.D.S.D., and J.P.F. were supported by Fundação para a Ciência e Tecnologia (FCT) through national funds and by FEDER through COMPETE2020 - Programa Operacional Competitividade e Internacionalização by these grants: UID/FIS/04434/2019, UIDB/04434/2020, UIDP/04434/2020, PTDC/FIS-AST/32113/2017 and POCI-01-0145-FEDER-032113, and PTDC/FIS-AST/28953/2017 and POCI-01-0145-FEDER-028953. V.A., E.D.M., N.C.S., and S.G.S. acknowledge support from FCT through Investigador FCT contracts IF/00650/2015/CP1273/CT0001, IF/00849/2015/CP1273/CT0003, IF/00169/2012/CP0150/CT0002, and IF/00028/2014/CP1215/CT0002, respectively, and POPH/FSE (EC) by FEDER funding through the program "Programa Operacional de Factores de Competitividade - COMPETE." C.D. acknowledges support from the Swiss National Science Foundation (SNSF) under grant PZ00P2_174028. C.D. and C.M. acknowledge support from the National Center for Competence in Research PlanetS supported by the SNSF. O.D.S.D. and J.P.F. are supported by contracts (DL 57/2016/CP1364/CT0004 and DL57/2016/CP1364/CT0005, respectively) funded by FCT. E.M. acknowledges funding from the French National Research Agency (ANR) under contract ANR-18-CE31-0019 (SPlaSH). B.B. was supported by the European Research Council (ERC Starting Grant 757448-PAMDORA). **Author contributions:** V.A. led the data analysis and wrote the paper, with contributions from C.D. C.D. led the planetary interior analysis, and S.G.S. performed the stellar parameter analysis. V.A. determined the composition of the stars, with contributions from E.D.M. N.C.S., B.B., C.M., B.M.T.B.S., and G.I. contributed to the discussion of the implications of the data. S.C.C.B., O.D.S.D., A.A.H., and M.O. contributed to the general discussion of the results. C.D., S.G.S., N.C.S., G.I., C.M., E.D.M., M.K., Y.T., E.J., R.P., and E.M. gathered the spectroscopic observations. J.P.F. and P.F. performed the statistical analysis. All authors discussed the results and commented on the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The archival observations used were obtained from the sources listed in table S1. Our own observations are available in the following observatory archives: the Gemini Observatory Archive at https://archive.gemini.edu/searchform under program ID GN-2017A-FT-20 and the Subaru Mitaka Okayama Kiso Archive at https://smoka.nao.ac.jp/objectSearch.jsp by selecting Subaru HDS, then searching for Kepler-20 and Kepler-78. Our final co-added spectra for all stars in the sample are available on Figshare (*30*). The Python code we used to calculate $f_{\mathrm{iron}}^{\mathrm{star}}$ is provided in data S1. The output values for the properties of the stars and exoplanets are listed in tables S3 and S4, respectively. The numerical values of the correlations we found using ODR are listed in tables S6 and S7.


## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abg8794
Materials and Methods
Supplementary Text
Figs. S1 to S4
Tables S1 to S8
References (*31*–*96*)
Data S1

10 February 2021; accepted 27 August 2021
10.1126/science.abg8794

TROPHIC CASCADES

# Physical disturbance by recovering sea otter populations increases eelgrass genetic diversity

Erin Foster[1,2]*, Jane Watson[3], Matthew A. Lemay[1], M. Tim Tinker[4,5], James A. Estes[5], Rebecca Piercey[1], Lauren Henson[2,6], Carol Ritland[7], Allyson Miscampbell[7], Linda Nichol[8], Margot Hessing-Lewis[1], Anne K. Salomon[9], Chris T. Darimont[1,2,6]

**Most knowledge regarding the role of predators is ecological in nature. Here, we report how disturbance generated by sea otters (*Enhydra lutris*) digging for infaunal prey in eelgrass (*Zostera marina*) meadows increases genetic diversity by promoting conditions for sexual reproduction of plants. Eelgrass allelic richness and genotypic diversity were, respectively, 30 and 6% higher in areas where recovering sea otter populations had been established for 20 to 30 years than in areas where they had been present <10 years or absent >100 years. The influence of sea otter occupancy on the aforementioned measures of genetic diversity was stronger than those of depth, temperature, latitude, or meadow size. Our findings reveal an underappreciated evolutionary process by which megafauna may promote genetic diversity and ecological resilience.**

Predators can have profound effects on ecosystems, but most of what is known about these effects stems from changes in community structure brought about by trophic interactions (*1*, *2*). By contrast, nontrophic genetic consequences of predation have rarely been examined (*3*, *4*). Disturbance is a major driver of genetic diversity (*5*) and often results from foraging activity (*6*). However, most inquiries into the genetic consequences of disturbance focus on abiotic forces that enhance diversity by removing barriers to gene flow (*7*) or reduce diversity by diminishing effective population size (*8*). Moreover, the known evolutionary consequences of biotic disturbance on plants are limited to direct effects of herbivores and pathogens (*9*); indirect effects are rarely considered (*10*). Here, we describe genetic changes in a plant (eelgrass, *Zostera marina*) resulting from substrate disturbance by a predator (sea otter, *Enhydra lutris*).

Sea otters regularly dig for infaunal prey (e.g., bivalves) in eelgrass meadows (*11*), inadvertently breaking up rhizomes, creating pits (*12*), and leaving bare patches where seeds may settle (Fig. 1, A to C). Eelgrass meadows can consist of clonal perennial plants, which propagate vegetatively through rhizomes; sexually reproducing annual plants, which produce flowers and seeds; or a combination of the two (*13*). Most contemporary eelgrass populations are perennial (*13*) but sexually reproducing annual plants, which increase genetic variation through recombination, may be favored in disturbed environments (*14*). We thus hypothesized that digging by otters would increase eelgrass sexual reproduction and allelic richness. Alternatively, we hypothesized that disturbance could decrease genotypic diversity by favoring specific genotypes.

History provides an opportunity to test these hypotheses. Sea otters, which once occupied the coastal margins of the North Pacific, are recovering in British Columbia after being extirpated during the maritime fur trade, which lasted from the mid-1700s to the early 1900s. To test the prediction that eelgrass genetic diversity was influenced by otter-generated disturbance, we used a natural experiment, replicated across two eelgrass populations (Fig. 2 and figs. S1 to S3) (*15*). We documented otter-imposed disturbance and collected eelgrass shoots ($n$ = 462) for genetic analyses from meadows ($n$ = 15) where sea otters had been absent >100 years, had arrived within the past 10 years, or had been present for decades (hereafter absent, recent, or established). We genotyped 13 microsatellite loci (table S1) and quality-checked the results before analysis (tables S2 to S5) (*15*). We accounted for meadow size, temperature, latitude, and depth (all of which can influence eelgrass reproduction and diversity); we also considered other biotic disturbances such as grazing waterfowl (*15*).

Allelic richness was of primary interest because it is relevant to resilience (*16*), increases with sexual reproduction (*17*), and is especially valuable to the evolutionary potential of clonal organisms (*18*). Our genetic analysis targeted microsatellite loci because they had been previously developed for local populations of eelgrass, and a pilot analysis proved them sufficiently polymorphic to address our research questions. Moreover, other studies of eelgrass genetic diversity have shown consistent patterns between microsatellites and single-nucleotide polymorphisms (SNPs) (*19*). Genotypic diversity can be correlated with allelic richness at low diversity (*18*), although the strength and significance of this correlation varied across our samples (fig. S4). Heterozygosity was of limited interest to our hypotheses because its interpretation is uncertain for clonal organisms (*20*).

[1]Hakai Institute, Heriot Bay, BC V0P 1H0, Canada. [2]Department of Geography, University of Victoria, Victoria, BC V8W 2Y2, Canada. [3]Department of Biology, Vancouver Island University, Nanaimo, BC V9R 5S5, Canada. [4]Nhydra Ecological Consulting, St. Margaret's Bay, NS B3Z 2G6, Canada. [5]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95060, USA. [6]Raincoast Conservation Foundation, Bella Bella, BC V0T 1Z0, Canada. [7]Genetic Data Centre, University of British Columbia, Vancouver, BC V6T 1Z4, Canada. [8]Cetacean Research Program, Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, BC V9T 6N7, Canada. [9]School of Resource and Environmental Management, Simon Fraser University, Burnaby, BC V5A 1S6, Canada.
*Corresponding author. Email: erin.foster@hakai.org

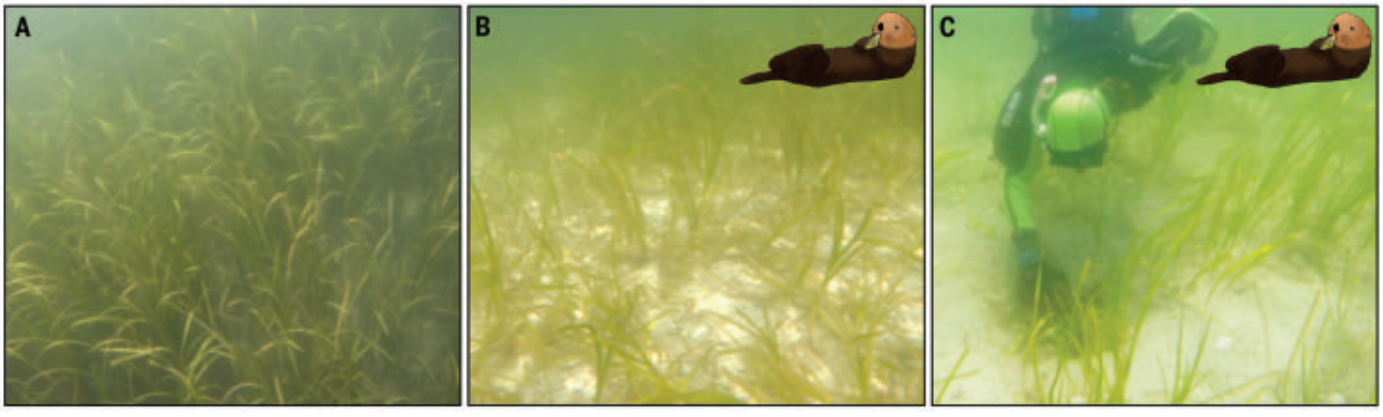



**Fig. 1. Digging by otters changes eelgrass meadows.** (**A**) Otters absent; (**B**) otters established; and (**C**) an otter-foraging pit.

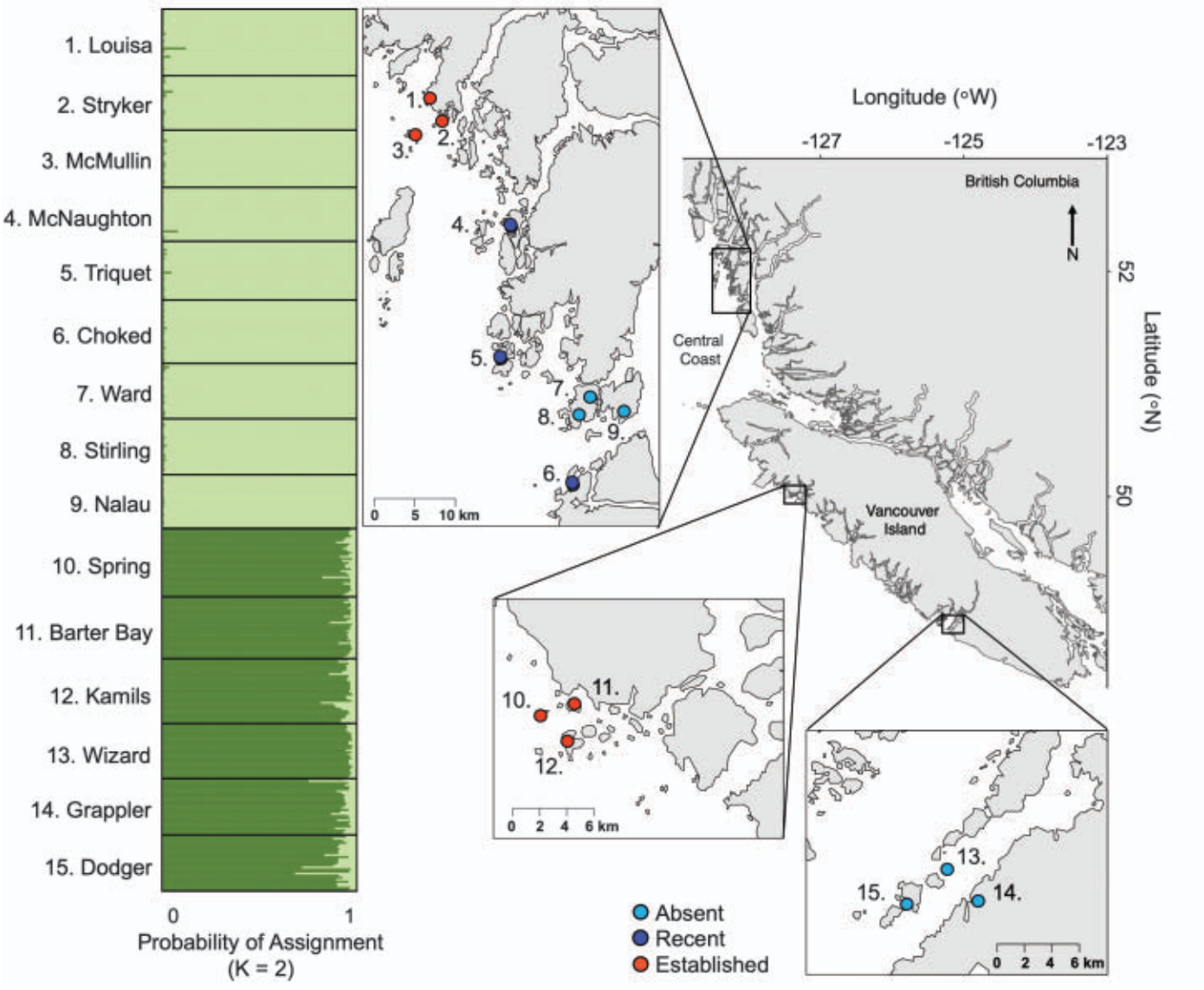



**Fig 2. Eelgrass populations and locations.** Meadow locations with sea otters absent (pale blue circles), recent (<10 years; dark blue circles), and established (20 to 30 years; red circles). STRUCTURE analyses identified two eelgrass populations ($K$ = 2); genetic divergence is likely driven by distance. Pale and dark green bars show the probability of individual population assignment.

We found that otter-imposed disturbance increases eelgrass genetic diversity. In established areas, ~5% of the seafloor area was measurably disturbed by otters. Because foraging pits are ephemeral, this assessment provides a conservative temporal snapshot of actual disturbance (*12*). The mean number of alleles in established areas (46, range 42 to 50) was 21 to 44% greater than that of recent or absent areas (32, range 31 to 34; 38, range 34 to 45; Fig. 3A). In both regions, allelic richness and genotypic diversity were greater in established than either recent or absent areas (Fig. 3, B to C). There were no associations between heterozygosity and other metrics (Fig. 3, D to E).

Allelic richness was best described by the model that included otter occupancy time [relative variable importance (RVI) = 1.00] and water depth (RVI = 0.99) (tables S6 and S7). The magnitude of the effect of otters established for 20 to 30 years was 2.5 times that of the the effect associated with a unit change in depth (Fig. 4A). The mean expected allelic richness in an area with established otters [4.25, 95% credible interval (CI) = 3.40 to 5.38] was ~30% greater than in an equivalent area without otters (3.27, 95% CI = 2.59 to 4.15; table S7 and Fig. 4B). Allelic richness increased by 12% with each 1-m increase in depth (fig. S5), a pattern potentially related to otter behavior as 68% of 342 observed foraging dives in eelgrass occurred at subtidal depths (*21*). Genotypic diversity was best described by the model that included meadow size, otter occupancy, and temperature (tables S8 and S9; RVIs 0.78, 0.74, 0.58). Mean expected genotypic diversity in an area with otters established for 20 to 30 years (0.97, 95% CI = 0.93 to 0.98) was 6% greater than that of an equivalent area without otters (0.91, 95% CI = 0.85 to 0.95); the magnitude of the effect of established otters was twice that of variation in meadow size (~2.3% increase per square kilometer increase) or temperature range (1.3% decrease per degrees Celsius increase; Fig. 4, C to D and fig. S6). For areas in which sea otters had been present <10 years, there was no effect on allelic richness or genotypic diversity, suggesting a lag between otter occupancy time and increasing diversity, likely related to sea otter niche diversification (*21*) and eelgrass generation time (*13*).

The relationship between disturbance, flowering, and seed set is well established, but in this study we implicate digging by

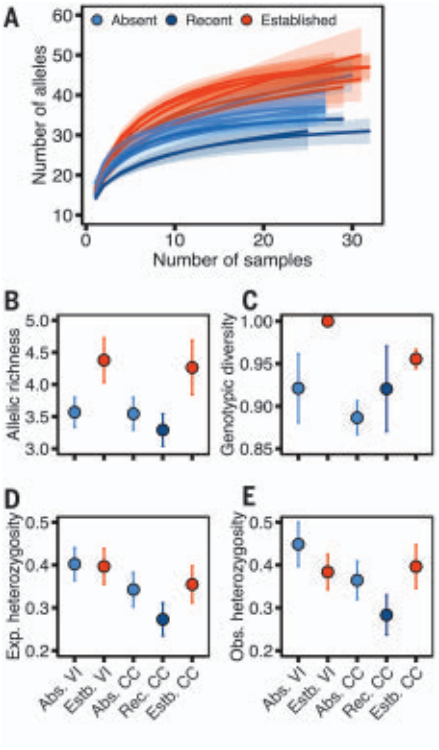



**Fig. 3. Eelgrass genetic data.** (**A**) Allele discovery curves and 95% confidence intervals. (**B**) Mean ± SEM of allelic richness, (**C**) genotypic diversity, (**D**) expected heterozygosity, and (**E**) observed heterozygosity. Abs, otters absent; Rec, otters recent; Estb, otters established; VI, Vancouver Island; CC, Central Coast.

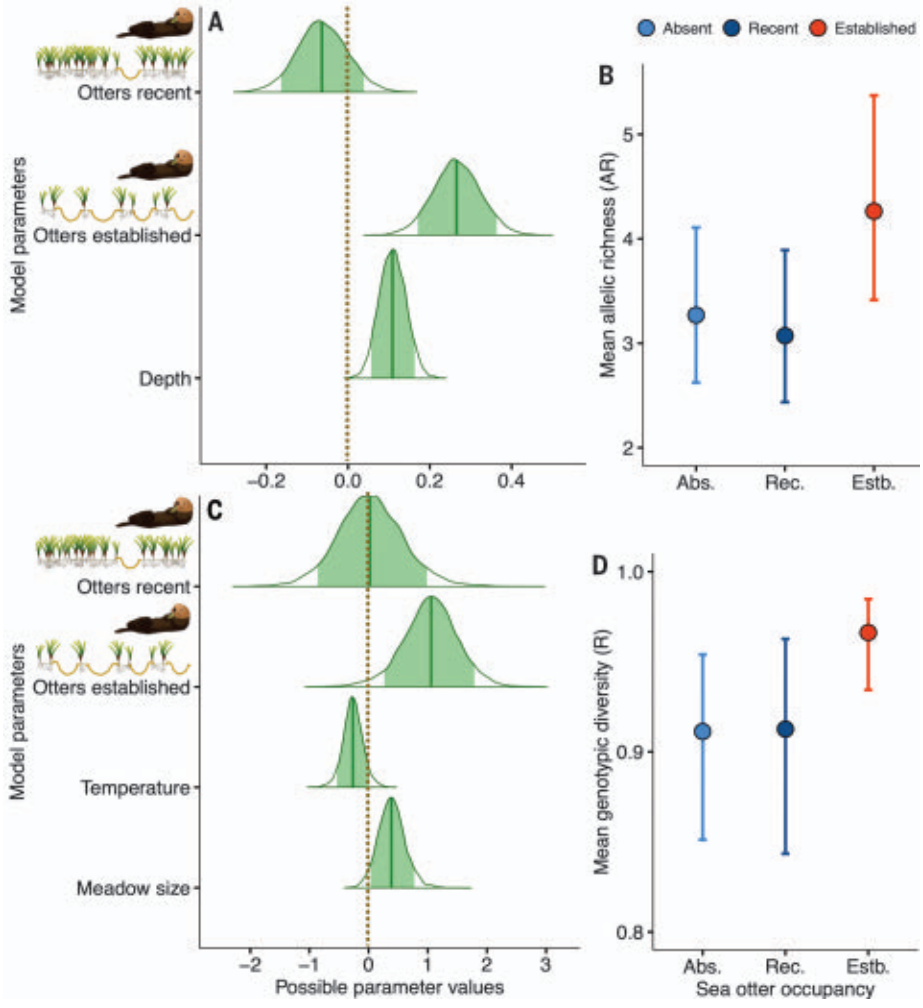



**Fig. 4. Established otters enhance eelgrass genetic diversity.** Posterior distributions of estimated parameters from the best-supported model predicting eelgrass allelic richness (**A**) and genotypic diversity (**C**). Shaded areas show 90% credible intervals (CI). Marginal means of sea otter effects on allelic richness (**B**) and genotypic diversity (**D**); bars indicate 95% CI. Colors and abbreviations as in Fig. 3.

megafauna. Globally, eelgrass meadows that experience sediment disturbance exhibit up to fivefold increases in flowering (*14*), as well as increased seed set (*13*). Disturbance was correlated with otter occupancy time ($r$ = 0.86; fig. S7). Where otters were absent, eelgrass was largely undisturbed, and clones reached lengths of up to 50 m (table S10). We could not assess flowering because flowering periods are highly variable regionally and temporally (*13*).

Sea otters have likely exerted evolutionary and ecological effects on North Pacific eelgrass meadows since at least the mid-Pleistocene (~600 to 700 thousand years ago) (*19, 22*). Digging would have stimulated sexual reproduction and increased seed set, an interaction that unraveled with otter extirpation. The resulting diminished disturbance regime likely favored clonal growth, which led to a decline in allelic richness. Flowering may have also been promoted by traditional harvesting of *Z. marina* rhizomes (*23*) and seeds (*24*) by Indigenous peoples, a practice that declined with European colonization.

Sea otter recovery provides insights into how genetic diversity can be affected when species interactions are lost. These findings have far-reaching implications. Genetic diversity increases the adaptive potential of *Z. marina* populations, which enhances resilience (*25*) and resistance (*26*) to perturbations. Thus, the recovery of sea otters confers both genetic and ecological (*27*) benefits to eelgrass.

Effects on genetic diversity resulting from megafauna-induced disturbance could be widespread. During the Pliocene, purported interactions between gomphotheres and local flora influenced fruit and nut qualities, plant reproductive traits, and dispersal (*4*). Trampling by large vertebrates increases flower production in forbs (*28*), as well as seed germination rates (*29*). In the ocean, foraging grey whales (*Eschrichtius robustus*) may influence genetic diversity in nontarget organisms when disturbing sediment (*6*). In areas where megafauna have recovered, such interactions may be restored. This suggests a heretofore unrecognized role for megafauna: as agents of disturbance with potentially cascading genetic effects, a process lost and overlooked in the many places where large-bodied animals have disappeared.

## REFERENCES AND NOTES:


1. J. A. Estes, M. Heithaus, D. J. McCauley, D. B. Rasher, B. Worm, *Annu. Rev. Environ. Resour.* **41**, 83–116 (2016).
2. J. A. Estes *et al.*, *Science* **333**, 301–306 (2011).
3. J. A. Estes, P. D. Steinberg, *Paleobiology* **14**, 19–36 (1988).
4. D. H. Janzen, P. S. Martin, *Science* **215**, 19–27 (1982).
5. S. C. Banks *et al.*, *Trends Ecol. Evol.* **28**, 670–679 (2013).
6. J. S. Oliver, P. N. Slattery, *Ecology* **66**, 1965–1975 (1985).
7. A. R. Templeton, H. Brazeal, J. L. Neuwald, *Ecology* **92**, 1736–1747 (2011).
8. D. A. Potvin *et al.*, *J. Appl. Ecol.* **54**, 945–954 (2017).
9. A. A. Agrawal, A. P. Hastings, M. T. J. Johnson, J. L. Maron, J. P. Salminen, *Science* **338**, 113–116 (2012).
10. J. A. Estes, J. S. Brashares, M. E. Power, *Am. Nat.* **181** (suppl. 1), S76–S99 (2013).
11. M. Hessing-Lewis *et al.*, *Mar. Pollut. Bull.* **134**, 134–144 (2018).
12. S. B. Traiger, B. Konar, A. Doroff, L. McCaslin, *Mar. Ecol. Prog. Ser.* **560**, 73–86 (2016).
13. S. E. Blok, B. Olesen, D. Krause-Jensen, *Mar. Ecol. Prog. Ser.* **590**, 79–93 (2018).
14. S. Cabaço, R. Santos, *Ecol. Indic.* **23**, 116–122 (2012).

15. Materials and methods are available as supplementary materials.
16. E. Foster *et al*., *Zostera marina* microsatellite and environmental data, Dryad (2021); doi:10.5061/dryad.ns1rn8pt4.
17. A. Caballero, A. García-Dorado, *Genetics* **195**, 1373–1384 (2013).
18. J. M. Ruiz *et al*., *Mar. Pollut. Bull.* **134**, 49–54 (2018).
19. E. U. Rechsteiner *et al*., *Ecol. Evol.* **9**, 3321–3334 (2019).
20. S. I. Massa, C. M. Paulino, E. A. Serrão, C. M. Duarte, S. Arnaud-Haond, *BMC Ecol.* **13**, 39 (2013).
21. S. Arnaud-Haond, C. M. Duarte, F. Alberto, E. A. Serrão, *Mol. Ecol.* **16**, 5115–5139 (2007).
22. R. W. Boessenecker, *J. Mamm. Evol.* **25**, 27–35 (2016).
23. J. A. Coyer *et al*., *Syst. Biodivers.* **11**, 271–284 (2013).
24. S. Cullis-Suzuki *et al*., *Aquat. Bot.* **127**, 26–34 (2015).
25. R. Felger, M. B. Moser, *Science* **181**, 355–356 (1973).
26. T. B. Reusch, A. Ehlers, A. Hämmerli, B. Worm, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 2826–2831 (2005).
27. A. R. Hughes, J. J. Stachowicz, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8998–9002 (2004).
28. B. B. Hughes *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 15313–15318 (2013).
29. A. Elson, D. C. Hartnett, *Am. Midl. Nat.* **178**, 245–259 (2017).

**ACKNOWLEDGMENTS**

We thank the Tula Foundation for support and the Ka:'yu:'k't'h'/Che:k:tles7et'h', Huu-ay-aht, Haíłzaqv, and Wuikinuxv Nations in whose territories we worked. R. Hughes, K. Kay, J. Maron, M. Peterson, D. Rasher, and T. Williams provided early reviews. Hakai Nearshore, G. Ellis, G. Lemieux, and L. Saville provided field support. L. Barrett-Lennard, B. Cameron, C. Crossman, R. Gryba, T. Hanley, R. Hughes, R. Joy, C. Kellogg, N. Kollars, J. Olsen, and J. Stachowicz provided advice. L. Reshitnyk and K. Holmes provided geospatial support. M. Foster and M. Morgan Henderson managed data. **Funding:** E.F. was supported by an NSERC Vanier Award and the Tula Foundation; C.T.D. received an NSERC Discovery Grant. **Author contributions:** E.F. and J.W. were responsible for idea conception. All authors contributed to hypotheses development. E.F. led fieldwork. R.P., C.R., and A.M. led lab work. E.F., C.R., A.M., M.T.T., M.A.L., and L.H. conducted data analyses. All authors contributed to writing. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data and annotated code are publicly available at Dryad (*16*).

**SUPPLEMENTARY MATERIALS**

science.org/doi/10.1126/science.abf2343
Materials and Methods
Figs. S1 to S7
Tables S1 to S10
References (*30*–*75*)
MDAR Reproducibility Checklist

13 October 2020; accepted 17 August 2021
10.1126/science.abf2343

PHYSICS

# Electrically tunable Feshbach resonances in twisted bilayer semiconductors

**Ido Schwartz[1,2]*†, Yuya Shimazaki[1,3]†, Clemens Kuhlenkamp[1,4,5], Kenji Watanabe[6], Takashi Taniguchi[7], Martin Kroner[1], Ataç Imamoğlu[1]***

**Moiré superlattices in transition metal dichalcogenide bilayers provide a platform for exploring strong correlations with optical spectroscopy. Despite the observation of rich Mott-Wigner physics stemming from an interplay between the periodic potential and Coulomb interactions, the absence of tunnel coupling–induced hybridization of electronic states has ensured a classical layer degree of freedom. We investigated a $MoSe_2$ homobilayer structure where interlayer coherent tunneling allows for electric field–controlled manipulation and measurement of the ground-state hole-layer pseudospin. We observed an electrically tunable two-dimensional Feshbach resonance in exciton-hole scattering, which allowed us to control the strength of interactions between excitons and holes located in different layers. Our results may enable the realization of degenerate Bose-Fermi mixtures with tunable interactions.**

The ability to tune the sign and the strength of interactions between particles is key for realization of quantum simulators (*1*). In ultracold-atom experiments, the requisite tunability of interactions is achieved by using Feshbach resonances (*2*–*5*); this has enabled many of the impressive developments in the field, including the preparation of ultracold molecules and the observation of BEC (Bose-Einstein condensate) to BCS (Bardeen-Cooper-Schrieffer state) transition (*6*–*9*). As a consequence of their rich internal energy level structure, two cold atoms prepared in a given internal state with total spin $S$ can be brought into resonance with a bound molecular state of the two atoms with $S' \neq S$ by applying an external magnetic field. Finite hyperfine coupling allows two atoms in a low-energy scattering state (open channel) to hybridize with a bound molecular state (closed channel), resulting in fully tunable scattering phase shifts in the open channel.

Recently, twisted bilayers of two-dimensional (2D) materials have emerged as a new platform for investigating the physics of strongly correlated electrons (*10*–*17*). In these synthetic quantum materials, partial quenching of kinetic energy in emerging flat bands and strong Coulomb interactions play a central role. The role of optical excitations, by contrast, has been mainly restricted to spectroscopy, revealing features such as incompressibility and emergent charge order in Mott-Wigner states of the interacting electron system (*14*–*16*, *18*). Enhancing the strength of exciton-electron interactions could drastically enrich the physics of this system by allowing for photoinduced correlated states in degenerate Bose-Fermi mixtures.

Here, we demonstrate an electrically tunable Feshbach resonance in a twisted homobilayer $MoSe_2$ heterostructure, providing the basis for tunable interactions between bosonic (exciton or polariton) and fermionic (electron or hole) particles. While the presence of a layer pseudospin degree of freedom allows us to obtain open and closed channels that can be tuned in and out of resonance, coherent hole tunneling between the layers provides the counterpart of hyperfine interaction that leads to hybridization of scattering and bound molecular states.

The interplay among strong Coulomb interactions, the moiré superlattice, and coherent interlayer hole tunneling is key to understanding the rich physics of the twisted $MoSe_2$/hBN (hexagonal boron nitride)/$MoSe_2$ heterostructure that we study. To unravel its properties, we use the electric field ($E_z$) and hole density dependence of the exciton-polaron spectrum (*19*, *20*). We find a striking moiré site and filling factor dependence of the many-body system: For unity filling of the moiré superlattice ($\nu = 1$), we observe Mott-like correlated insulator states in the top or bottom layers, as evidenced by the observation of an umklapp resonance (*18*). More generally for $\nu \leq 1$, the hole states show no evidence for layer hybridization. For $\nu > 1$, however, the excess holes occupy layer hybridized sites in the moiré lattice, leading to strong $E_z$ dependence of the associated attractive exciton-polaron resonance energy. Upon further increasing $E_z$ for $\nu > 1$, we find that the umklapp resonance disappears, which suggests that the holes are no longer confined to high-symmetry points of the moiré lattice. Precisely in this parameter regime, we observe an asymmetric avoided crossing in the optically excited state; this demonstrates the hybridization of the top-layer trion state with a continuum of

[1]Institute for Quantum Electronics, ETH Zürich, CH-8093 Zürich, Switzerland. [2]Physics Department and Solid State Institute, Technion–Israel Institute of Technology, 32000 Haifa, Israel. [3]Center for Emergent Matter Science, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. [4]Department of Physics and Institute for Advanced Study, Technical University of Munich, 85748 Garching, Germany. [5]Munich Center for Quantum Science and Technology, 80799 Munich, Germany. [6]Research Center for Functional Materials, National Institute for Materials Science, Tsukuba, Ibaraki 305-0044, Japan. [7]International Center for Materials Nanoarchitectonics, National Institute for Materials Science, Tsukuba, Ibaraki 305-0044, Japan.
***Corresponding author. Email: idosch@technion.ac.il (I.S.); imamoglu@phys.ethz.ch (A.I.)**
†These authors contributed equally to this work.

scattering states composed of a top-layer exciton and a bottom-layer hole. This observation provides unequivocal evidence for a 2D Feshbach resonance (*21*) and the presence of attractive and repulsive polaron branches associated with the emerging interlayer Feshbach molecule.

Our device consists of two $MoSe_2$ layers encapsulated between two thick hBN layers and separated by a monolayer hBN tunnel barrier (Fig. 1A). Top- and bottom-gate voltages, $V_{tg}$ and $V_{bg}$, are applied using transparent thin graphene sheets while the two transition metal dichalcogenide (TMD) layers are grounded (*16*). Figure 1B shows a typical differential reflectance spectrum [$\Delta R/R_0 \equiv (R - R_0)/R_0$] as function of top-gate voltage with a fixed back-gate voltage $V_{bg} = -4$ V. In the range $V_{tg} \approx \pm 6$ V, the excitation spectrum is dominated by the neutral exciton resonances of the top and bottom layers at $E^0_{X_{top}} = 1.632$eV and $E^0_{X_{bot}} = 1.64$eV, indicating that both layers are charge-free. By setting $|V_{tg}| > 6$ V, the top-layer exciton evolves into the repulsive polaron (RP). In addition, an attractive polaron (AP) branch appears at lower energies ($E_{AP_{top}} = 1.607$eV) (*19*).

We use $V_E \equiv 0.5V_{tg} - 0.5V_{bg}$ and $V_\mu \equiv 0.528V_{tg} + 0.472V_{bg}$ to denote a linear combination of $V_{tg}$ and $V_{bg}$ that results in a minimal change to $E_z$ and to the chemical potential μ, respectively (*22*). Figure 1C shows $V_E$-dependent $\Delta R/R_0$ at $\nu = 1$ ($V_\mu = -4.985$ V). By scanning across $V_E = -0.2$ V (corresponding to $E_z = 0$), $AP_{top}$ abruptly disappears in conjunction with the appearance of $AP_{bot}$. Similarly to the electron-doped regime (*16*), this abrupt transfer of all holes from one layer to the other indicates a first-order phase transition between incompressible Mott-like hole states in the top and bottom layers. This incompressible state results in a periodic potential for excitons and the emergence of an umklapp band (*18*). Figure 1D shows the derivative of $\Delta R/R_0$ with respect to energy as a function of $V_E$ at $\nu = 1$. The umklapp resonances appear 2.8 meV above the RPs and are labeled $U_{top}$ and $U_{bot}$. We emphasize that the dominant AP resonance energy remains, to first order, insensitive to the emergence of a periodic charge distribution inherent in the Mott-like correlated state in weak moiré potentials. This observation is consistent with the interpretation of an AP quasiparticle as a collective trion excitation out of electrons or holes in the ground state (*23*); the ground-state charge mobility does not have a measurable effect on the energy of the AP resonance.

Increasing the carrier density to $\nu \geq 1$ reveals a qualitative difference between electron and hole doping: In contrast to changes in electron densities in which a periodic structure of incompressible states appears in the spectra (*16*),

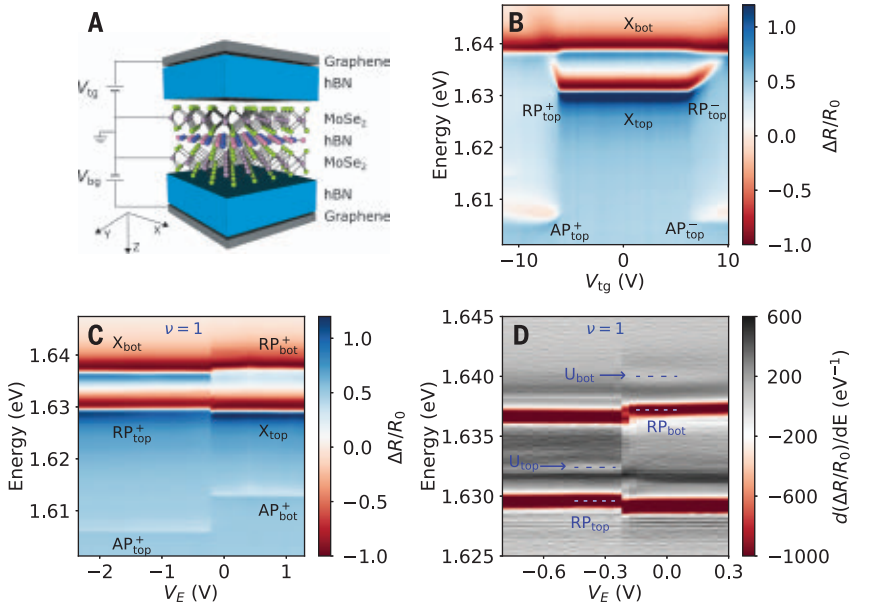



**Fig. 1. Device structure and basic characteristics.** (**A**) Schematic of the device structure. The applied electric field and the optical axis are along the *z* direction. (**B**) Top gate–dependent differential reflectance spectra at a fixed back-gate voltage $V_{bg} = -4$ V. $X_{top}$ and $X_{bot}$, top- and bottom-layer excitons [split as a result of strain (*16*)]; $RP^+_{top}$ and $RP^-_{top}$, top-layer positive and negative repulsive polarons; $AP^+_{top}$ and $AP^-_{top}$, top-layer positive and negative attractive polarons. (**C**) $V_E$-dependent differential reflectance spectra at a fixed chemical potential for unity filling of the moiré superlattice $\nu = 1$. $RP^+_{bot}$ and $AP^+_{bot}$, bottom-layer positive repulsive and attractive polaron resonances. (**D**) $V_E$-dependent differential reflectance spectra differentiated with respect to photon energy at a fixed chemical potential for $\nu = 1$. $U_{top}$ and $U_{bot}$, top- and bottom-layer umklapp resonances. Blue dashed lines mark the energy of the repulsive polarons and the associated umklapp resonances.

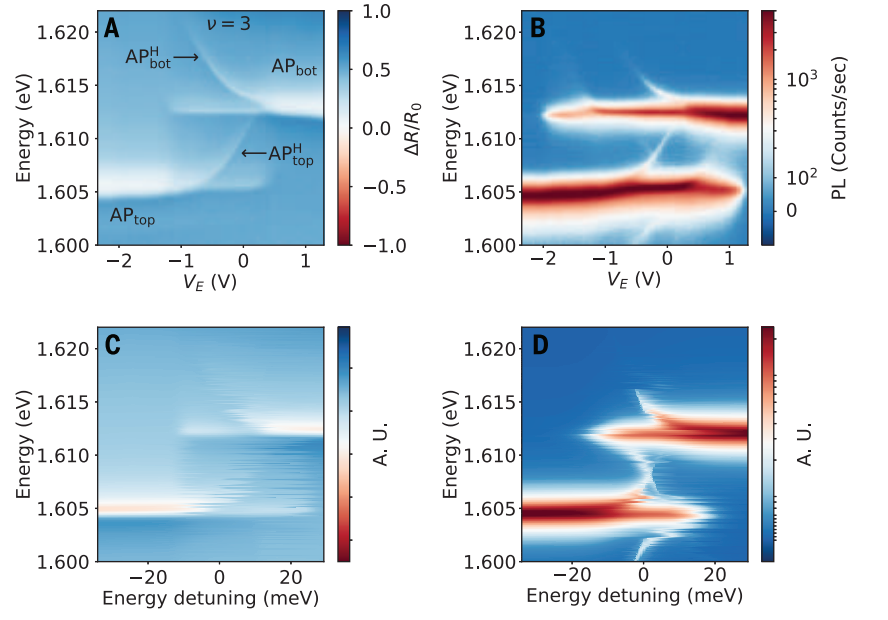



**Fig. 2. Coherent hole tunneling at moiré filling factor $\nu = 3$.** (**A**) $V_E$-dependent differential reflectance spectra at a fixed chemical potential for $\nu = 3$, where each lattice site accommodates three holes. $AP_{top}$ and $AP_{bot}$, top- and bottom-layer attractive polarons; $AP^H_{top}$ and $AP^H_{bot}$, top- and bottom-layer attractive polarons originating from the hybridized hole state, (**B**) $V_E$-dependent PL spectra at a fixed chemical potential for $\nu = 3$. (**C** and **D**) Calculated reflectance spectra (C) and PL (D) at $\nu = 3$ as function of energy detuning between top and bottom layers. A.U., arbitrary units.

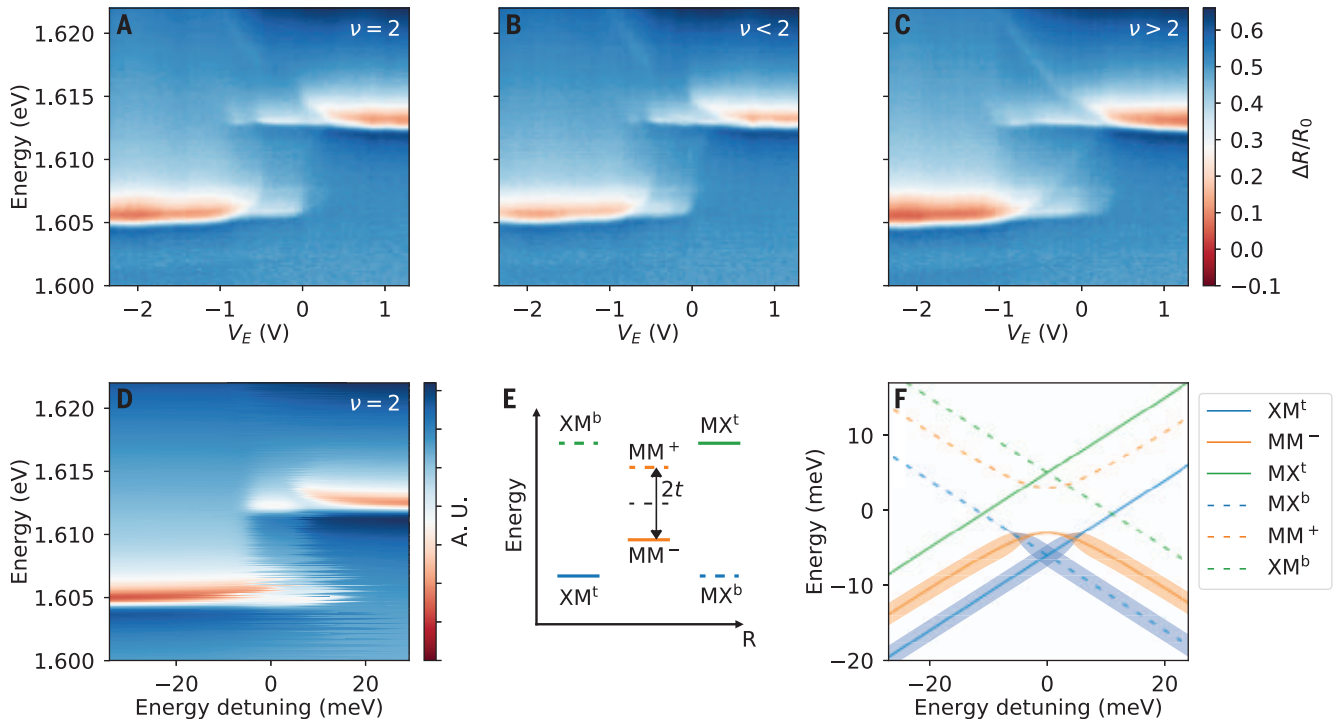



**Fig. 3. Coherent hole tunneling around ν = 2.** (**A** to **C**) $V_E$-dependent differential reflectance spectra at a fixed chemical potential for ν = 2 (A), ν < 2 (B), and ν > 2 (C). (**D**) Calculated reflectance spectra at ν = 2 as function of energy detuning between top and bottom layers. (**E**) Hole energy levels at the three high-symmetry points XM, MX, and MM in the moiré lattice at zero detuning for both top and bottom layers. (**F**) Calculated hole energy levels as function of energy detuning between top and bottom layers. In $XM^t$ and $MX^t$, the holes occupy the top layer; in $XM^b$ and $MX^b$, the holes occupy the bottom layer. $MM^-$ and $MM^+$ represent layer-hybridized hole states. The widths of the semitransparent lines represent hole occupancy at ν = 2.

upon increasing the hole density, both $AP_{top}$ and $AP_{bot}$ split and signatures of an avoided crossing appear in one of their branches. Figure 2A shows $V_E$-dependent $\Delta R/R_0$ of AP resonances at ν = 3 ($V_\mu$ = −5.325 V). Varying $V_E$ from negative to positive, the $AP_{top}$ splits into two branches at $V_E \approx -0.7$ V, where the energy of one branch remains unchanged and the other branch blue-shifts. These observations show that a subset of ground-state holes occupy layer-hybridized states, whereas the others are localized within a layer. A complementary picture is observed for the $AP_{bot}$ by scanning $V_E$ in the opposite direction. Figure 2B shows the $V_E$-dependent photoluminescence (PL) spectra at ν = 3, where the avoided crossing structure is fully visible in both $AP_{top}$ and $AP_{bot}$ with a coupling strength of $2t \approx 6$ meV. The two radiative decay channels, ensuring the observation of both PL branches in the avoided crossing, provide clear evidence for the existence of hybridized states. The absence of the lower-energy branch in $\Delta R/R_0$ in turn allows us to conclude that the higher-energy AP resonances exhibiting strong $V_E$ dependence ($AP^H_{top}$ and $AP^H_{bot}$) originate from the holes occupying the lower-energy hybridized state. Figure 2, C and D, shows the calculated spectra as function of energy detuning between the top and bottom layer for reflection and PL, respectively (*22*).

The blue-shifting branches of $AP_{top}$ and $AP_{bot}$ have comparable strength for $V_E$ = −0.2 V. Upon increasing $E_z$, we observe that the $\Delta R/R_0$ strength of $AP^H_{top}$ decreases while that of $AP^H_{bot}$ increases; likewise, upon decreasing $E_z$, the $\Delta R/R_0$ strength of $AP^H_{bot}$ decreases while that of $AP^H_{top}$ increases. This observation can be understood as arising from the $E_z$ dependence of the probability amplitude for finding the hybridized hole state in the top or bottom layer: For large and positive $E_z$, the hybridized hole state is predominantly in the bottom layer, and for large and negative $E_z$, the hybridized hole state is predominantly in the top layer. In the language of spins, the Bloch vector of the collective hole-layer pseudospin rotates from the equator to the south pole by increasing $E_z$, and from the equator to the north pole by decreasing $E_z$. By contrast, different $AP_{top}$ and $AP_{bot}$ resonance energies allow us to measure the projection of the collective layer pseudospin of holes occupying the hybridized states.

Figure 3A shows $V_E$-dependent $\Delta R/R_0$ at ν = 2 ($V_\mu$ = −5.212 V). For small detuning between the single-particle states of the two layers (−0.5 V < $V_E$ < 0 V), $AP^H_{top}$ and $AP^H_{bot}$ are not visible, indicating that hybridized hole states are not occupied. This gap in the AP spectra is prominent for 1 < ν ≤ 2 (Fig. 3B). In contrast, $AP^H_{top}$ and $AP^H_{bot}$ resonances can be observed for all $V_E$ for ν > 2 ($V_\mu$ = −5.250 V), as can be seen in Fig. 3C. The emergence of such a gap in the optical excitation spectrum provides clear evidence that the holes occupy different sites in the moiré lattice at large ($V_E$ > 0 V or −0.5 V > $V_E$) and small (−0.5 V < $V_E$ < 0 V) detunings for fixed ν.

The observations reported in Figs. 2 and 3 for small $E_z$ can be captured by a simple single-particle model that assumes that holes are subject to a superlattice potential with three minima, strong on-site repulsion, and vanishing hopping between the different sites within a layer. The three minima are attributed to different sites in the moiré superlattice for which the local stacking of the metal (M) and chalcogen (X) atoms of the two layers are denoted by MM, MX, and XM. We conclude, from the lack of avoided crossing signatures below ν ≤ 1, that the lowest single-particle moiré subbands are located at $XM^t$ ($MX^b$) for $V_E$ < −0.2 V ($V_E$ > −0.2 V) (*24*), with the superscripts indicating top- or bottom-layer holes. When 1 < ν ≤ 2, the second moiré subband starts to fill: For large energy detuning

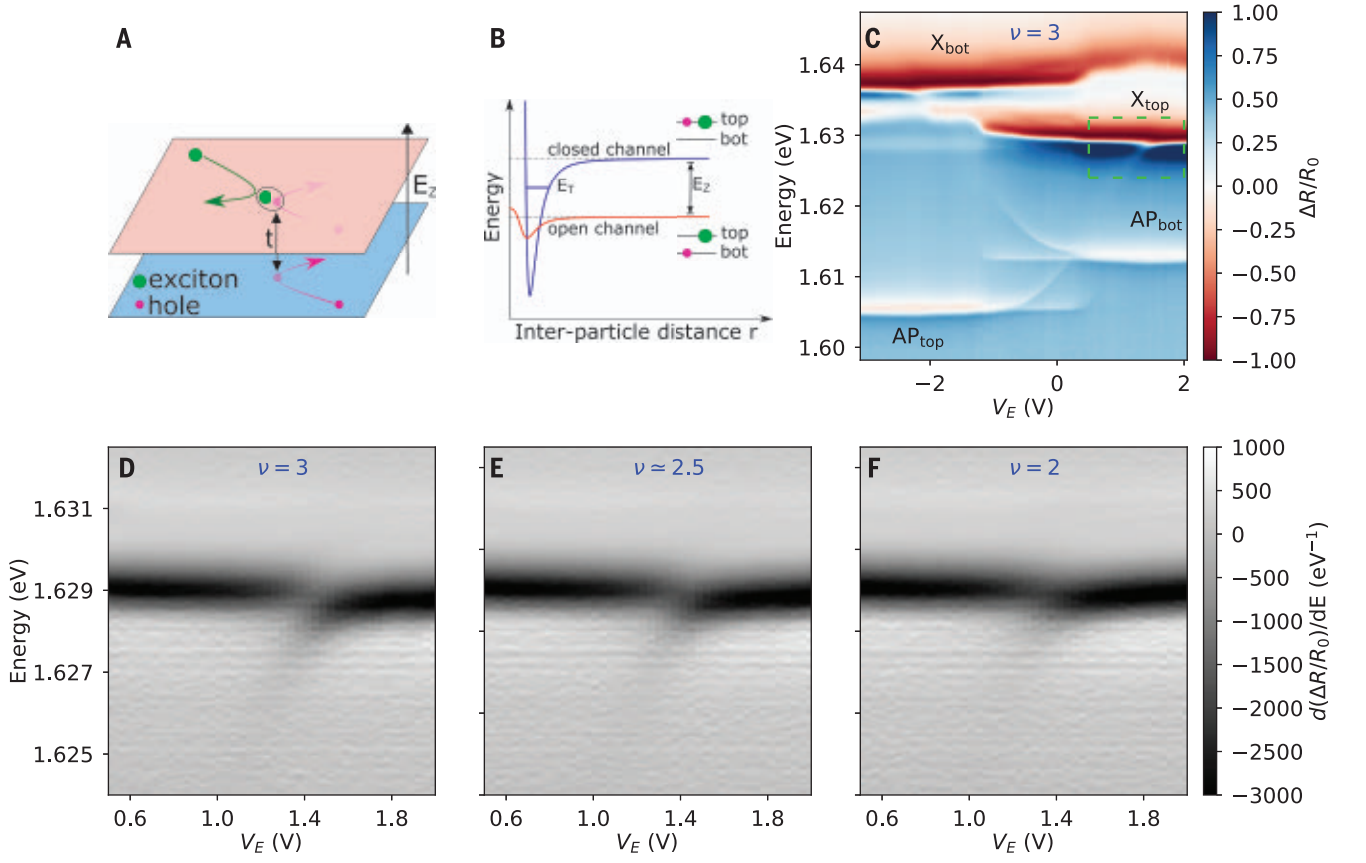



**Fig. 4. Feshbach resonance.** (**A**) Illustration of resonantly enhanced scattering of an exciton and a hole in a bilayer TMD. The Feshbach resonance is achieved by tuning $E_z$ and setting the electrostatic potential difference between the layers to match the intralayer exciton-hole bound-state (trion) energy $E_T$. (**B**) Schematic depiction of the potential energy of an exciton and a hole in the open channel (red) and the closed channel (blue). (**C**) $V_E$-dependent differential reflectance spectra at a fixed chemical potential for ν = 3. (**D** to **F**) $V_E$-dependent differential reflectance spectrum differentiated with respect to energy $E$, in the area marked with dashed green box in (C) for ν = 3, ν ≈ 2.5, and ν = 2, respectively. The spectrum around the $X_{top}$ resonance is visibly asymmetric and depends on the density of holes in the opposite (bottom) layer.

between the layers, this subband is located at the hybridized MM site, and for small energy detuning, the two lowest-energy states are $XM^t$ and $MX^b$. For ν > 2, the third moiré subband starts to fill up, leading to the closing of the gap in the avoided crossing in $\Delta R/R_0$ and indicating that the MM sites are occupied for all energy detunings between the layers. Figure 3E is a schematic of the model showing the energy of the hole states at the three high-symmetry points of the two layers for $E_z = 0$. Figure 3F shows the calculated hole energy levels using the simplified model (*22*).

We emphasize, however, that although the general spectral features for small $|E_z|$ are qualitatively captured by this model, the absence of an umklapp resonance for ν ≥ 2 at large $|E_z|$ indicates that when all holes reside in one layer, they do not form a periodic charge distribution (*22*). We speculate that delocalization of holes is attributable to finite hole hopping among MM, MX, and XM sites, together with screening of the moiré potential.

Having identified the ground-state properties arising from an interplay between Coulomb interactions and interlayer coherent hole tunneling using optical excitations as a spectroscopic tool, we address the Feshbach physics that emerges in the excited state. The relevant scattering process is illustrated in Fig. 4A. By changing the electric field, the potential energy difference of holes in the top and bottom layers can be adjusted. This in turn allows us to adjust the interlayer and intralayer exciton-hole potentials with respect to one another, as depicted schematically in Fig. 4B. We show a wide range of the $V_E$-dependent $\Delta R/R_0$ spectra at ν = 3, covering both AP and RP resonances in Fig. 4C. For concreteness, we focus on the $AP^H_{top}$ branch that is associated with holes originally at the MM sites of the moiré lattice: As $V_E$ is increased, this branch continues to blue-shift while losing oscillator strength. The latter feature is a consequence of the fact that as $V_E$ is increased, the probability of finding a lowest-energy MM site hole in the top layer is decreased: Because the hole in the initial state of the optical transition is predominantly in the bottom layer, the photon energy required to create a top-layer AP increases linearly with $V_E$. Naturally, for large $V_E$, the probability of finding the hole in the top-layer MM site is small, as evidenced by the very faint $AP_{top}$ signal.

Remarkably, for $V_E \geq 1.3$ V, the oscillator strength of the $V_E$-dependent top-layer AP transition starts to increase and exhibits a highly asymmetric anti-crossing with the top-layer exciton transition at $V_E = 1.4$ V. The corresponding avoided crossing between the bottom-layer AP and the bottom-layer exciton transitions is observed at $V_E = -2.1$ V. If this had been a simple zero-dimensional system, such as a quantum-dot molecule, we could have explained the avoided crossing as stemming from hybridization of a top-layer trion with a state consisting of a top-layer exciton

and a bottom-layer hole (*25*, *26*). The experiments depicted in Fig. 4C, by contrast, are carried out in a 2D system with a layer pseudospin degree of freedom (*27*, *28*): The resonance responsible for the avoided crossing is between a bound molecular state (trion) and a continuum of scattering states consisting of 2D excitons and holes with finite probability density at the MM sites of the moiré lattice. This is in stark contrast to the ground-state anti-crossing stemming from interlayer hybridization of two hole states with the same in-plane momentum. Two important distinctions need to be made: (i) Because we are dealing with particles in 2D, the coupled system supports a bound interlayer molecular state for all electric fields $V_E$; this molecule originates from intralayer trions resonantly coupled to a continuum of interlayer exciton-hole states and is referred to as a Feshbach molecule (*4*). (ii) The observed hole density dependence of the avoided crossing can be explained in terms of RP and AP formation associated with this interlayer molecular state (*29*, *30*).

Figure 4D shows the derivative of $\Delta R/R_0$ with respect to energy in the spectral region highlighted by the green dashed box in Fig. 4C. We observe that the splitting is much smaller than the coherent hole tunnel rate. Moreover, we observed strongly asymmetric anti-crossing in the optically excited state, which is a purely many-body effect and is due to the formation of AP and RP resonances associated with the emerging interlayer Feshbach molecule. Shown in Fig. 4, E and F, are derivatives of $\Delta R/R_0$ with respect to energy at $\nu \approx 2.5$ and $\nu = 2$; at lower hole densities, the splitting is further reduced. This reduction is consistent with the observation that both the oscillator strength of the Feshbach AP resonance and the splitting between the AP and RP resonances scale with the density of holes. We also emphasize that the continuous increase of the excited-state splitting from $\nu = 2$ to $\nu = 3$ suggests that all holes injected by increasing $V_\mu$ beyond $\nu = 2$ have a finite probability for occupying the MM sites of the bottom layer. This observation provides further evidence for the 2D nature of bottom-layer holes; if the holes had been localized at high-symmetry points of the moiré lattice, the AP-RP splitting at a fixed $V_E$ would have remained unchanged for $\nu \geq 2$.

Our observation of strongly asymmetric splitting between the AP and RP resonances as the system is tuned across a Feshbach resonance is typical for scattering processes with negligible finite-range corrections (*4*). These features are in contrast to experiments on polariton Feshbach resonances, where signatures of the characteristic scattering physics were not conclusively demonstrated (*21*); there, the cavity-like dispersion of the polaritons leads to a strongly reduced density of states and consequently to prominent finite-range corrections. Recent theoretical work shows that the AP-RP splitting for narrow Feshbach resonances strongly resembles a symmetric anti-crossing (*31*).

Feshbach resonances in 2D differ qualitatively from their 3D counterparts. The 2D Feshbach molecule remains bound for all values of the electric field that is used to tune the system (*32*); moreover, strong correlations between excitons and holes are obtained when the average separation between the bottom-layer holes is comparable to the Bohr radius of the interlayer Feshbach molecule (*30*, *33*). In contrast to cold-atom experiments in a quasi-2D setting where the size of the closed-channel molecule is much smaller than the atomic confinement length scale, the motion of holes and excitons in monolayer $MoSe_2$ is strictly restricted to 2D.

Our experiments establish that bilayer TMD systems with coherent interlayer hole tunneling exhibit 2D Feshbach resonances that induce interlayer Feshbach molecules even when the energy of the intralayer trion is higher than the energy of the scattering state. Such a resonance is used to enhance the interactions between excitons and holes residing in different layers by electrically tuning the binding energy of the Feshbach molecule. The interaction strength is valley-selective and can induce exciton-mediated ferromagnetic interactions between holes in neighboring MM sites. Consequently, we expect our findings to play a role in exploring quantum magnetism in moiré lattices (*34*). Even though our experiments explore enhancement of exciton-hole scattering, the underlying concept is more general; an extension of the concepts and techniques introduced would be provided by an electric field–controlled Feshbach resonance between a scattering state of two electrons in a single layer and the bound interlayer trimer state (*35*, *36*). Furthermore, the reported experiments establish that the combination of applied electric field, interlayer hole tunneling, and layer-selective optical excitations can in principle allow for arbitrary rotation of the layer pseudospin on the Bloch sphere together with projective measurements in the up-down basis. The degree of achievable control could enable a new set of quantum optics experiments in 2D materials, including optical pumping of valley and layer pseudospin and electromagnetically induced transparency exploiting layer coherence.

## REFERENCES AND NOTES


1. R. P. Feynman, *Int. J. Theor. Phys.* **21**, 467–488 (1982).
2. H. Feshbach, *Ann. Phys.* **19**, 287–313 (1962).
3. S. Inouye *et al.*, *Nature* **392**, 151–154 (1998).
4. C. Chin, R. Grimm, P. Julienne, E. Tiesinga, *Rev. Mod. Phys.* **82**, 1225–1286 (2010).
5. I. Bloch, J. Dalibard, S. Nascimbène, *Nat. Phys.* **8**, 267–276 (2012).
6. M. Bartenstein *et al.*, *Phys. Rev. Lett.* **92**, 120401 (2004).
7. C. A. Regal, M. Greiner, D. S. Jin, *Phys. Rev. Lett.* **92**, 040403 (2004).
8. M. W. Zwierlein *et al.*, *Phys. Rev. Lett.* **92**, 120403 (2004).
9. T. Bourdel *et al.*, *Phys. Rev. Lett.* **93**, 050401 (2004).
10. Y. Cao *et al.*, *Nature* **556**, 80–84 (2018).
11. A. L. Sharpe *et al.*, *Science* **365**, 605–608 (2019).
12. X. Lu *et al.*, *Nature* **574**, 653–657 (2019).
13. X. Liu *et al.*, *Nature* **583**, 221–225 (2020).
14. E. C. Regan *et al.*, *Nature* **579**, 359–363 (2020).
15. Y. Tang *et al.*, *Nature* **579**, 353–358 (2020).
16. Y. Shimazaki *et al.*, *Nature* **580**, 472–477 (2020).
17. L. Wang *et al.*, *Nat. Mater.* **19**, 861–866 (2020).
18. Y. Shimazaki *et al.*, *Phys. Rev. X* **11**, 021027 (2021).
19. M. Sidler *et al.*, *Nat. Phys.* **13**, 255–261 (2016).
20. D. K. Efimkin, A. H. MacDonald, *Phys. Rev. B* **95**, 035417 (2017).
21. N. Takemura, S. Trebaol, M. Wouters, M. T. Portella-Oberli, B. Deveaud, *Nat. Phys.* **10**, 500–504 (2014).
22. See supplementary materials.
23. A. Imamoglu, O. Cotlet, R. Schmidt, *C. R. Phys.* **22**, 1–8 (2021).
24. H. Yu, M. Chen, W. Yao, *Natl. Sci. Rev.* **7**, 12–20 (2019).
25. E. A. Stinaff *et al.*, *Science* **311**, 636–639 (2006).
26. H. J. Krenner *et al.*, *Phys. Rev. Lett.* **97**, 076403 (2006).
27. H. A. Fertig, *Phys. Rev. B* **40**, 1087–1095 (1989).
28. A. H. MacDonald, P. M. Platzman, G. S. Boebinger, *Phys. Rev. Lett.* **65**, 775–778 (1990).
29. C. Kohstall *et al.*, *Nature* **485**, 615–618 (2012).
30. R. Schmidt, T. Enss, V. Pietilä, E. Demler, *Phys. Rev. A* **85**, 021602 (2012).
31. C. Kuhlenkamp, M. Knap, M. Wagner, R. Schmidt, A. Imamoglu, arXiv 2105.01080 [cond-mat.mes-hall] (3 May 2021).
32. D. S. Petrov, G. V. Shlyapnikov, *Phys. Rev. A* **64**, 012706 (2001).
33. I. Bloch, J. Dalibard, W. Zwerger, *Rev. Mod. Phys.* **80**, 885–964 (2008).
34. K. S. Burch, D. Mandrus, J.-G. Park, *Nature* **563**, 47–52 (2018).
35. K. Slagle, L. Fu, *Phys. Rev. B* **102**, 235423 (2020).
36. Y. Zhang, T. Liu, L. Fu, *Phys. Rev. B* **103**, 155142 (2021).


## ACKNOWLEDGMENTS


We thank M. Knap, R. Schmidt, and P. A. Murthy for fruitful discussions. **Funding:** Supported by Swiss National Science Foundation (SNSF) grant 200021-178909/1; Japan Society for the Promotion of Science (JSPS) overseas research fellowships (Y.S.); and the Elemental Strategy Initiative conducted by MEXT, Japan, A3 Foresight by JSPS and CREST (grant no. JPMJCR15F3) and JST (K.W. and T.T.). **Author contributions:** I.S., Y.S., and A.I. supervised the project. I.S. and Y.S. carried out the measurements. Y.S. designed and fabricated the sample. M.K. helped to prepare the experimental setup. K.W. and T.T. grew the hBN crystal. I.S. and Y.S. performed the data analysis. C.K. provided input on the data analysis. I.S., Y.S., C.K., and A.I. wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The data presented in this study are available in the ETH Research Collection (www.research-collection.ethz.ch/handle/20.500.11850/482685).


## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj3831
Materials and Methods
Supplementary Text
Figs. S1 to S8
References (*37*–*42*)


9 May 2021; accepted 31 August 2021
10.1126/science.abj3831

OCEAN CIRCULATION

# The Gulf Stream and Kuroshio Current are synchronized

Tsubasa Kohyama[1]*, Yoko Yamagami[2], Hiroaki Miura[3], Shoichiro Kido[4], Hiroaki Tatebe[5], Masahiro Watanabe[2]

**Observations show that sea surface temperatures along the Gulf Stream and the Kuroshio Current tend to synchronize at decadal time scales. This synchronization, which we refer to as the boundary current synchronization (BCS), is reproduced in global climate models with high spatial resolution. Both in observations and model simulations, BCS is associated with meridional migrations of the atmospheric jet stream. Changes in the strength and path of the ocean currents associated with the jet shifts lead to the synchronicity of surface temperatures. Numerical simulations using a conceptual model and an atmospheric general circulation model are consistent with a notion that BCS is an interbasin air-sea coupled mode. Air temperature patterns similar to the one associated with BCS have been repeatedly observed, including in July of 1994 and 2018.**

The two warm ocean currents, the Gulf Stream and the Kuroshio, are located in the western boundaries of the Atlantic and the Pacific, respectively, so they are referred to as the western boundary currents (WBCs) (*1*–*3*). Meanderings of WBCs and the associated sea surface temperature (SST) variations have long been known to affect local weather and climate in the coastal metropolitan areas, mainly because WBCs transport heat from the tropics to the extratropics and modulate cyclogenesis (*4*) and low cloud formation (*5*). More recently, high-resolution satellite observations have helped reveal that heat released from WBCs has profound effects on the entire troposphere (*6*–*8*). The Gulf Stream and the Kuroshio Current are known as centers of action in the midlatitude intrinsic variability (*9*–*11*) and also serve as surface fingerprints of low-frequency natural climate variability [e.g., the Atlantic Meridional Overturning Circulation (*12*) and the Pacific Decadal Oscillation (*13*)]. Thus, understanding WBCs has major implications for paleoclimatology (*14*), climate modeling (*15*, *16*), and disentangling natural variability from anthropogenic climate change (*17*, *18*).

Although both WBCs have experienced warming during recent decades (*17*), tight linkages between the Gulf Stream and the Kuroshio Current have never been discovered. More than a decade ago, a monograph by Kelly and Dong (*19*) found a hint of the WBC covariability in the upper-ocean heat-content data. They estimated that 26% of heat-content variations over the entire North Atlantic and Pacific were in phase. Nevertheless, because the data length and the spatial resolution were limited at that time, it was difficult to detect a fine structure or well-defined covariability between the two WBCs. Although some climatologists have mentioned this potential WBC covariability as an outstanding issue (*2*), this monograph (*19*) is, to the best of our knowledge, the only observational effort that has been taken to explore a possible linkage between the Gulf Stream and the Kuroshio.

In the present day, satellite-based high-resolution SST data records (*20*, *21*) have become long enough to begin thorough analyses in this vein. The SST distribution in early 2018 may initiate speculations about a linkage between the two WBCs (Fig. 1A). During this time span, SSTs at many points near both of the WBC regions were warmer than those from 1971 to 2000 climatology by ~3° to 5°C, which corresponds to 2 to 3 monthly standard deviations. This simultaneous warm event would be rarely experienced by random chance (i.e., a 1 to 5% chance, assuming an ideally normal distribution). One could partly attribute these record-breaking warm currents to the increasing greenhouse gas forcing (*17*), but this explanation appears to not be the whole story.

In this study, we show that the regional-mean SSTs over the two ocean currents are synchronized for interannual to decadal time scales, which we refer to as the boundary current synchronization (BCS). First, BCS is statistically demonstrated on the basis of observed and modeled data. Next, we define the BCS index and highlight the effect of model resolutions on the fidelity of simulated BCS. Then, the physical mechanism is investigated on the basis of model experiments. Finally, implications are discussed.

Simple regional-mean SST time series are sufficient to suspect the existence of covariability. On the basis of satellite observations and output from a high-resolution global climate model (GCM), GFDL-CM4C192, Fig. 1B shows 5-month running-mean, standardized time series of regional-mean SST anomalies over the Gulf Stream (35°N to 45°N, 80°W to 50°W) and the Kuroshio (35°N to 45°N, 140°E to 170°E) regions. Although it is not necessarily obvious that the model perfectly simulates observed features, the modeled time series at least robustly reproduces the observed synchronicity. We acknowledge that many definitions have been proposed to characterize the variability of the Kuroshio and Gulf Stream regions (*22*–*24*), but because our focus here is interannual to decadal variability of WBCs, we have chosen the Kuroshio and Gulf Stream boxes after a close inspection of low-frequency variance. These boxes can adequately capture the collective behaviors of each WBC while retaining fine structures of frontal variability, which we will demonstrate later.

In both observations and high-resolution GCMs, the SSTs averaged over the two regions exhibit significant coherence at the 95% confidence level only in frequencies lower than 0.05 per month (Fig. 1C). Here, we plot the squared coherences for observations and four GCMs. GFDL-CM4C192, which has finer atmospheric resolution than GFDL-CM4, exhibits higher coherence at low frequency. MIROC6, which has lower oceanic resolution than MIROC6subhires, does not exhibit a statistically significant coherence throughout all frequencies. Although three GCMs, except for MIROC6, qualitatively reproduce BCS, the two GFDL models overestimate the coherence, whereas MIROC6subhires underestimates it.

The BCS covariability is almost simultaneous. For observations and the four GCMs, Fig. 1D shows lag correlations between the two regional-mean SST anomalies. As for observations, GFDL-CM4C192 and MIROC6subhires, the highest correlations are realized within 1-month lag and are significant at the 95% confidence level. In observations and MIROC6subhires, although Gulf Stream leads the Kuroshio by 1 month, this small lag is within an acceptable uncertainty considering the time scale of the phenomenon. This simultaneity is also confirmed by the phase spectra (fig. S1). By contrast, correlations in the low-resolution models are lower than those observed in the high resolution models (correlations in MIROC6 are not significant at the 95% confidence level), and no strong peak at zero-lag is detected.

We define the BCS index as the average of the low-pass filtered, standardized regional-mean SST anomalies over the Gulf Stream ($\tilde{G}$) and the Kuroshio ($\tilde{K}$) regions—i.e., BCS ≡ ($\tilde{G}$ + $\tilde{K}$) / 2, where a tilde denotes performing a 5-month running-mean filter and then

[1]Department of Information Sciences, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan. [2]Atmosphere and Ocean Research Institute, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa-shi, Chiba 277-0882, Japan. [3]Department of Earth and Planetary Sciences, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. [4]Application Laboratory (APL), Japan Agency for Marine-Earth Science and Technology, 3173-25 Showa-cho, Kanazawa-ku, Yokohama-shi, Kanagawa 236-0001, Japan. [5]Research Center for Environmental Modeling and Application (CEMA), Japan Agency for Marine-Earth Science and Technology, 3173-25 Showa-cho, Kanazawa-ku, Yokohama-shi, Kanagawa 236-0001, Japan.
*Corresponding author. Email: tsubasa@is.ocha.ac.jp

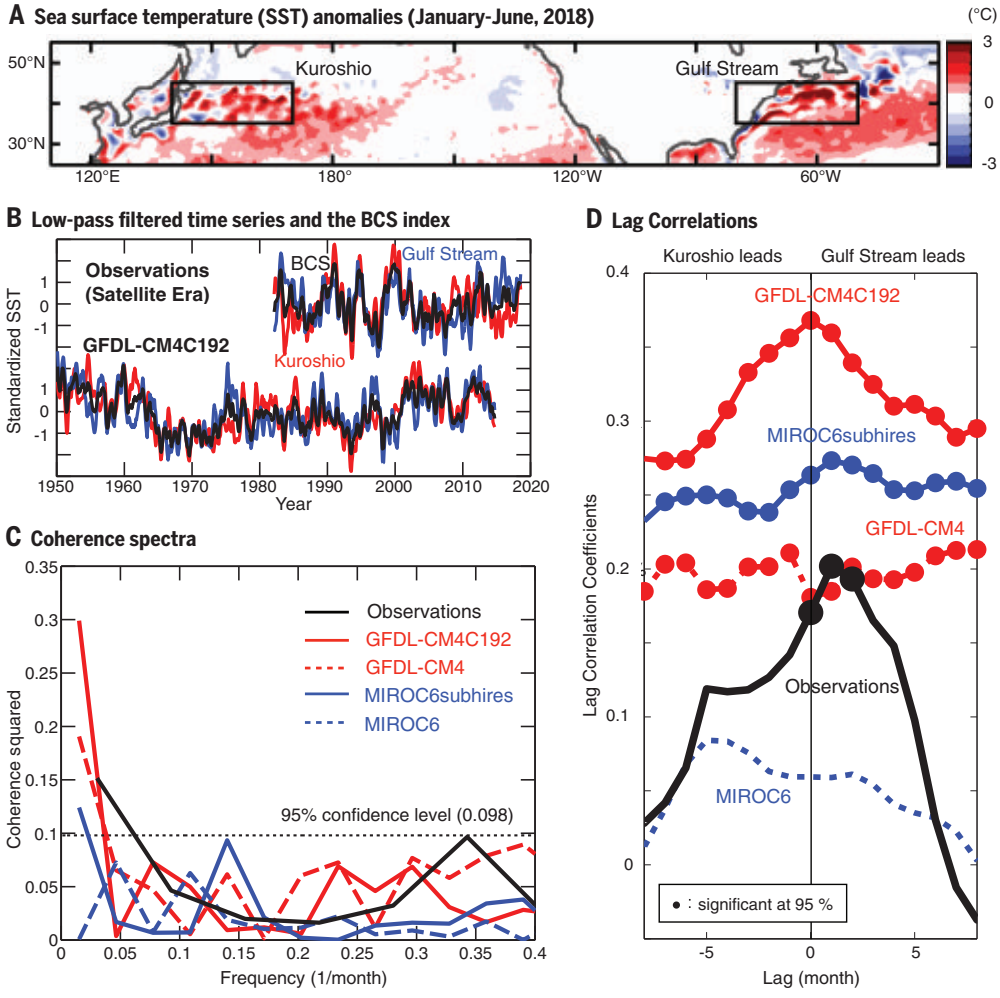



**Fig. 1. Statistical demonstration of the BCS.** (**A**) SST anomalies averaged from January through June of 2018. In the original data, before taking the temporal average at each point, anomalies are calculated relative to 1971 to 2000 climatology and are not detrended. Boxes show the locations of the two WBCs. (**B**) (Top) The observed BCS index (black) and 5-month running-mean, standardized SST anomaly time series averaged over the Gulf Stream (35°N to 45°N, 80°W to 50°W) (blue) and the Kuroshio (35°N to 45°N, 140°E to 170°E) (red) regions defined as the boxes in (A). (Bottom) As in top, but for a high-resolution GCM. (**C**) Coherence spectra between the two SST anomaly time series averaged over the Gulf Stream and Kuroshio regions. The black line is observed data, the solid red and blue lines are from high-resolution GCMs, and the dashed lines are from low-resolution GCMs. The statistical significance calculated based on Amos and Koopmans (*43*) is shown as the black dotted line. (**D**) Same as in (C), but lag correlation coefficients for the 5-month running-mean time series. Positive lags mean that the Gulf Stream leads Kuroshio. Statistically significant correlations are shown as filled circles.

normalizing by its own standard deviation. As is demonstrated in Fig. 1B, the BCS index is defined to capture the temporal variations of SST over both the Gulf Stream and the Kuroshio regions.

One might suspect that the regions used to define the index are too subjective and too large to capture specific features of the boundary currents and that they might reflect SST variability in broader regions rather than in the currents themselves. Therefore, to present a counterargument, we verify here that the BCS index is almost equivalent to more objective and precise time series, which highlights boundary current variability. Figure 2 shows results from the singular value decomposition (SVD) between SST fields in the western North Pacific and Atlantic. This analysis, also known as the maximum covariance analysis, extracts the SST pattern that maximizes the covariance between the two projected time series.

In two out of three models with eddy-permitting oceanic resolutions (i.e., GFDL-CM4C192 and MIROC6subhires), the first SVD (SVD1) mode captures the SST variability of the narrow boundary currents (Fig. 2, A and B). The two projected time series exhibit statistically significant correlations of 0.65 and 0.52 for GFDL-CM4C192 and MIROC6subhires, respectively, which confirms the suitability to perform SVD for these particular fields. Moreover, the mean of the two projected time series exhibits a correlation larger than 0.9 with the BCS index. This high correlation assures us that the simple definition of the BCS index is virtually identical to a more objectively defined index that reflects temperature variations confined to the narrow boundary current regions.

The SVD1 of GFDL-CM4, which also has an eddy-permitting ocean but has a coarser atmosphere than that of GFDL-CM4C192, does not capture boundary currents well (Fig. 2C). On the basis of this result, the low correlations shown in Fig. 1D are the result of the ill-defined SST fronts that originate from the coarse atmospheric resolution. Considering that the MIROC6subhires and GFDL-CM4 are comparable in their resolution, high ocean resolution seems to be a necessary condition but by no means a guarantee for the existence of a well-defined spatial structure of BCS. In MIROC6, which does not resolve oceanic eddies, the SVD1 does not capture the narrow boundary currents at all (Fig. 2D).

The correlations of projected time series in low-resolution models are lower than the high-resolution counterparts. These low correlations support a notion that high-resolution models that adequately resolve mesoscale eddies and their interactions with the atmosphere are essential for an accurate representation of BCS. The BCS index represents boundary current variability only when the spatial grids have sufficiently high resolutions in both atmospheric and oceanic components. This resolution dependence of BCS simulations suggests that the BCS index is not simply reflecting the forcing of large-scale SST anomalies by intrinsic atmospheric variability through air-sea heat fluxes. If the mid-latitude ocean is perfectly passive to the atmosphere, low-resolution models, or even slab-ocean models, should also reproduce the observed BCS.

When we perform the same analysis for observations, BCS variability is captured as the third SVD mode (fig. S3). This result shows

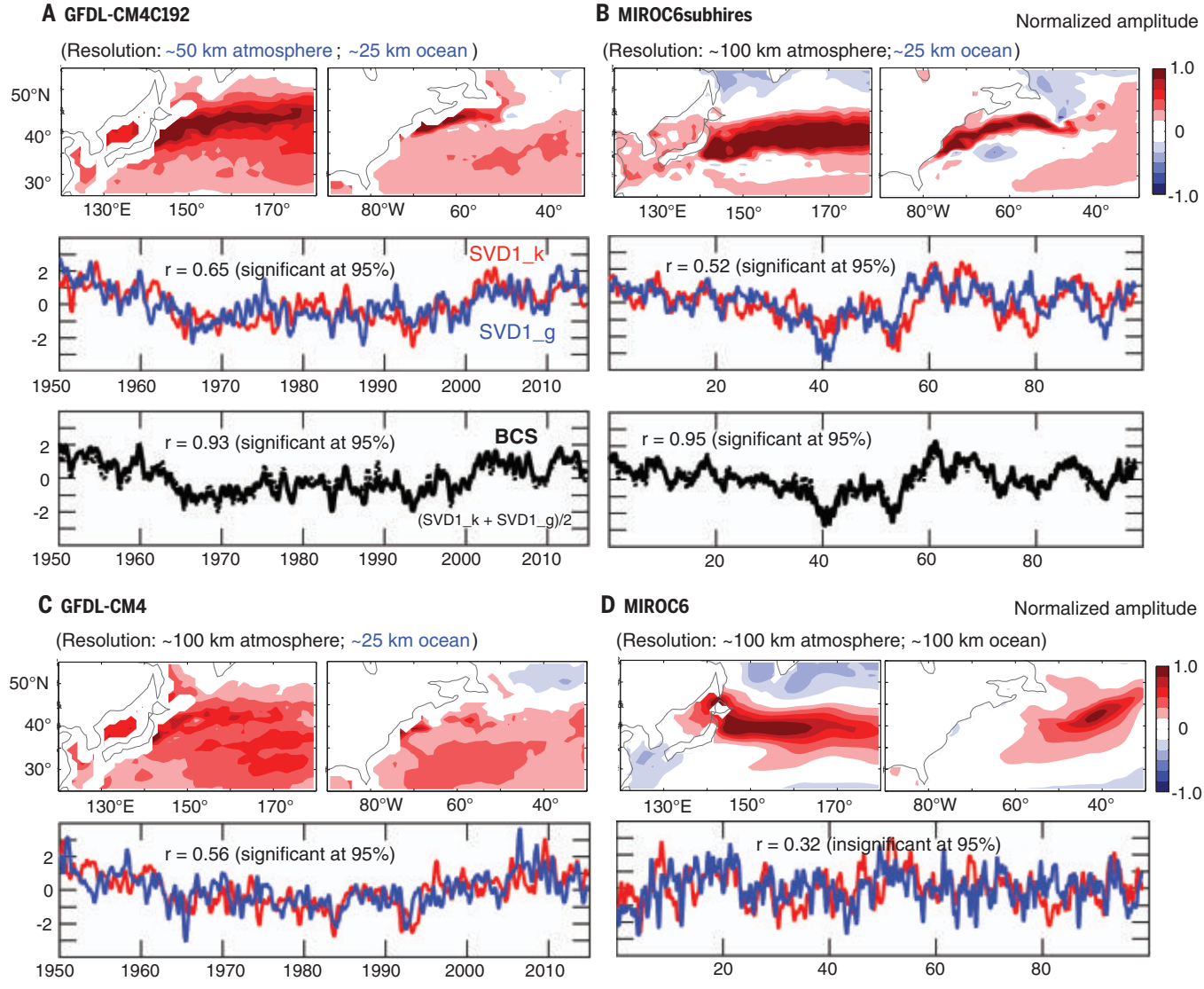



**Fig. 2. Definition of the BCS index and its dependence on spatial resolution.** (**A**) (Top) SST patterns of GFDL-CM4C192 extracted as the first mode of the SVD (SVD1) between the northwest Pacific (25°N to 55°N, 120°E to 180°) and the northwest Atlantic (25°N to 55°N, 90°W to 30°W) regions. (Middle) Projected SST time series onto SVD1 for the northwest Pacific ($SVD1_k$) and the northwest Atlantic ($SVD1_g$). Five-month running-mean filtering is performed. Also shown are the correlation coefficient *r* between the two time series and its statistical significance estimated in fig. S2. (Bottom) Same as in middle, but the BCS index (solid) and the average of $SVD1_k$ and $SVD1_g$ (dashed). (**B**) Same as in (A), but for MIROC6subhires. (**C**) Same as in (A), but for GFDL-CM4. The bottom panel is omitted. (**D**) Same as in (C), but for MIROC6.

that BCS is a more subtle (but existent) signal in the real world—at least after the satellite era—than in the high-resolution models. Although the less dominant SVD signal could be partly the result of the short length of the observational data record over the satellite era, the narrowness of SST variability is consistent with the difficulty at finding the BCS signal in low-resolution GCMs. From GFDL-CM4 and MIROC6, the well-defined boundary current variability is not detected even in the second and third SVD modes.

Meridional migrations of the tropospheric westerly jet stream serve as an essential component of BCS. The regression map of observed SST and zonal winds on the BCS index shows that, when the boundary current regions are anomalously warm, the tropospheric jet stream tends to migrate northward and vice versa (Fig. 3A). The same relationship is reproduced by MIROC6subhires (Fig. 3B). Considering that the two ocean currents are separated by the North American continent, they cannot exchange heat through oceanic pathways within the decadal time scale. Therefore, it is virtually certain that the atmosphere ties temperature variations of two distant ocean currents, and a promising candidate is the eddy-driven jet stream, particularly because tropical signals associated with BCS are much smaller than the extratropical signals. Again, this relationship with the jet stream supports a notion that BCS is not simply passive oceanic responses to intrinsic atmospheric variability because only high-resolution GCMs reproduce BCS, despite the fact that most low-resolution conventional GCMs also exhibit an atmospheric annular mode.

Ocean dynamics is of first-order importance for BCS, which is consistent with previous studies (*25*). The composite maps of geostrophic current strength anomalies show that, when the BCS index is positive, the positions for the boundary currents to be separated from the shores tend to shift northward and vice versa (Fig. 3C). For example, when BCS reaches +2 standard deviations, the Gulf Stream is separated from the North American continent

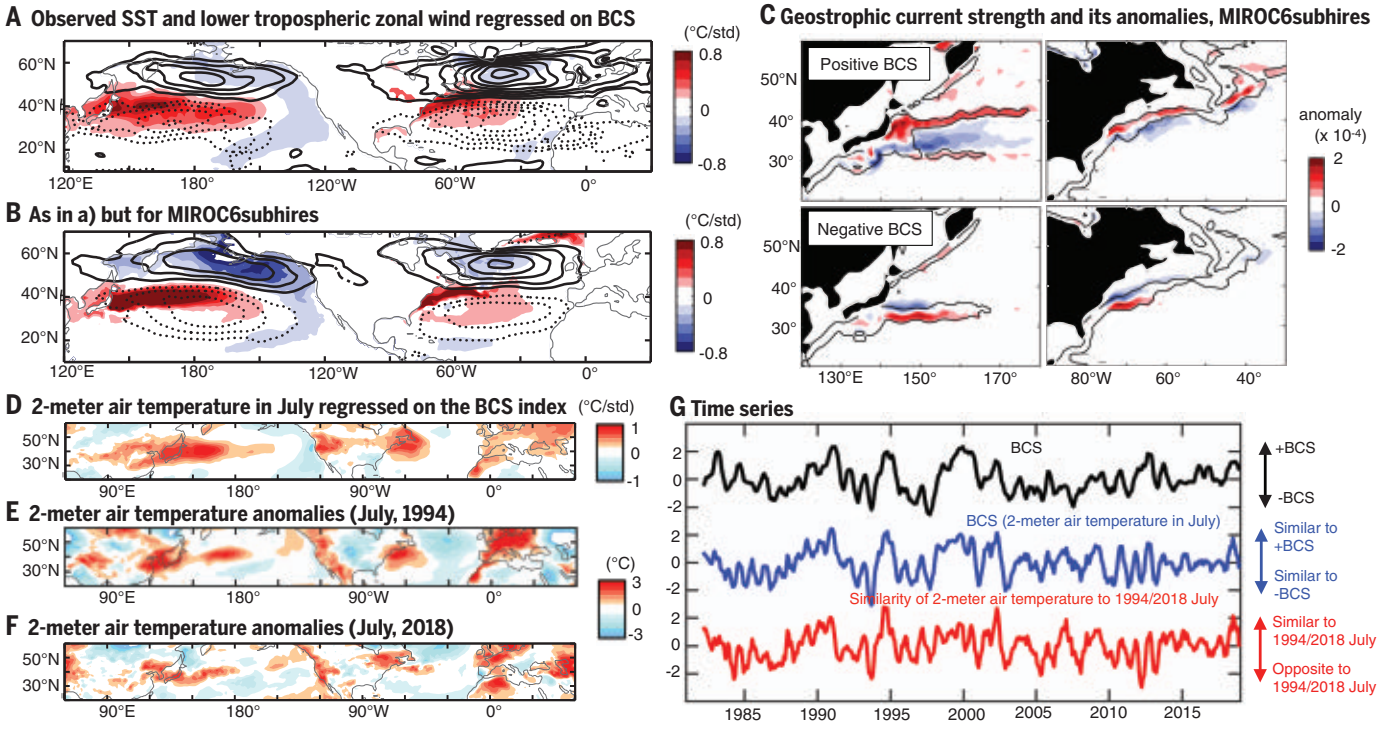



**Fig. 3. Dynamical and thermodynamical manifestations of BCS.** (**A**) Observed anomalies of SST (shaded areas) and zonal winds (contours) at 850 hPa regressed on the BCS index. Contour interval is 0.15 m/s per standard deviation (std). Solid and dashed contours show positive and negative anomalies, respectively, and zero contours are omitted. (**B**) Same as in (A), but for the MIROC6subhires model. (**C**) Composite maps of geostrophic current strength (contours) and its anomalies (shaded areas), which are estimated as anomalies of the absolute value of gradient of the sea level, for extremely positive (top) and negative (bottom) BCS events. Extreme BCS events are defined as the months when the BCS index exceeds its ±2 standard deviations. Contour interval is 1.2 $\times 10^{-4}$. (**D**) Two-meter air temperature anomalies in July regressed on the BCS index calculated using the July-only data. (**E**) Monthly-mean surface temperature anomalies for July of 1994. (**F**) Same as in (E), but for 2018. (**G**) (Black) Same as in the top black curve in Fig. 1B, but standardized by its own standard deviation. (Blue) Two-meter air temperature projected on (D) over 35°N to 45°N, standardized by its own standard deviation. Five-month running-mean filtering is performed. (Red) Same as in blue, but projected on the mean spatial pattern between (E) and (F).

near the Virginia state (38°N). By contrast, when BCS reaches −2 standard deviations, the Gulf Stream is diverted eastward near the South Carolina state (32°N). Similarly, the so-called ripping point of the Kuroshio also varies meridionally from the Iwate prefecture (40°N) to the Aogashima Island (32°N), depending on the phase of BCS.

These meridional shifts of the current pathways are consistent with the temperature variations shown in Fig. 3B, considering that both the Gulf Stream and the Kuroshio Current transport warm water from the tropics. A positive BCS event lets the boundary currents convey more heat to the north, whereas a negative BCS event keeps heat in the south. This synchronization of WBCs is an essential component that gives rise to the synchronization of large-scale SST anomalies, even though they are not equivalent. Figure S4 shows the linear dynamical temperature advection anomalies by the geostrophic current and the net downward heat flux anomalies in units of temperature tendency. Even if the mixed layer depth is conservatively estimated to be 100 m, the role of ocean advection is stronger than that of net surface heat flux by an order of magnitude. This result is consistent with a previous study (*23*) that showed that the geostrophic advection anomaly determines the low-frequency SST variability of this region to the first order. The spatial patterns of ocean advection, particularly by Kuroshio rather than Oyashio, are also more suggestive of the BCS anomalies than those of heat flux.

Because the BCS index exhibits covariability with the atmospheric jet stream (Fig. 3, A and B), large BCS events coincide with mid-latitude extreme weather. The regression map of 2-m air temperature anomalies on the BCS index calculated using the July-only data (Fig. 3D) corresponds well to the spatial patterns of 2-m air temperature anomalies observed in July of 1994 (Fig. 3E) and 2018 (Fig. 3F). East Asia, the west and east coasts of North America, Europe, and Northwest Africa experienced hot summers in 1994 and 2018, and these features are typical anomalies associated with positive BCS events. In Fig. 3G, the covariability between BCS and its 2-m air temperature signature is also shown in the form of time series, which are produced by projecting the temperature data onto the aforementioned spatial patterns. The correlations between the black and blue, black and red, and blue and red curves are 0.72, 0.52, and 0.86, respectively, all of which are statistically significant at the 95% confidence level. These high correlations among the time series support a notion that the air temperature patterns similar to the one associated with BCS have been repeatedly observed, including in July of 1994 and 2018.

Many previous studies have convincingly shown possible physical processes of interactions between large-scale tropospheric winds and WBCs (*3*, *7*, *26*–*28*). In particular, at the beginning of this century, a possibly related theoretical work was presented by Gallego and Cessi (GC01) (*29*). They developed an idealized model of two ocean basins, each having its own WBC, whose stream function is determined through the time-dependent Sverdrup-balance (Fig. 4A). In their model, the two WBCs are coupled to each other only through the zonally symmetric atmosphere.

The idealized model by GC01 illustrates possible mechanisms of WBC covariability

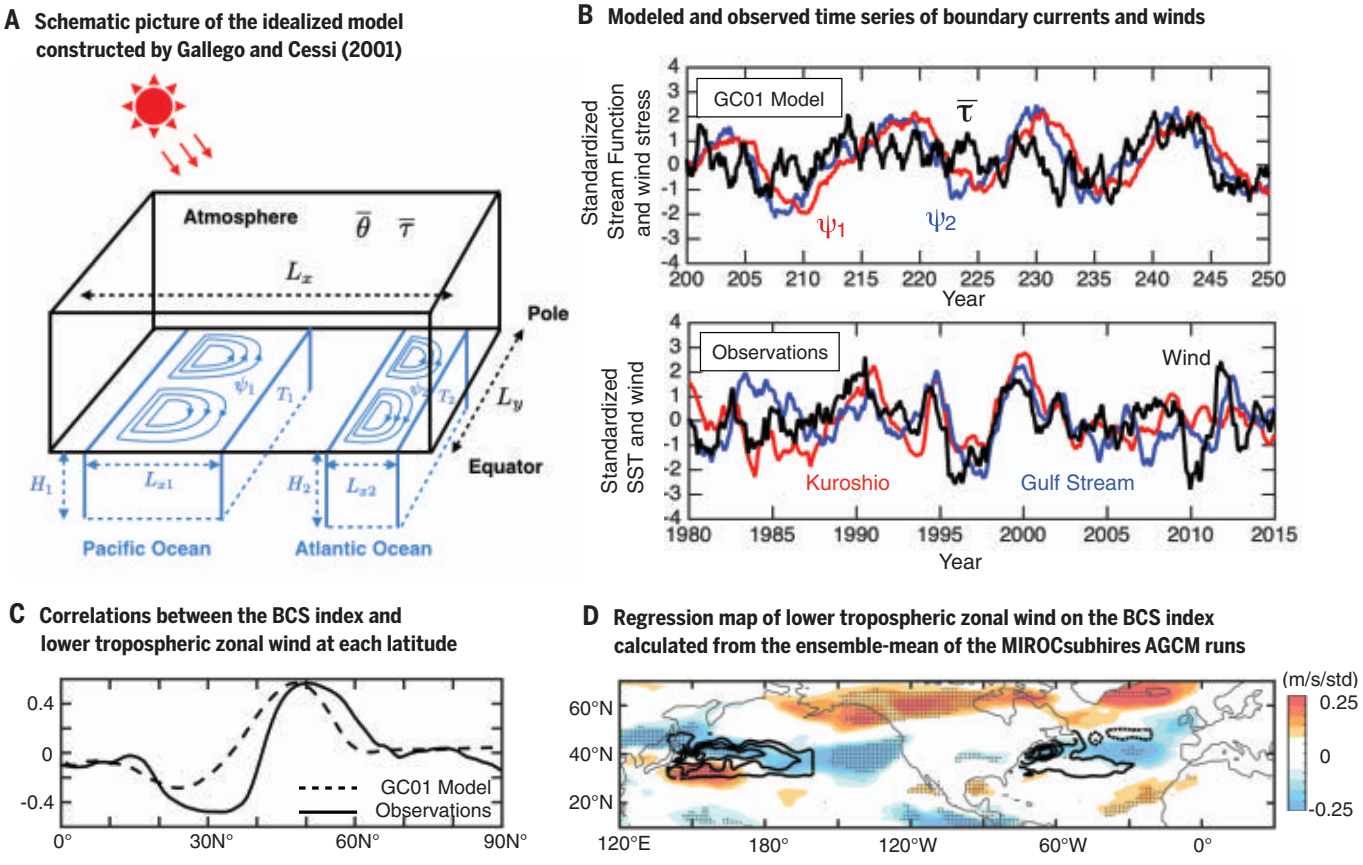



**Fig. 4. Model experiments.** (**A**) Schematic diagram of the idealized model formulated by Gallego and Cessi (*29*) (GC01). (**B**) (Top) Modeled stream function anomalies $\psi_1$ (red) and $\psi_2$ (blue) at 4000 km north of the southern boundary. Also shown in black are the zonal-mean westerly wind stress ($\tau$) anomalies at 5200 km north of the equator. (Bottom) Observed SST anomaly time series averaged over the Gulf Stream (35°N to 45°N, 80°W to 50°W) (blue) and the Kuroshio (35°N to 45°N, 140°E to 170°E) (red) regions. Also shown in black is the observed zonal-mean zonal wind at 850 hPa averaged over the Northern Pacific and Atlantic (45°N to 60°N, 140°E to 0°). Each time series is normalized by its own standard deviation. (**C**) (Solid) Observed correlation coefficients between the BCS index and 11-month running-mean zonal wind at 850 hPa zonally averaged over the Pacific and the Atlantic (140°E to 0°). (Dashed) Same as in the solid curve, but between modeled temperature averaged over the two basins $[(T_1 + T_2) / 2]$ and the zonal wind stress ($\tau$). The southern and northern boundaries of the model are set to be 20°N and 90°N, respectively. (**D**) Ensemble-mean anomalies of SST (contours) and zonal winds (shaded areas) at 850 hPa regressed on the BCS index in MIROCsubhires AGCM runs with SST anomalies added to climatology in the Kuroshio and Gulf Stream regions. Contour interval is 0.2°C per standard deviation. Solid and dashed contours show positive and negative anomalies, respectively, and zero contours are omitted. Dotted areas exhibit statistically significant correlations at the 95% confidence level.

through wind stress and heat fluxes and theoretically predicts the existence of a chaotic regime that exhibits BCS-like variability. Using a modified version of the GC01 model with realistic choices of parameters (table S1), the observed synchronicity of the two strengthening WBCs and the westerly jet are reproduced (Fig. 4B). In particular, this conceptual model predicts that the warm phases of WBCs are associated with a northward shift of the jet stream and vice versa, which is consistent with observations and the high-resolution GCMs (Fig. 4C). As a promising candidate to explain BCS, this model presents a physical process where zonally symmetric atmosphere synchronizes the Pacific and Atlantic with different intrinsic frequencies by mediating the information of the two ocean currents. This interbasin coupling mechanism is also consistent with Omrani *et al.* (*11*), who showed, using semi-idealized atmospheric general circulation models (AGCMs), that forcings from both Kuroshio and the Gulf Stream are necessary to maintain the zonally quasi-symmetric Northern Annular Mode variability.

To quantify the role of SST forcing further, we also performed an AGCM experiment. Figure 4D shows the regression map of zonal wind at 850 hPa on the BCS index calculated from the ensemble-mean of the MIROCsubhires AGCM runs with observed SST anomalies added to the model climatology only in the Kuroshio and Gulf Stream regions (which we call the BCS experiment; see the supplementary text). This AGCM experiment is designed to capture a strongly positive BCS event. As a forced response to warm WBC SST anomalies, the jet stream is shifted northward, which is consistent with both observations and the idealized model (Fig. 4C). Similar results are obtained by subtracting zonal winds of the control experiment from those of the BCS experiment.

These results qualitatively support the air-sea coupled mechanism proposed by GC01 in the sense that zonally symmetric atmospheric variability is forced by SST anomalies in the western boundary regions, which in turn feeds back to the ocean through dynamical and thermodynamical processes. In this regard, our additional sensitivity experiments using the GC01 model (supplementary text) show that, when the dynamical (thermodynamical) coupling is artificially removed, correlations between the Gulf Stream and the Kuroshio decline from 0.84 to 0.26 (0.41) (fig. S5). Both dynamical and thermodynamical coupling

mechanisms could play fundamental roles in BCS (*30*–*33*).

Quantitatively, however, zonal wind anomalies simulated by our AGCM experiment are weaker than those of observations and the air-sea coupled model experiments. This small magnitude of the anomalies can at least be partially explained by the following two reasons. First, because the AGCM experiment prescribes the observed SST anomalies as the lower boundary conditions, amplifications of the BCS mode through interactive air-sea coupling are underestimated. For example, air-sea couplings between the storm track and the northern hemispheric WBCs (*34*) would enhance atmospheric variability to realize a realistic BCS amplitude. Similarly, air-sea coupling processes could also operate for locating the SST front to an optimum position where surface heating can enhance the persistence of the atmospheric intrinsic variability (*35*). Second, as is also discussed by Smirmov *et al.* (*16*), the resolution of observational SST data that are added to the lower boundary condition of the AGCM could be too coarse to capture the full role of SST fronts. As shown above, GCMs with high spatial resolution exhibit statistically significant BCS variability, whereas those with low resolution do not. This result is consistent with a notion that sharp SST fronts are another essential ingredient of BCS, in addition to the interbasin coupling mechanism proposed by GC01. By adding idealized SST fronts to an AGCM experiment, Ogawa *et al.* (*36*) also showed that the meridional position of the eddy-driven jet can be anchored by the SST fronts.

Additionally, as a general caveat of this study, the current ocean model resolution (with horizontal resolution of ~1/4 degree) may still not be enough to determine the detailed BCS mechanism with certainty. Some previous studies have suggested that eddy-resolving ocean resolution (~1/10 degree) is needed to properly simulate WBC behavior and their interaction with the atmosphere (*28*, *37*). Under the constraint of current computational resources, however, it is still challenging to obtain air-sea coupled model ensembles of 100-year eddy-resolving simulations in a global domain. In the near future, when computational resources suffice, it will be necessary to revisit our hypothesis that BCS forms an interbasin air-sea coupled mode.

Understanding BCS has immediate implications for human lives because the Gulf Stream and the Kuroshio Current transport heat from the tropics to the extratropics, and their temperature variations affect extreme weather of densely populated areas in the Northern Hemisphere (*4*–*6*). The hot summer experienced in 2018 is a good example of the kind of extreme weather associated with BCS (Fig. 3F). In particular, because other prominent climate modes (e.g., the El Niño Southern Oscillation) were relatively inactive in 2018, the BCS signature may have emerged in the observed air temperature over the entire northern hemispheric extratropics.

BCS also has implications for fishery productions because the variability of WBCs modulates marine ecosystems (*38*–*41*). Warm SST associated with a northward shift of the Gulf Stream increases the mortality of Atlantic cod (*Gadus morhua*) (*38*), whereas migrations of pelagic fish, such as Japanese sardine (*Sardinops melanostictus*) (*39*) and Pacific saury (*Cololabis sairai*) (*40*, *41*), are influenced by the Kuroshio variability because they use the Kuroshio region as spawning and nursery grounds (*42*). Continuous monitoring of the two WBCs with a finer observational network, as well as development of high-resolution GCMs, are necessary for accurate understanding and prediction of BCS in a changing climate.

## REFERENCES AND NOTES


1. N. G. Hogg, W. E. Johns, *Rev. Geophys.* **33**, 1311–1334 (1995).
2. Y.-O. Kwon *et al.*, *J. Clim.* **23**, 3249–3281 (2010).
3. K. A. Kelly *et al.*, *J. Clim.* **23**, 5644–5667 (2010).
4. F. Sanders, *Mon. Weather Rev.* **114**, 1781–1794 (1986).
5. G. S. Young, T. D. Sikora, *Mon. Weather Rev.* **131**, 2177–2191 (2003).
6. S. Minobe, A. Kuwano-Yoshida, N. Komori, S.-P. Xie, R. J. Small, *Nature* **452**, 206–209 (2008).
7. H. Nakamura *et al.*, *J. Oceanogr.* **71**, 463–467 (2015).
8. R. Masunaga, H. Nakamura, H. Kamahori, K. Onogi, S. Okajima, *Sci. Online Lett. Atmos.* **14**, 6–13 (2018).
9. M. Latif, T. P. Barnett, *Science* **266**, 634–637 (1994).
10. M. Latif, T. P. Barnett, *J. Clim.* **9**, 2407–2423 (1996).
11. N.-E. Omrani *et al.*, *Sci. Rep.* **9**, 3014 (2019).
12. R. Zhang, *Geophys. Res. Lett.* **35**, L20705 (2008).
13. N. J. Mantua, S. R. Hare, Y. Zhang, J. M. Wallace, R. C. Francis, *Bull. Am. Meteorol. Soc.* **78**, 1069–1079 (1997).
14. M. Yamamoto, T. Oba, J. Shimamune, T. Ueshima, *Geophys. Res. Lett.* **31**, L16311 (2004).
15. A. Paul, C. Schäfer-Neth, *Paleoceanography* **18**, 1058 (2003).
16. D. Smirnov, M. Newman, M. A. Alexander, Y.-O. Kwon, C. Frankignoul, *J. Clim.* **28**, 1126–1147 (2015).
17. L. Wu *et al.*, *Nat. Clim. Chang.* **2**, 161–166 (2012).
18. R. C. Wills, T. Schneider, J. M. Wallace, D. S. Battisti, D. L. Hartmann, *Geophys. Res. Lett.* **45**, 2487–2496 (2018).
19. K. A. Kelly, S. Dong, in *Earth's Climate: The Ocean-Atmosphere Interaction*, C. Wang, S. P. Xie, J. A. Carton, Eds., vol. 147 of *Geophysical Monograph Series* (American Geophysical Union, 2004), pp. 347–363.
20. R. W. Reynolds *et al.*, *J. Clim.* **20**, 5473–5496 (2007).
21. D. P. Dee *et al.*, *Q. J. R. Meteorol. Soc.* **137**, 553–597 (2011).
22. B. Qiu, *J. Phys. Oceanogr.* **30**, 1486–1502 (2000).
23. S. Dong, K. A. Kelly, *J. Phys. Oceanogr.* **34**, 1214–1231 (2004).
24. S. Kida *et al.*, in *"Hot Spots" in the Climate System*, H. Nakamura, A. Isobe, S. Minobe, H. Mitsudera, M. Nonaka, T. Suga, Eds. (2016), pp. 1–30.
25. S. P. Bishop, R. J. Small, F. O. Bryan, R. A. Tomas, *J. Clim.* **30**, 8207–8221 (2017).
26. R. Seager, Y. Kushnir, N. H. Naik, M. A. Cane, J. Miller, *J. Clim.* **14**, 4249–4265 (2001).
27. M. Nonaka, S.-P. Xie, *J. Clim.* **16**, 1404–1413 (2003).
28. X. Ma *et al.*, *Nature* **535**, 533–537 (2016).
29. B. Gallego, P. Cessi, *J. Clim.* **14**, 2815–2832 (2001).
30. C. Frankignoul, P. Mu¨ller, E. Zorita, *J. Phys. Oceanogr.* **27**, 1533–1546 (1997).
31. F.-F. Jin, *J. Clim.* **10**, 1821–1835 (1997).
32. R. Saravanan, J. C. McWilliams, *J. Clim.* **11**, 165–188 (1998).
33. S. Kravtsov, W. K. Dewar, P. Berloff, J. C. McWilliams, M. Ghil, *Dyn. Atmos. Oceans* **43**, 123–150 (2007).
34. B. J. Hoskins, P. J. Valdes, *J. Atmos. Sci.* **47**, 1854–1864 (1990).
35. S. Peng, W. A. Robinson, *J. Clim.* **14**, 2943–2959 (2001).
36. F. Ogawa, H. Nakamura, K. Nishii, T. Miyasaka, A. Kuwano-Yoshida, *Geophys. Res. Lett.* **39**, L05804 (2012).
37. E. P. Chassignet, D. P. Marshall, in *Ocean Modeling in an Eddying Regime*, M. W. Hecht, H. Hasumi, Eds., vol. 177 of *Geophysical Monograph Series* (American Geophysical Union, 2008), pp. 39–61.
38. A. J. Pershing *et al.*, *Science* **350**, 809–812 (2015).
39. Y. Watanabe, H. Zenitani, R. Kimura, *Can. J. Fish. Aquat. Sci.* **53**, 55–61 (1996).
40. S.-i. Saitoh, S. Kosaka, J. Iisaka, *Deep Sea Res. A* **33**, 1601–1615 (1986).
41. Y. Tian, T. Akamine, M. Suda, *Fish. Res.* **60**, 439–454 (2003).
42. H. Kuroda *et al.*, *Open J. Mar. Sci.* **7**, 62–90 (2017).
43. D. E. Amos, L. H. Koopmans, *Tables of the Distribution of the Coefficient of Coherence for Stationary Bivariate Gaussian Processes* (Sandia Corporation, vol. 483, 1963).
44. V. Eyring *et al.*, *Geosci. Model Dev.* **9**, 1937–1958 (2016).
45. T. Kohyama *et al.*, Data for Kohyama et al. (2021), data set, Zenodo (2021); https://doi.org/10.5281/zenodo.5166678.


## ACKNOWLEDGMENTS


We would like to thank Y. Masumoto, T. Tozuka, and four anonymous reviewers for useful discussions and comments. In particular, we thank the reviewer who reviewed our manuscripts three times and provided impressively useful comments every round. **Funding:** T.K. is supported by JSPS-Kakenhi grant nos. 19K23460 and 20K14554. Y.Y., H.M., and M.W. are supported by JSPS-Kakenhi grant no. 20H05729. Y.Y., H.T., and M.W. are supported by the Integrated Research Program for Advancing Climate Models (TOUGOU) grant no. JPMXD0717935457 from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. H.M. is supported by JSPS-Kakenhi grant no. 16H04048. S.K. was supported by JSPS-Kakenhi grant no. 17J04384. **Author contributions:** T.K., H.M., and S.K. conceived the general idea. T.K. led the project, analyzed the data, and wrote the manuscript except for the model description. Y.Y. performed the AGCM runs and analyzed the MIROC6subhires and MIROC6 pictl/AGCM runs. S.K. performed the idealized model runs, analyzed the output, and wrote the idealized model description in the materials and methods. H.T. led the MIROC6 model project, developed the MIROC6subhires model, performed the MIROC6subhires and MIROC6 pictl runs, and wrote the MIROC6 model description in the materials and methods. M.W. supervised the MIROC6 model project and conceived the application of the SVD analysis. All authors discussed the results and contributed to the final manuscript. **Competing interests:** All authors declare no competing interests. **Data and materials availability:** Observed SST data are downloaded from the National Oceanic and Atmospheric Administration (NOAA) Optimum Interpolation SST (OISST) (*20*) available at https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.html. The high-resolution OISST data available at www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.highres.html are also used. Observational zonal wind fields at 850 hPa and 2-m air temperature are from the European Center for Medium range Weather Forecasting (ECMWF) ERA-Interim reanalysis data (*21*) at http://apps.ecmwf.int/datasets/data/interim-full-moda/levtype=sfc/. The SST output of the GFDL models is from the Coupled Model Intercomparison Project Phase 6 (CMIP6) data (*44*), available at the website of the World Climate Research Programme (https://esgf-node.llnl.gov/search/cmip6/). The data from the idealized model and the MIROC6 models used in this study are available from Zenodo (*45*).


## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abh3295
Materials and Methods
Supplementary Text
Figs. S1 to S6
Tables S1
References (*46*–*51*)


2 March 2021; accepted 23 August 2021
10.1126/science.abh3295

CELL CYCLE

# $G_1$ cyclin–Cdk promotes cell cycle entry through localized phosphorylation of RNA polymerase II

Mardo Kõivomägi[1], Matthew P. Swaffer[1], Jonathan J. Turner[1], Georgi Marinov[2], Jan M. Skotheim[1]*

**Cell division is thought to be initiated by cyclin-dependent kinases (Cdks) inactivating key transcriptional inhibitors. In budding yeast, the $G_1$ cyclin Cln3-Cdk1 complex is thought to directly phosphorylate the Whi5 protein, thereby releasing the transcription factor SBF and committing cells to division. We report that Whi5 is a poor substrate of Cln3-Cdk1, which instead phosphorylates the RNA polymerase II subunit Rpb1's C-terminal domain on $S_5$ of its heptapeptide repeats. Cln3-Cdk1 binds SBF-regulated promoters and Cln3's function can be performed by the canonical $S_5$ kinase Ccl1-Kin28 when synthetically recruited to SBF. Thus, we propose that Cln3-Cdk1 triggers cell division by phosphorylating Rpb1 at SBF-regulated promoters to promote transcription. Our findings blur the distinction between cell cycle and transcriptional Cdks to highlight the ancient relationship between these two processes.**

The eukaryotic cell cycle is driven by a series of cyclin-dependent kinase (Cdk) complexes that promote cell cycle progression by phosphorylating key substrates. The first step of the cell cycle, from the $G_1$ to the S phase, is thought to require the inactivation of a transcriptional inhibitor. In human cells, cyclin D-Cdk4 and Cdk6 complexes phosphorylate the retinoblastoma protein Rb, and in budding yeast, Cln3-Cdk1 is thought to phosphorylate the transcriptional repressor Whi5. This results in the activation of the E2F and SBF transcription factors in animal and yeast cells, respectively, that commit cells to cell cycle entry through positive-feedback loops in which downstream cyclin-Cdk complexes hyperphosphorylate Whi5 and Rb (Fig. 1A) (*1*). However, in early to mid $G_1$, cyclin D-Cdk4,6 constitutively hypophosphorylates Rb (*2*), which is unlikely to account for its complete inactivation (*3*), and Whi5 phosphorylation by Cln3-Cdk1 has not been observed in vivo. Thus, key mechanistic aspects of the eukaryotic $G_1$-S transition model remain either unknown or untested.

To examine the prevailing model that Cln3-Cdk1 phosphorylates Whi5 (*4*, *5*), we used Phos-tag–supplemented SDS–polyacrylamide gel electrophoresis (*6*) to separate distinct phospho-isoforms of Whi5, including multiple hypophosphorylated species not previously observed (Fig. 1B and fig. S1, B and F). We examined early $G_1$ cells collected by elutriation as well as cells released from $G_1$ pheromone arrest. Whi5 hypophosphorylation is slightly reduced upon release from pheromone but is otherwise constant throughout early and mid $G_1$ (Fig. 1B and fig. S1, A to F). Subsequently, at the $G_1$-S transition, Whi5 becomes fully hyperphosphorylated, presumably by the newly synthesized Cln1/2-Cdk1 complexes (*4*, *7*). To determine whether the early $G_1$ hypophosphorylation of Whi5 is caused by Cln3-Cdk1, we conditionally depleted Cln3 from the nucleus in early $G_1$ using the rapamycin-dependent "anchor-away" technique (*8*). After Cln3 depletion, Whi5 hypophosphorylation was similar to that in untreated early and mid $G_1$ cells (Fig. 1B and fig. S1, A to E). We observed the same pattern in wild-type (WT) and *cln3*Δ cells after release from pheromone (fig. S1, F and G). Rapid Whi5 hyperphosphorylation in late $G_1$ is delayed by Cln3 depletion or *cln3*Δ, as reported previously (*9*, *10*) (Fig. 1B and fig. S1, B, F, and G), likely because of Cln1/2-Cdk1 expression being delayed in the absence of upstream Cln3-Cdk1 activity. On the basis of these data and additional supporting experiments (fig. S1, H to L), we propose that Cln1/2-Cdk1 complexes complete Whi5 hyperphosphorylation in late $G_1$, but Whi5 hypophosphorylation in early $G_1$ is not predominantly caused by Cln3-Cdk1.

If Cln3 promotes cell cycle entry, but not through Whi5 phosphorylation, it is possible that Cln3 and Whi5 act as separate inputs regulating SBF. To test this, we performed chromatin immunoprecipitation sequencing (ChIP-seq) for Cln3 and the SBF components Swi4 and Swi6. Cln3 peaks were found at 68 gene promoters, 67 of which were also bound by SBF (table S3), including key SBF-binding sites in the *CLN1* and *CLN2* promoters (Fig. 1C and fig. S2, A, B, D, and E) (*9*). Cln3 localization to SBF-binding sites did not depend on Stb1 or Whi5, Cln3's previously assumed targets (Fig. 1C and fig. S2D). Consistent with the separate input model, *cln3*Δ*whi5*Δ cells are larger than *whi5*Δ cells (*4*, *5*), Cln3-dependent activation of the *CLN2* promoter precedes loss of Whi5 protein from the promoter (*9*), and a hyperactive *CLN3* allele (*CLN3ΔC*) reduces cell size more than *whi5*Δ alone (Fig. 1D).

That Whi5 is so poorly phosphorylated by Cln3-Cdk1 in vivo indicates that Cln3-Cdk1 may phosphorylate a different target involved in SBF-dependent transcription. To explore this, we examined Cln3-Cdk1 kinase activity toward SBF-interacting proteins in vitro. To purify Cln3-Cdk1, we fused *CLN3* to *CDK1* with a glycine-serine linker (*CLN3-L-CDK1*) because Cln3 is not as tightly bound to Cdk1 as other cyclins. When the endogenous *CLN3* was replaced by this fusion, cells exhibited no cell cycle defects, and the removal of the linker after purification had little effect in vitro (figs. S3, A and H to J, and 4, A to C). Thus, this fusion complex appeared to function similarly to the WT complex. We note that kinase activity detected in a previously reported purification of Cln3-Cdk1 from insect cells (*5*), which we repeated, was not caused by Cln3 because activity was still present in a control lacking Cln3 (see the materials and methods; fig. S3, B and C). Consistent with our in vivo data, Cln3-L-Cdk1 poorly phosphorylated Whi5 in vitro, whereas Cln2-L-Cdk1 readily hyperphosphorylated Whi5 (Fig. 1E). Moreover, Cln3-L-Cdk1 was also a very poor kinase toward the SBF-associated proteins Swi6, Stb1, and Msa1 in vitro (Fig. 1F) (*10*, *11*).

Cln3-Cdk1's lack of in vitro activity against SBF-associated proteins and the model Cdk substrate H1 (Fig. 1F) suggested that Cln3 may promote the $G_1$–S transition independently of Cdk1 kinase activity. To test this, we fused *CLN3* to a catalytically inactive (kinase-dead, KD) *CDK1* allele (*12*) (*CDK1*$^{KD}$; fig. S3D). Although *CLN3-L-CDK1*$^{KD}$ rescued the effects of deleting *CLN3*, immunoprecipitation revealed endogenous Cdk1 bound to Cln3-L-Cdk1 (fig. S3C). To prevent this, we introduced a cyclin box mutation (CBM) in *CLN3* (*13*) (*CLN3*$^{CBM}$; Fig. 2A). Replacing *CLN3* with *CLN3*$^{CBM}$*-L-CDK1* resulted in WT-sized cells, indicating that the fusion allows Cln3$^{CBM}$ to activate Cdk1 (Fig. 2B). *CLN3*$^{CBM}$*-L-CDK1*$^{KD}$ cells were as large as *cln3*Δ cells, suggesting that Cln3-Cdk1 kinase activity is indeed required for Cln3 function (Fig. 2B). To confirm this, we examined Cln3-Cdk1 function in a simplified Cdk network driven exclusively by *CLN3-L-CDK1, CLB5-L-CDK1*, and *CLB2-L-CDK1* fusions (Fig. 2, C and D). In this background, the addition of *CLN3-L-CDK1*$^{KD}$ instead of *CLN3-L-CDK1* was insufficient for proliferation, confirming that Cln3-Cdk1's kinase activity promotes the $G_1$-S transition (Fig. 2, C and D, and fig. S3, E to G).

Having established that Cln3-Cdk1's kinase activity promotes the $G_1$-S transition, we next sought to identify its substrates through a candidate-based in vitro screen. We measured the specificity of purified Cln3-L-Cdk1 and other yeast cyclin-Cdk1 complexes toward >20 candidate targets (Fig. 3, A and B; fig. S4, A to D; and table S4). By far the most specific

[1]Department of Biology, Stanford University, Stanford, CA 94305, USA. [2]Department of Genetics, Stanford University, Stanford, CA 94305, USA.
*Corresponding author. Email: skotheim@stanford.edu

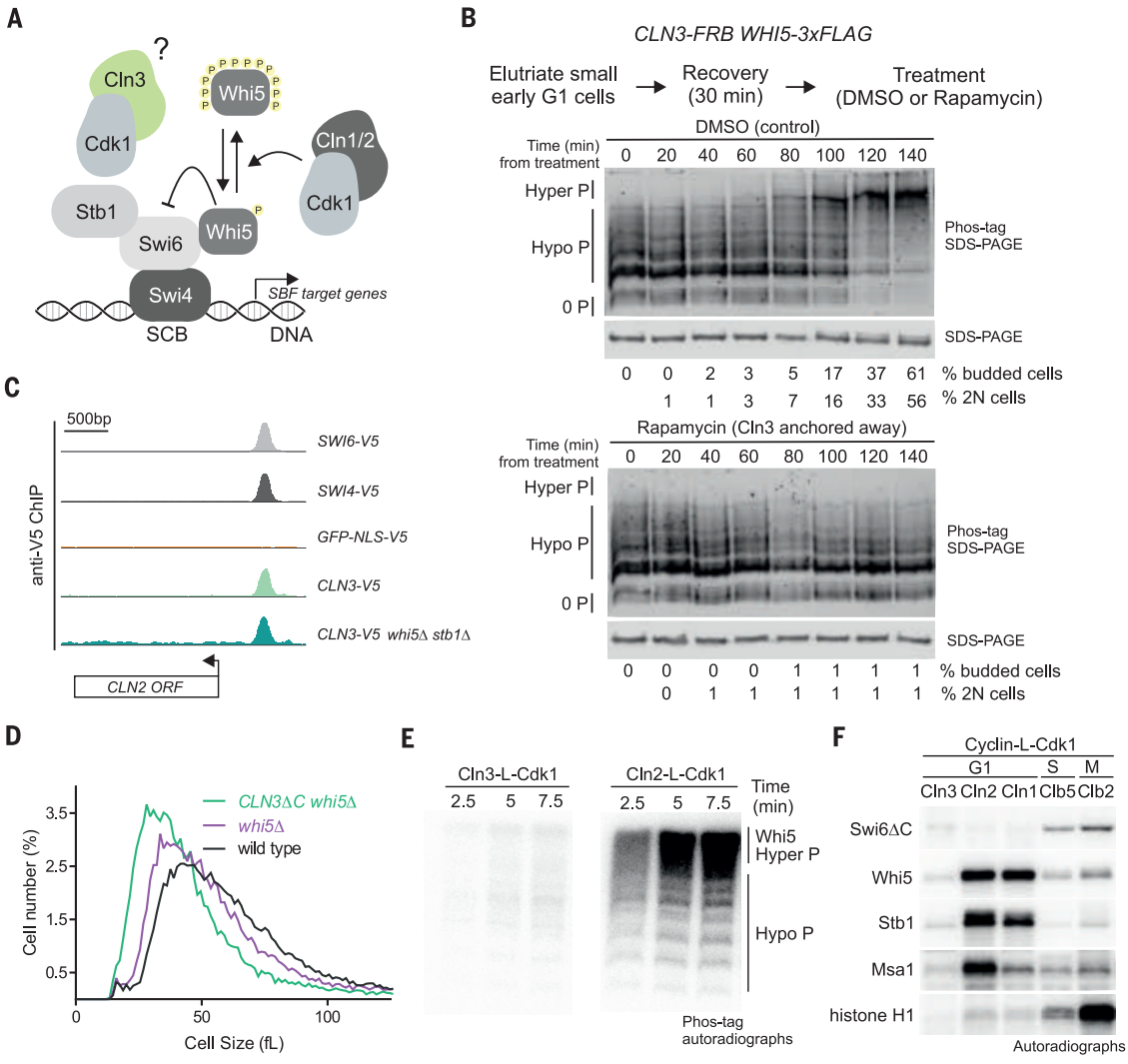



**Fig. 1. Whi5 is a poor substrate of Cln3-Cdk1.** (**A**) Budding yeast $G_1$-S network. (**B**) Immunoblots showing Whi5-3xFlag phosphorylation in $G_1$ cells collected by centrifugal elutriation treated with rapamycin to deplete Cln3-FRB from the nucleus or with dimethyl sulfoxide (see fig. S1, A to E). (**C**) Anti-V5 ChIP-seq signal of the indicated genotypes at the *CLN2* locus. (**D**) Cell size distributions for the indicated genotypes. (**E** and **F**) Autoradiographs of in vitro phosphorylation of SBF-interacting proteins and histone H1 by the indicated cyclin-L-Cdk1 complexes.

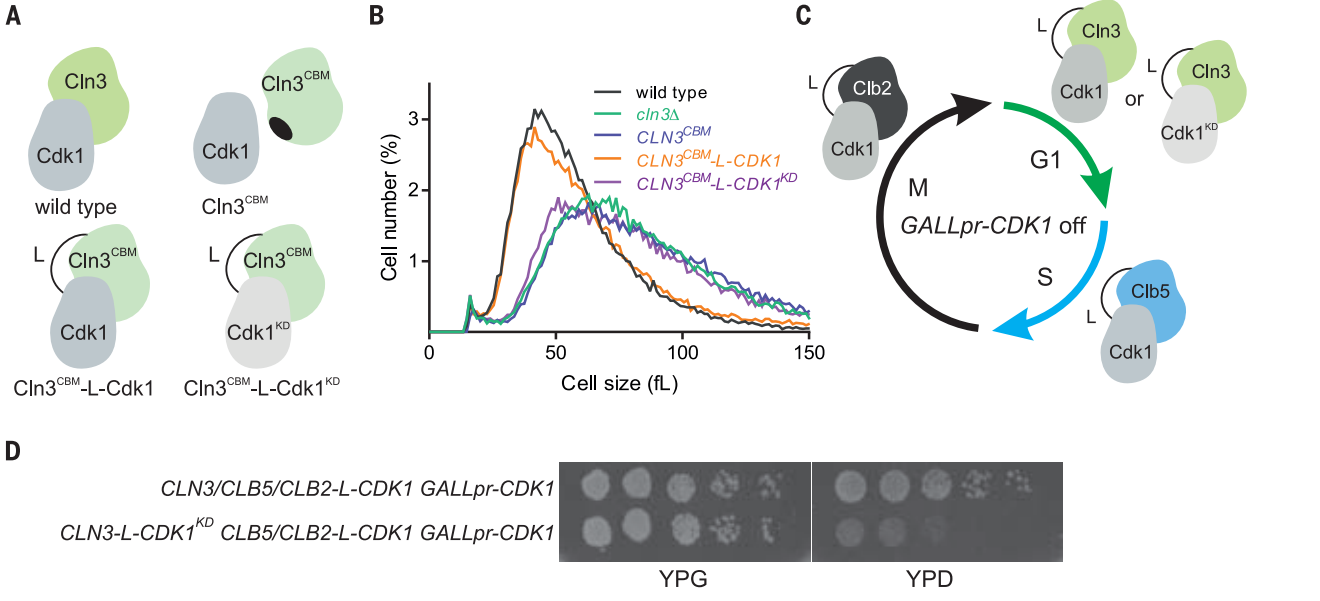



**Fig. 2. Cln3-Cdk1 kinase activity promotes $G_1$-S.** (**A**) Schematic of the different Cln3-Cdk1 complexes used. The CBM mutant cannot bind Cdk1 unless fused through the linker L. (**B**) Cell size distributions for the indicated genotypes. (**C**) Schematic of the experimental design for (D). The endogenous *CDK1* promoter was replaced with a *GALLpr*. *CLN3*, *CLB5*, and *CLB2* are fused to the indicated *-L-CDK1*. (**D**) Spot viability assays of the indicated genotypes on YPG (*GALLpr* ON) or YPD (*GALLpr* OFF).

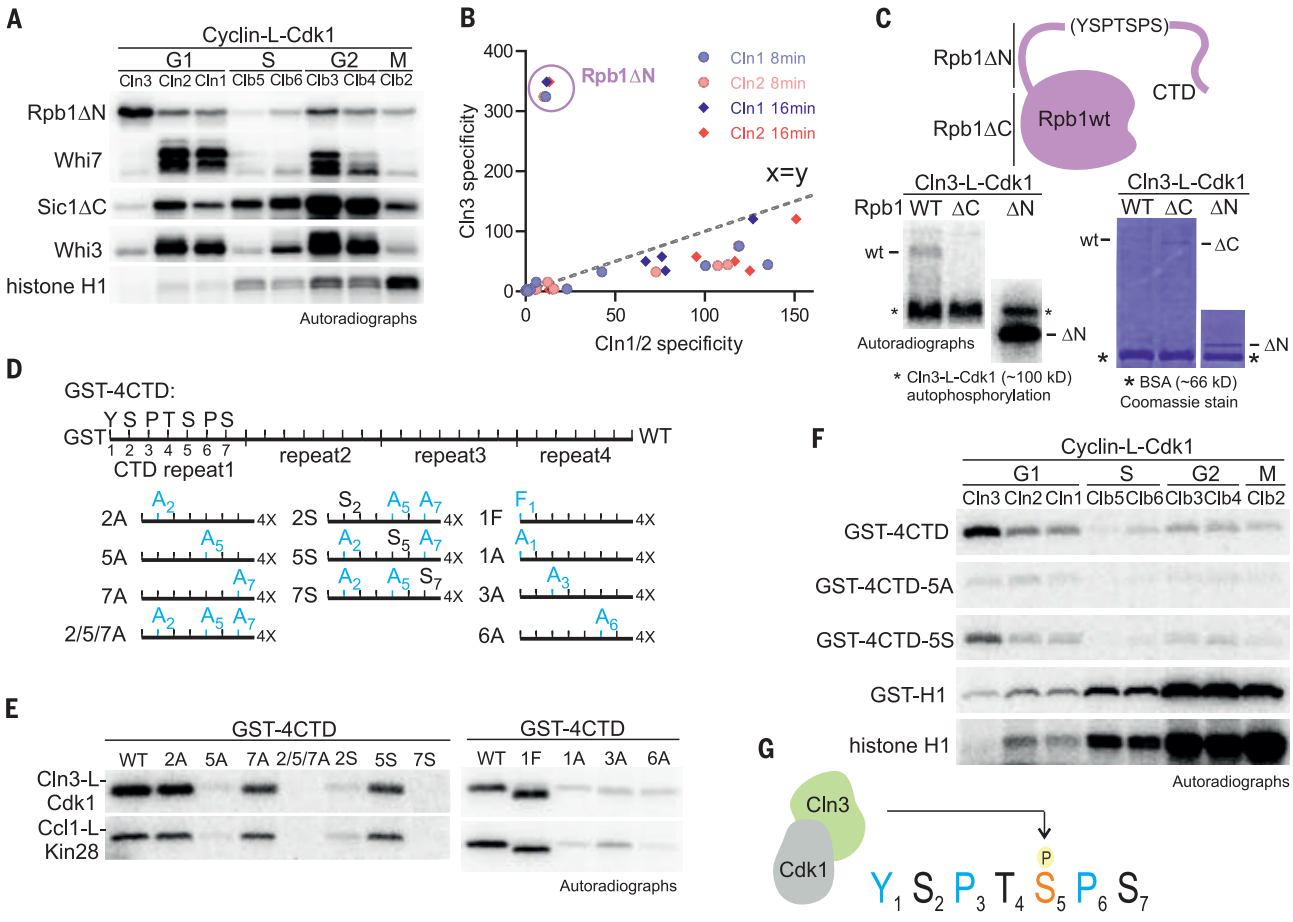



**Fig. 3. Cln3-Cdk1 phosphorylates $S_5$ in Rpb1's CTD repeats.** (**A**) Autoradiographs of in vitro phosphorylation of candidate substrates by the indicated cyclin-L-Cdk1 complexes. (**B**) Cyclin-L-Cdk1 specificity is the ratio of activity toward a substrate relative to its activity toward histone H1. Each point corresponds to one substrate. (**C**) Autoradiographs of in vitro phosphorylation by Cln3-L-Cdk1 of Rpb1 truncations (WT denotes full length; ΔC denotes 1453 N-terminal amino acids; ΔN denotes 280 C-terminal amino acids). Note that the low purification yield of WT (~192 kD) results in reduced signal. (**D** to **F**) Autoradiographs of in vitro phosphorylation of indicated GST-CTD substrates by the indicated kinases. (**G**) Schematic showing $S_5$-specific phosphorylation by Cln3-Cdk1.

target for Cln3-L-Cdk1 was the RNA polymerase II subunit Rpb1 (Fig. 3C), which contains a C-terminal unstructured region (CTD) with multiple heptapeptide repeats ($Y_1S_2P_3T_4S_5P_6S_7$) (*14*). Truncations to isolate Rpb1's unstructured C-terminal region, and then to remove the regions on either side of the CTD repeats, did not reduce phosphorylation. This implies that Cln3-Cdk1 targets residues inside the repeats that are also targeted by the canonical transcriptional kinases (except Bur1) (*15*) (fig. S4G).

Phosphorylation of the different residues within these heptad repeats by the canonical transcriptional kinases regulates transcriptional initiation, elongation, and termination (*14*). We therefore sought to identify Cln3-Cdk1's specific target residues by generating a series of model CTD substrates (Fig. 3D). Only mutation of the serine 5 residue prevented phosphorylation by Cln3-L-Cdk1. Conversely, the addition of $S_5$ to repeats lacking all serines restored phosphorylation (Fig. 3E). Phosphorylation was decreased by $Y_1$, $P_3$, and $P_6$ mutations but not by $T_4$ mutation (Fig. 3E and fig. S4E). The effect of the $P_6$ alanine substitution was expected because Cdk1 is a proline-directed kinase, whereas the $P_3$ requirement is similar to the -2P enhancement of Cln2-Cdk1 activity (*16*). Thus, Cln3-Cdk1, in contrast to other cell cycle cyclin-Cdk1 complexes, specifically phosphorylates $S_5$ in vitro, and this depends on the local amino acid sequence (Fig. 3, F and G).

That Cln3-Cdk1 colocalizes with SBF on the genome and functions as an $S_5$ CTD kinase in vitro suggests that Cln3-Cdk1 should be responsible for $S_5$ phosphorylation specifically at SBF-regulated promoters. To test this, we depleted Cln3 from the nucleus (*8*) and then performed ChIP-seq against Ceg1, a protein for which chromatin association depends on binding $pS_5$ (*17*) (Fig. 4A). This resulted in reduced Ceg1-FLAG signal specifically at SBF-regulated promoters (Fig. 4, B to D). Because the time scale required for nuclear depletion (>10 min) cannot distinguish between direct and indirect effects, we also performed a time course after chemical-genetic inhibition of $CDK1^{as}$ (Fig. 4E). We observed an immediate and rapid decrease in the Ceg1-FLAG signal at SBF-regulated genes already within 1 min (Fig. 4, F to H, and fig. S7, A to D). We estimated SBF $pS_5$ turns over with a half-life of ~1 min (fig. S7D), comparable to that previously determined for other Cdk1 substrates (*18*). Although Rpb3 at SBF-regulated genes also decreases after 1NM-PP1 treatment, this occurred more slowly, suggesting that the loss of Ceg1 is unlikely to be caused by a loss of polymerase occupancy (fig. S7, E and F). Further supporting Cln3-Cdk1's role as an $S_5$ kinase, *cln3*Δ reduced global $S_5$ phosphorylation in whole-cell extracts (fig. S5).

We reasoned that if Cln3-Cdk1 functions as an Rpb1-CTD-$S_5$ kinase at SBF-regulated genes, then we should be able to bypass the requirement for Cln3 by providing an alternative $S_5$ kinase at SBF-regulated promoters. To test this, we used a rapamycin-dependent binding system (*8*) (fig. S6, A and B) to conditionally recruit a fusion protein of the canonical CTD $S_5$ kinase, Ccl1-L-Kin28, to SBF (Fig. 4I). Recruitment of Ccl1-L-Kin28 to SBF fully rescued the

size and cell cycle phenotypes of *cln3Δ* cells, and this rescue depended on Kin28 kinase activity (Fig. 4J and fig. S6, C to I). We therefore propose that Cln3-Cdk1 promotes the first step in the budding yeast cell cycle by directly phosphorylating $S_5$ on Rpb1's CTD at SBF-regulated genes (Fig. 4K).

Our findings here add Cln3-Cdk1 to the list of transcriptional kinases that promote gene expression through CTD phosphorylation (*14*). Unexpectedly, Whi5 and the other SBF-associated factors that we tested were very poor substrates of Cln3-Cdk1, and our in vitro screen suggested that Cln3-Cdk1 has few, if any, targets other than the Rpb1. Such extreme cyclin specificity may help to order the cell cycle (*16*). If, as we suspect, Cln3 has no targets driving another cell cycle event, then the level of Cln3 can be used to exclusively modulate the $G_1$-S transition without prematurely triggering downstream cell cycle events.

Although the cell cycle Cdks and transcriptional Cdks are all part of the same kinase family, their functions are generally thought to have diverged (*19*). Our findings break this dichotomy by showing that cell cycle Cdks can directly activate transcription at specific genes. Unexpectedly, we observed that Kin28 also binds SBF promoters (fig. S2, C and D), indicating that both Kin28 and Cln3-Cdk1 cooperate to phosphorylate Rpb1, consistent with the reported genetic interactions between *KIN28* and *CDK1* (*20*). The functional overlap of cell cycle and transcriptional Cdks has also been suggested by the dual function of the Kin28 orthologs in other eukaryotes that both activate the cell cycle Cdks and function as global $S_5$ CTD kinases (*21*). In addition, Cdk1 and Cdk2 were identified as the first RNA polymerase II CTD kinases in vitro, but whether they function as such in vivo is unknown (*22*). Regardless, it is now apparent

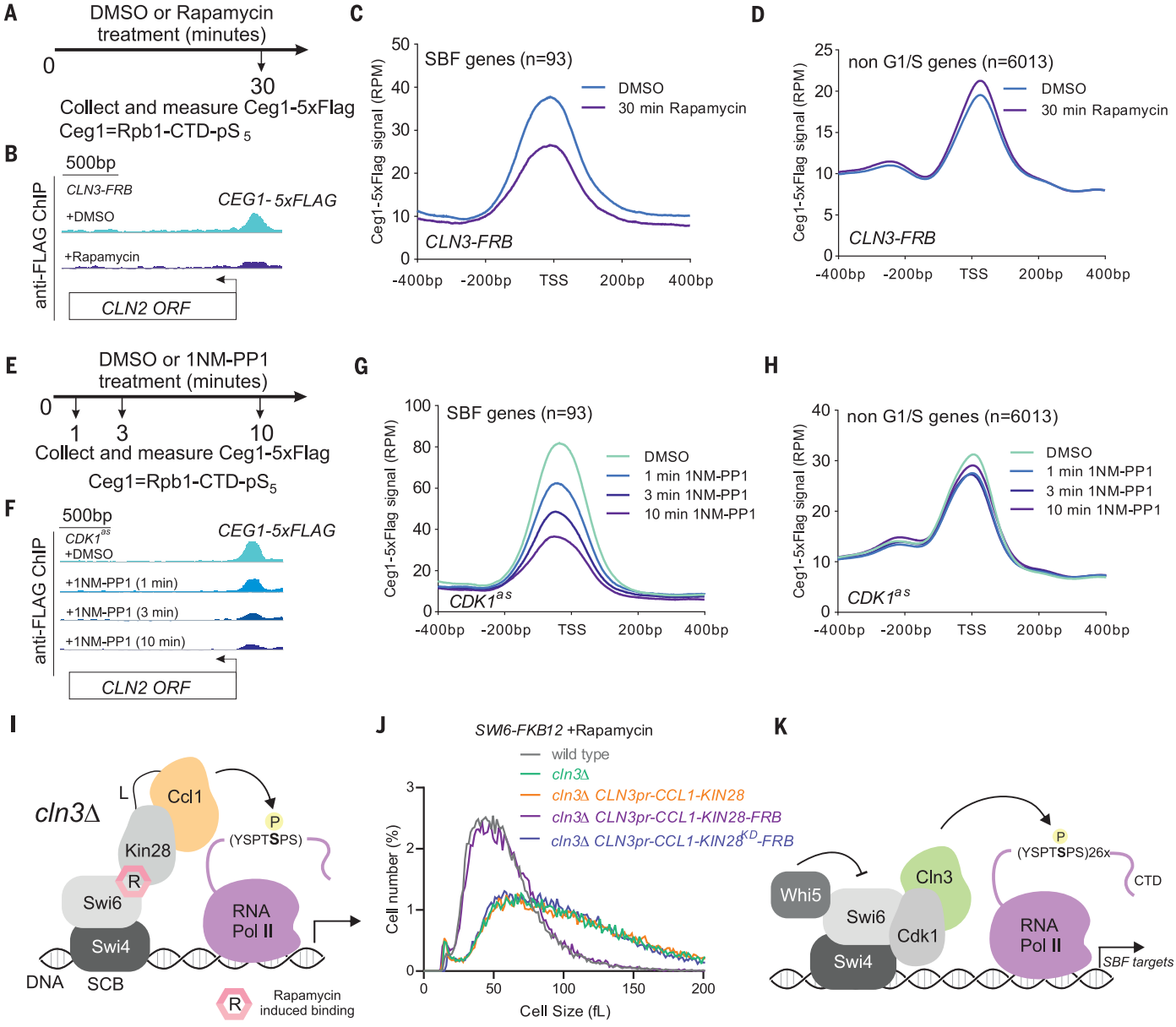



**Fig. 4. Cln3-Cdk1 drives the $G_1$-S transition by phosphorylating $S_5$ on Rpb1's CTD at SBF-regulated genes.** (**A**) Cln3-FRB is conditionally anchored away from the nucleus by rapamycin. Ceg1-5xFLAG ChIP-seq measures Rpb1 $pS_5$. (**B** to **D**) Ceg1-5xFLAG ChIP-seq signal at the *CLN2* promoter (B) and average coverage at SBF-regulated genes (C) and non–$G_1$/S genes (D). (**E** to **H**) As in (A) to (D) except $CDK1^{as}$ is inhibited by 1NM-PP1. (**I** and **J**) Cell size distributions for the indicated genotype. Ccl1-L-Kin28 fusion proteins were expressed from a *CLN3* promoter and recruited to SBF through Swi6 by rapamycin-induced binding. (**K**) Summary model.

that the two branches of Cdks that regulate cell division and transcription have overlapping functions, which suggests the possibility that their primordial ancestor regulated both processes. Such an ancient link may be reflected here in our discovery that Cln3-Cdk1 drives cell cycle progression by directly activating transcription.

## REFERENCES AND NOTES


1. C. Bertoli, J. M. Skotheim, R. A. M. de Bruin, *Nat. Rev. Mol. Cell Biol.* **14**, 518–528 (2013).
2. A. M. Narasimha *et al.*, *eLife* **3**, e02872 (2014).
3. B. R. Topacio *et al.*, *Mol. Cell* **74**, 758–770.e4 (2019).
4. R. A. M. de Bruin, W. H. McDonald, T. I. Kalashnikova, J. Yates 3rd, C. Wittenberg, *Cell* **117**, 887–898 (2004).
5. M. Costanzo *et al.*, *Cell* **117**, 899–913 (2004).
6. E. Kinoshita, E. Kinoshita-Kikuta, K. Takiyama, T. Koike, *Mol. Cell. Proteomics* **5**, 749–757 (2006).
7. J. M. Skotheim, S. Di Talia, E. D. Siggia, F. R. Cross, *Nature* **454**, 291–296 (2008).
8. H. Haruki, J. Nishikawa, U. K. Laemmli, *Mol. Cell* **31**, 925–932 (2008).
9. H. Wang, L. B. Carey, Y. Cai, H. Wijnen, B. Futcher, *PLOS Biol.* **7**, e1000189 (2009).
10. R. A. M. de Bruin, T. I. Kalashnikova, C. Wittenberg, *Mol. Cell. Biol.* **28**, 6919–6928 (2008).
11. M. Ashe *et al.*, *J. Biol. Chem.* **283**, 6040–6049 (2008).
12. L. Kao *et al.*, *Mol. Cell. Proteomics* **13**, 594–605 (2014).
13. M. E. Miller, F. R. Cross, A. L. Groeger, K. L. Jameson, *Yeast* **22**, 1021–1036 (2005).
14. K. M. Harlen, L. S. Churchman, *Nat. Rev. Mol. Cell Biol.* **18**, 263–273 (2017).
15. Y. Chun *et al.*, *Mol. Cell. Biol.* **39**, 395 (2019).
16. M. Kõivomägi *et al.*, *Mol. Cell* **42**, 610–623 (2011).
17. P. Komarnitsky, E. J. Cho, S. Buratowski, *Genes Dev.* **14**, 2452–2460 (2000).
18. M. P. Swaffer, A. W. Jones, H. R. Flynn, A. P. Snijders, P. Nurse, *Cell* **167**, 1750–1761.e16 (2016).
19. L. Cao *et al.*, *BMC Evol. Biol.* **14**, 10 (2014).
20. P. Chymkowitch *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 10450–10455 (2012).
21. R. P. Fisher, *Transcription* **10**, 47–56 (2019).
22. Y.-X. Xu, Y. Hirose, X. Z. Zhou, K. P. Lu, J. L. Manley, *Genes Dev.* **17**, 2765–2776 (2003).


## ACKNOWLEDGMENTS


We thank T. Stearns for Sf9 cells; M. Tyers for baculovirus stocks; J. Ewald, K. Schmoller, D. Morgan, S. Hahn, and S. Buratowski for yeast strains; M. Loog, J. Ferrell, F. Cross, P. Pryciak, A. Ansari, D. Morgan, and members of the Skotheim laboratory for constructive feedback; and B. Reyes Topacio for help with enzyme purifications. **Funding:** This work was supported by the National Institutes of Health (grant R35 GM134858), the Howard Hughes Medical Institute–Simons (JMS Faculty Scholars Program), the Human Frontier Science Program (postdoctoral fellowship to M.K.), and the Life Sciences Research Foundation (Simons Foundation Fellowship to M.S.). **Author contributions:** M.K., J.T., and M.S. performed experiments. G.M. and M.S. analyzed sequencing data. M.K., M.S., and J.S. conceptualized the study, designed experiments, interpreted data, wrote the manuscript, and supervised the study. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** ChIP-seq data are available at the Gene Expression Omnibus (GEO) repository under no. GSE169271. Other materials are available upon request.


## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.aba5186
Materials and Methods
Figs. S1 to S7
Tables S1 to S4
References (*23*–*38*)
MDAR Reproducibility Checklist


10 December 2019; resubmitted 22 March 2021
Accepted 24 August 2021
10.1126/science.aba5186


DEVELOPMENTAL BIOLOGY

# Bioelectric signaling and the control of cardiac cell identity in response to mechanical forces


**Hajime Fukui[1,2], Renee Wei-Yan Chow[1], Jing Xie[3], Yoke Yin Foo[4], Choon Hwai Yap[4,5], Nicolas Minc[3], Naoki Mochizuki[2], Julien Vermot[1,5]***



**Developing cardiovascular systems use mechanical forces to take shape, but how ubiquitous blood flow forces instruct local cardiac cell identity is still unclear. By manipulating mechanical forces in vivo, we show here that shear stress is necessary and sufficient to promote valvulogenesis. We found that valve formation is associated with the activation of an extracellular adenosine triphosphate (ATP)–dependent purinergic receptor pathway, specifically triggering calcium ion ($Ca^{2+}$) pulses and nuclear factor of activated T cells 1 (Nfatc1) activation. Thus, mechanical forces are converted into discrete bioelectric signals by an ATP-$Ca^{2+}$-Nfatc1–mechanosensitive pathway to generate positional information and control valve formation.**


The cardiovascular system is continuously exposed to mechanical forces such as fluid shear stresses and stretching forces generated by blood flow and heartbeat. In the heart, the mechanotransduction cascade translates forces into cellular biochemical signals in the endocardial cells (EdCs) to shape valves (*1*, *2*). These cell types present unique functional properties and are thought to differentially respond to forces depending on the heart regions to which they belong (*3*, *4*). Here, we examined the ability of mechanical forces to modulate specific endocardial fate toward valvulogenesis. We found that valvular cell identity is directed by mechanical forces through different mechanosensitive pathways: transient receptor potential (TRP)–mediated *klf2a* activation and adenosine triphosphate (ATP)–mediated $Ca^{2+}$ oscillations leading to Nfatc1 signaling in response to local high shear stresses.

Using zebrafish to allow for high-precision control of spatiotemporal mechanical parameters, we analyzed $Ca^{2+}$ dynamic patterns by live imaging of the fluorescent $Ca^{2+}$-sensor protein GCaMP7a expressed in EdCs using *Tg(fli1a:galff);(uas:GCaMP7a)* (*5*). We found that EdCs displayed $Ca^{2+}$ oscillations almost exclusively in the EdCs of the atrioventricular canal (AVC) that will form the AV valve (Fig. 1, A to C; fig. S1, A to E; and movies S1 and S2) between 32 and 102 hours postfertilization (hpf). To verify that the $Ca^{2+}$ oscillations were associated with valvulogenic differentiation, we searched for the decoder and the effector of $Ca^{2+}$ signaling. Nuclear factor of activated T cells 1 (Nfatc1) is an established $Ca^{2+}$-sensitive transcription factor known to modulate the endocardial-to-mesenchymal transition and subsequent heart valve morphogenesis (fig. S2, A and B) (*6*–*11*). We generated a Nfat binding element reporter line, *Tg(4xnfbr:d2EGFP)*, which expresses d2EGFP in response to the binding of nuclear-localized Nfat protein (fig. S2, C and D). Nfat reporter expression was observed in EdCs of the AVC from 34 hpf onward (Fig. 1, D and E, and fig. S2E), with a similar time course observed for the $Ca^{2+}$ response (Fig. 1, C and E). We confirmed these results by tracking Nfat nuclear relocalization of an endothelial- and endocardial-specific green fluorescent protein (GFP)–Nfatc1 reporter (fig. S2F). Thus, both spatial and temporal regulation of $Ca^{2+}$ signaling and Nfat activation are features of the valvulogenic EdCs.

To determine whether the observed $Ca^{2+}$ oscillations in EdCs are force responsive, we assessed the $Ca^{2+}$ signals in hearts stopped by MS-222 or p-amino blebbistatin treatment. We found that $Ca^{2+}$ oscillations were abrogated in EdCs when the heart stopped and recovered after the heartbeat was restarted (fig. S3, A to C). Because lower hematocrit affects shear stress and flow profile in the AVC (*12*, *13*), we analyzed the bloodless mutant Vlad Tepes (*gata1a*) (*14*). Both *gata1a* mutants and morphants exhibited fewer $Ca^{2+}$ spikes and a decrease in Nfat activity (fig. S3, D to G). Thus, both $Ca^{2+}$ oscillations and Nfat activation are stimulated in response to mechanical forces generated by the beating heart.

We next changed the mechanical force distribution in the living heart by artificially changing the boundary conditions and thus the flow forces in the system (Fig. 2A). We inserted a 30- to 60-μm-diameter agarose-based magnetic bead into the cardiac lumen by microsurgery without perturbing heart function (Fig. 2, A and B, and movie S3). To quantify the force generated around the bead, we performed a three-dimensional (3D) simulation of the wall shear stress and pressure around the bead


[1]Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Centre National de la Recherche Scientifique UMR7104, Institut National de la Santé et de la Recherche Médicale U1258 and Université de Strasbourg, Illkirch, France. [2]Department of Cell Biology, National Cerebral and Cardiovascular Center Research Institute, Suita, Japan. [3]Université de Paris, Centre National de la Recherche Scientifique UMR7592, Institut Jacques Monod, Paris, France. [4]Department of Biomedical Engineering, National University of Singapore, Singapore, Singapore. [5]Department of Bioengineering, Imperial College London, London, UK.
***Corresponding author. Email: jvermot@ic.ac.uk**

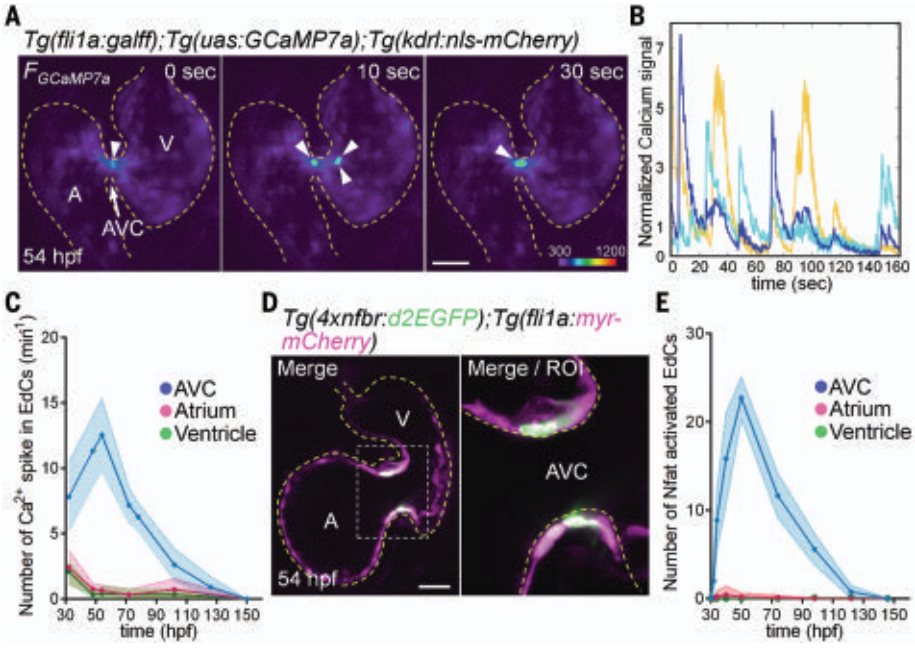



**Fig. 1. Endogenous $Ca^{2+}$ influx and Nfat signal activation occur in EdCs of the AVC during valvulogenesis.** (**A**) Representative single-scan beating heart images of the *Tg(fli1a:galff);Tg(uas:GCaMP7a);Tg(kdrl:nls-mCherry)* embryo. Panels are the heatmap-colored GCaMP7a images at the indicated time points. Arrowheads and dashed lines indicate GCaMP7a signal and outline of the endocardium, respectively. (**B**) Normalized $Ca^{2+}$ dynamics determined by $F_{GCaMP7a}/F_{nls\text{-}mCherry}$. Each colored line indicates the ratiometric change of individual cells (three cells) shown in movie S2. (**C**) The number of $Ca^{2+}$ spikes in EdCs of the indicated region from 32 to 150 hpf ($n > 5$). (**D**) Confocal single-scanned heart images of *Tg(4xnfbr:d2EGFP);Tg(fli1a:myr-mCherry)* embryo ($n = 20$). (**E**) Number of Nfat reporter–activated EdCs from 30 to 146 hpf ($n = 14$). A, atrium; V, ventricle. Scale bars, 30 μm.

using an image-based computational fluid dynamics simulation with dynamic wall motion (*15*) (Fig. 2C and fig. S4, A to C). Shear stress and pressure were increased in the vicinity of the bead (movie S4) and were close to the endogenous forces generated where the valve normally forms (*15*–*17*). In vivo, we observed $Ca^{2+}$ spikes in the EdCs next to the bead when trapped in the atrium or in the ventricle (Fig. 2D; fig. S4, D to G; and movies S5 and S6). To more precisely control shear forces, we directly manipulated the bead within the heart using magnetic tweezers (Fig. 2E). The force generated ranged from 100 to 2000 pN (fig. S4, H and I), a stimulus that is slightly higher than endogenous wall shear stresses generated in the AVC (*15*–*17*). Although the magnet probe alone did not affect $Ca^{2+}$ influx in EdCs, direct pulling of the magnetic bead led to the generation of $Ca^{2+}$ influx in the EdCs of the arrested heart (Fig. 2, F to H; fig. S4, J and K; and movie S7). Thus, $Ca^{2+}$ fluxes are induced in the EdCs by wall shear stress and associated wall pressure, whereas cardiac contractility is not necessary.

We assessed the biological relevance of the localized response initiated by bead grafting to find out whether it is associated with the formation of ectopic valves. Sixteen to 20 hours after bead grafting in the atrium, valve-like EdC clusters were observed near the bead. The valve-like structures were visible in the atrial region that protruded from the atrial surface to the atrial cavity or the cardiac jelly (Fig. 3A and fig. S5A). A grafted bead triggered the mislocalized protrusions into cardiac jelly (fig. S5A), which constitutes the first sign of valve morphogenesis in zebrafish (*18*). We additionally found that the grafted bead induced the expression of *twist1b* in the valve-like structures (Fig. 3B and fig. S6, A to D) and the ectopic expression of *klf2a*, *klf4a*, and *wnt9b* in the EdCs surrounding it (fig. S7, A to C and F). These observations indicate that ectopic valve–like structures share similarities with endogenous valves. In addition, Nfat activity was ectopically induced in EdCs located near the bead (fig. S5, B and C), which is consistent with the fact that Nfatc1 nuclear localization is activated in response to ectopic forces and reduced when heart contractility is affected (fig. S5, D to G). Nuclear translocation of Nfatc1 was inhibited by a chemical blocker of calcineurin (FK506) acting on the Ser/Thr dephosphorylation site of Nfat (fig. S5D). Thus, the nuclear localization of Nfatc1 is promoted by increased mechanical force in a Nfatc1 Ser/Thr dephosphorylation–dependent manner. Neither *klf2a* nor *klf4a* expression was affected when calcineurin was inhibited by either FK506 or VIVIT, a selective peptide inhibitor of Nfat (*19*) (fig. S7, C to F), suggesting that the $Ca^{2+}$-Nfat signaling pathway acts independently of the *klf2a* mechanotransduction pathway. Conversely, *egr1*, a known mechanosensitive gene expressed in the AVC (*20*), was down-regulated by calcineurin inhibition (fig. S7, G and H), suggesting that *egr1* acts downstream of Nfatc1. Thus, mechanical forces are key to specifying heart valve position during development by modulating both the Klf2a- and the Nfatc1-signaling pathways.

To characterize the factors responsible for the force-dependent $Ca^{2+}$-Nfat signal activation, we first focused on established mutant lines for the stretch-sensitive channels Trpp2, Trpv4, Piezo1, Piezo2a, and Piezo2b (*17*, *21*, *22*) and for ciliogenesis (*iguana*, Dzip1) (*23*), which have been implicated in flow sensing in the cardiovascular system (*24*). We found that all of these mutants showed normal $Ca^{2+}$ activation (fig. S8, A and B) as well as for embryos treated with gadolinium ion ($Gd^{3+}$), a nonspecific stretch-sensitive channel blocker (fig. S8C). Alternatively, ATP is released in cultured endothelial cells in response to mechanical stimuli (*25*). As a response, ATP activates $Ca^{2+}$ signaling through the purinergic receptor P2X channels (*26*) (Fig. 4A). Accordingly, $Ca^{2+}$ influx in EdCs of the AVC was abolished when cytosolic ATP levels were depleted by oligomycin A, a mitochondrial ATP synthase inhibitor (fig. S8D). We investigated whether extracellular ATP alters $Ca^{2+}$ influx in EdCs by microinjecting oligomycin A, apyrase (an ATPase), and ATPγS (an ATP analog) into the cardiac lumen (Fig. 4B, and movie S8), and found that the $Ca^{2+}$ influx was significantly affected in response to changes in extracellular ATP levels (Fig. 4C and fig. S8, E to I). Transcripts of the ATP receptors *p2rx1*, *p2rx4a*, and *p2rx7* were present in EdCs of the heart (fig. S9, A and B). Embryos treated with antagonists of P2X1, P2X4, and P2X7; embryos overexpressing a dominant-negative form of either human *P2X4* (*hP2X4C353W*) or zebrafish *p2rx4a* (*zP2rx4aC365W*) mRNA; and *p2rx4a* morphants displayed a significant reduction of $Ca^{2+}$ spikes in EdCs (Fig. 4D and fig. S9, C to G and I). These results show that the $Ca^{2+}$ influx in EdCs is modulated by P2X channels in response to changes in extracellular ATP levels.

We next assessed whether the activation and inhibition of P2X-mediated ATP signaling affected valve formation. Embryos treated with a combination of P2X1, P2X4, and P2X7 antagonists; a combination of P2X4 with P2Y2 antagonists; a P2X4 antagonist; or a P2Y2 antagonist suppressed Nfat reporter activation, whereas the other P2X or P2Y antagonists had no effect (fig. S10, A and B). Furthermore, overexpression of a dominant-negative form and knock-down of *p2rx4a* did not affect *klf2a*

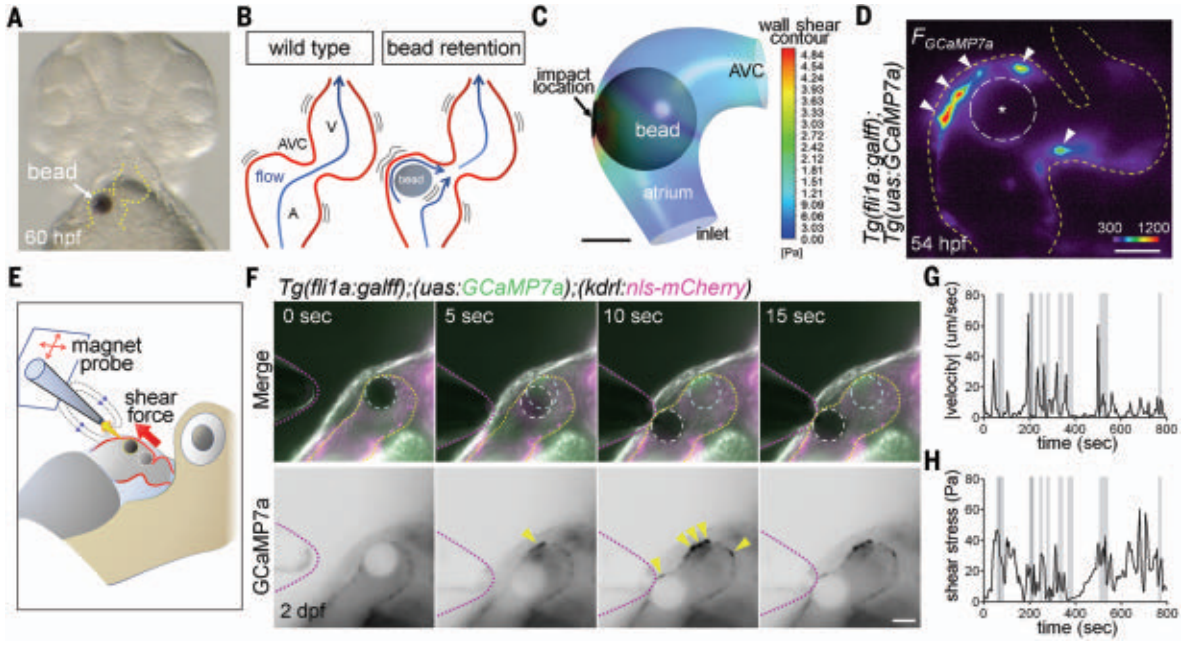



**Fig. 2. Mechanical forces directly trigger $Ca^{2+}$ influx in EdCs of zebrafish heart.** (**A**) Representative image of bead implantation into the heart. (**B**) Schematics of bead retention. (**C**) Force intensities from 3D dynamic mesh simulation of wall shear stress with a bead (clear circle) close to the impact location. (**D**) Representative GCaMP7a image of embryo grafted with a bead (asterisk). (**E**) Schematic of magnetic tweezer–based force application to a zebrafish heart. (**F**) Single-scan images of *Tg(fli1a:galff);Tg(uas:GCaMP7a);Tg(kdrl:nls-mCherry)* grafted embryo (original bead position is indicated by dashed cyan circles) pulled by the magnetic probe (dotted magenta lines), left view ($n$ = 6, yielding similar results). (**G** and **H**) Quantification of bead velocity (G) and shear stress (H) applied by the magnetic tweezer shown in movie S7. Gray shading indicates the period of activated GCaMP7a signal in EdCs. Arrowheads indicate the GCaMP7a signal and yellow dashed lines the endocardium. Scale bars, 15 μm in (C) and 30 μm in (D) and (F).

and *klf4a* expression but significantly suppressed Nfat reporter expression (Fig. 4E and fig. S9, H, J, and K), correlating well with the abnormal $Ca^{2+}$ activation (Fig. 4D and fig. S9G). In addition, embryos displayed abnormal valve structures with incomplete morphogenesis of the valve, absence of leaflet formation (Fig. 4F), and mislocalized Twist1b-positive cells in the endocardium when *p2rx* function was inhibited (Fig. 4F and fig. S11, A to C). EdCs were excluded from the luminal part of the valve when Nfat activity was suppressed in these cells (fig. S11, D and E), consistent with the fact that Nfatc1 is necessary for valve development in zebrafish (*7*). Although vascular endothelial growth factor (Vegf) signaling is known to activate Nfat signaling (*27*), a Vegfr2 inhibitor had no effect on either Nfat reporter expression or $Ca^{2+}$ activation in the EdCs (fig. S10, A to D). These results suggest that P2X and P2Y receptors coordinately modulate Nfat activity in the EdCs independently of VEGF signaling. In addition, we found that Nfat activity in the endocardium seemed to be unrelated to tyrosine phosphorylation (*28*) (fig. S10, A and E). Altogether, these results confirm that P2X functions act upstream of Nfat activity to control valve development in response to mechanical stimuli.

Mechanical forces have been implicated in a broad range of events during embryonic development (*29*–*31*). The presented work reframes

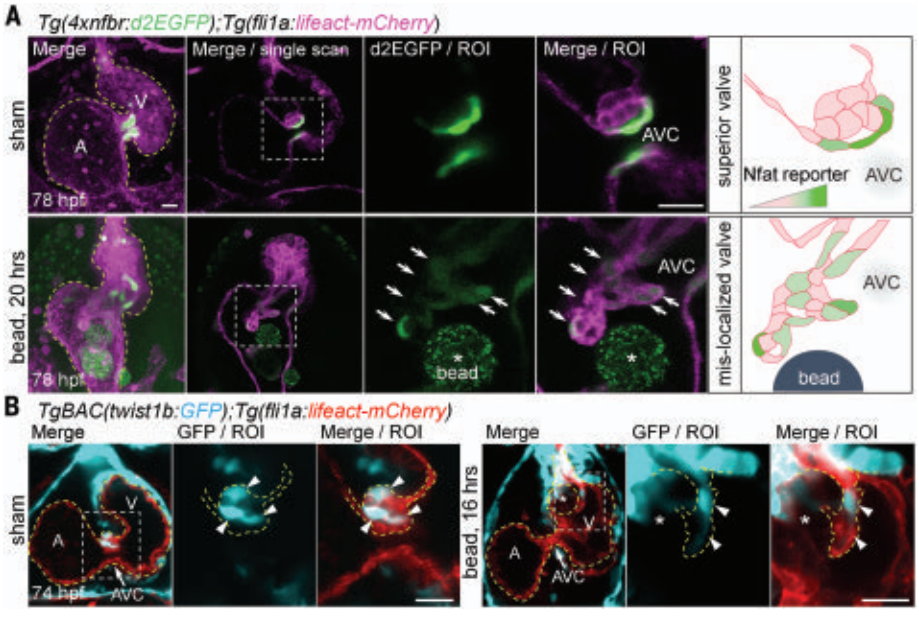



**Fig. 3. Artificial forces lead to the formation of ectopic valve–like structures.** (**A**) Confocal images of *Tg(4xnfbr:d2EGFP);Tg(fli1a:lifeact-mCherry)* embryo after sham treatment ($n$ = 11) and 20 hours after bead grafting ($n$ = 8) with the schematics of superior valve (top) and mislocalized valve (bottom) with Nfat reporter intensity ($n$ > 5, yielding similar results individually). Arrows indicate the ectopic structure. (**B**) Representative confocal images of *TgBAC(twist1b:GFP);Tg(fli1a:lifeact-mCherry)* embryo after sham treatment ($n$ = 10) and 16 hours after bead grafting ($n$ = 7). Arrowheads indicate Twist1b expression in EdCs, yellow dashed lines the outline of the endocardium, and asterisks and dashed circles the grafted bead. Scale bars, 20 μm.

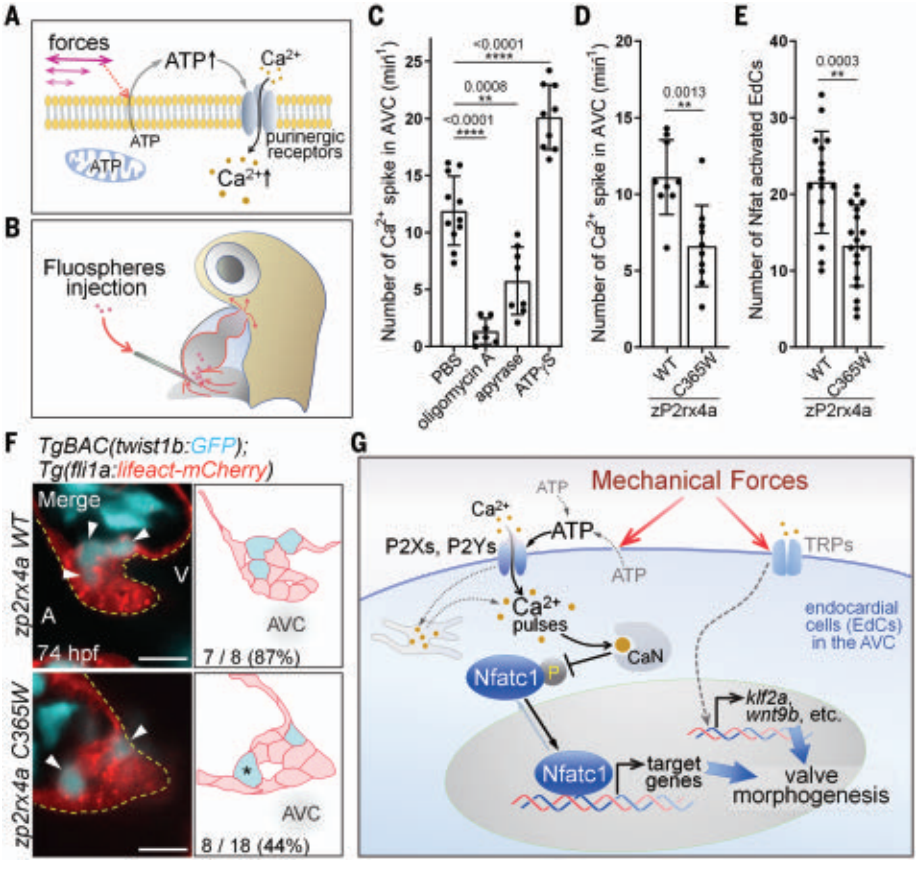



**Fig. 4. Extracellular ATP and P2rx channels function for valve development.** (**A**) ATP-dependent P2XRs $Ca^{2+}$ activation schematic. (**B**) Schematic of the microinjection approach to injecting drugs into cardiac lumen. (**C**) Number of $Ca^{2+}$ spikes in EdCs of the AVC for PBS-injected ($n$ = 11), oligomycin A–injected ($n$ = 7), apyrase-injected ($n$ = 8), and ATPγS-injected ($n$ = 9) embryos at 54 hpf. (**D** and **E**) Number of $Ca^{2+}$ influxes (D) and Nfat reporter activation (E) in EdCs of the AVC for embryos injected with *zp2rx4aWT* [$n$ = 9 in (D) and $n$ = 16 in (E)] and *zp2rx4aC365W* mRNA [$n$ = 10 in (D) and $n$ = 18 in (E)] at 54 hpf. (**F**) Confocal images of *TgBAC(twist1b:GFP);Tg(fli1a:lifeact-mCherry)* embryo injected with *zp2rx4aWT* ($n$ = 8) and *zp2rx4aC365W* mRNA ($n$ = 18). Arrowheads and percentages indicate *twist1b* expression in the EdCs and occurrence of each representative valve structure, respectively. Asterisks indicate mislocalized EdCs with ectopic *twist1b* expression. (**G**) Summary of endocardial mechanosensitive signal pathways for valve development. Note that forces trigger the ATP-transducer ($Ca^{2+}$)-decoder (NFAT) system, which is independent of the *klf2a* expression activation. Scale bars, 15 μm.

our view of how EdCs interpret mechanical forces (Fig. 4G) and supports the intriguing possibility that valve formation is mainly defined by the mechanosensitive inputs resulting from the local tissue constraints and flow forces (*7*, *18*, *32*–*34*). Piezo and Trp channels are well-established stretch-sensitive channels involved in cardiovascular morphogenesis and valve development (*17*, *21*, *35*–*37*). Our work identifies ATP as an additional mechanosensitive paracrine signal by which hemodynamic forces can direct heart valve development. Therefore, it is a compound that could be used to help grow heart valves in vitro and may be involved in congenital heart valve defects. More generally, this work shows that biological systems rely on multiple mechanosensitive pathways to precisely control morphogenetic processes. ATP release and stretch-sensitive channels are activated in response to discrete mechanical stimuli to precisely generate the permissive signal controlling the localization and timing of valve morphogenesis. We view this as a safeguard mechanism that increases the robustness of the mechanosensitive process to avoid ectopic valve formation in other regions of the heart where mechanical forces are also at work.

## REFERENCES AND NOTES


1. E. Steed, F. Boselli, J. Vermot, *Biochim. Biophys. Acta* **1863**, 1760–1766 (2016).
2. A. L. Duchemin, H. Vignes, J. Vermot, R. Chow, *Curr. Opin. Genet. Dev.* **57**, 106–116 (2019).
3. A. Paolini, S. Abdelilah-Seyfried, *Curr. Opin. Cell Biol.* **55**, 52–58 (2018).
4. F. Boselli, J. B. Freund, J. Vermot, *Cell. Mol. Life Sci.* **72**, 2545–2559 (2015).
5. Y. Yokota *et al.*, *eLife* **4**, e08817 (2015).
6. C. P. Chang *et al.*, *Cell* **118**, 649–663 (2004).
7. F. Gunawan, A. Gentile, S. Gauvrit, D. Stainier, A. Bensimon-Brito, *Circ. Res.* (2020).
8. A. M. Ranger *et al.*, *Nature* **392**, 186–190 (1998).
9. B. Zhou *et al.*, *Development* **132**, 1137–1146 (2005).
10. R. A. Schulz, K. E. Yutzey, *Dev. Biol.* **266**, 1–16 (2004).
11. B. Wu, H. S. Baldwin, B. Zhou, *Trends Cardiovasc. Med.* **23**, 294–300 (2013).
12. F. Boselli, E. Steed, J. B. Freund, J. Vermot, *Development* **144**, 4322–4327 (2017).
13. J. Vermot *et al.*, *PLOS Biol.* **7**, e1000246 (2009).
14. S. E. Lyons *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5454–5459 (2002).
15. Y. Y. Foo *et al.*, *Biomech. Model. Mechanobiol.* **19**, 221–232 (2020).
16. F. Boselli, J. Vermot, *Methods* **94**, 129–134 (2016).
17. E. Heckel *et al.*, *Curr. Biol.* **25**, 1354–1361 (2015).
18. E. Steed *et al.*, *Nat. Commun.* **7**, 11646 (2016).
19. J. Aramburu *et al.*, *Science* **285**, 2129–2133 (1999).
20. T. Banjo *et al.*, *Nat. Commun.* **4**, 1978 (2013).
21. A. L. Duchemin, H. Vignes, J. Vermot, *eLife* **8**, e44706 (2019).
22. J. Schottenfeld, J. Sullivan-Brown, R. D. Burdine, *Development* **134**, 1605–1615 (2007).
23. K. Sekimizu *et al.*, *Development* **131**, 2521–2532 (2004).
24. P. Campinho, A. Vilfan, J. Vermot, *Front. Physiol.* **11**, 552 (2020).
25. K. Yamamoto *et al.*, *J. Cell Sci.* **124**, 3477–3483 (2011).
26. K. Yamamoto, R. Korenaga, A. Kamiya, J. Ando, *Circ. Res.* **87**, 385–391 (2000).
27. A. L. Armesilla *et al.*, *Mol. Cell. Biol.* **19**, 2032–2043 (1999).
28. A. K. Patra *et al.*, *Nat. Immunol.* **14**, 127–135 (2013).
29. C. P. Heisenberg, Y. Bellaïche, *Cell* **153**, 948–962 (2013).
30. T. Mammoto, D. E. Ingber, *Development* **137**, 1407–1420 (2010).
31. T. Merle, E. Farge, *Curr. Opin. Cell Biol.* **55**, 111–118 (2018).
32. S. Donat *et al.*, *eLife* **7**, e28939 (2018).
33. F. Gunawan *et al.*, *J. Cell Biol.* **218**, 1039–1054 (2019).
34. J. Pestel *et al.*, *Development* **143**, 2217–2227 (2016).
35. J. Li *et al.*, *Nature* **515**, 279–282 (2014).
36. K. Nonomura *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 12817–12822 (2018).
37. S. S. Ranade *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 10347–10352 (2014).


## ACKNOWLEDGMENTS

We thank the members of J.V.'s and N.Mo.'s laboratories; the imaging and the fish facility staff of IGBMC; and M. Sone, A. Chiba, K. Hiratomi, and T. Babazono in N.Mo.'s laboratory for their assistance. **Funding:** This work was supported by the European Research Council (GA no. 682938), the Agence Nationale de la Recherche (ANR-15-CE13-0015-01, ANR-10-IDEX-0002–02, ANR-12-ISV2-0001-01, and ANR-10-LABX-0030-INRT to J.V.), the University of Strasbourg (USIAS-2017-097), the Takeda Medical Research Foundation, the Uehara Memorial Foundation, the Cell Science Research Foundation (H.F.), and JSPS KAKENHI (21K 06764 to H.F.). R.W.-Y. C. was supported by a Fondation Lefoulon-Delalande fellowship. J.X. acknowledges a fellowship from the Chinese Scholarship Council (201708070046) and N.Mi grants from La Ligue Contre le Cancer (EL2021.LNCC/ NiM) and the European Research Council (ERC CoG "Forcaster" no. 647073). **Author contributions:** H.F. and J.V. conceived the project. H.F. designed the work and performed experiments. R.W.C. performed the analysis of $Ca^{2+}$ signaling normalization. H.F., J.X. and N.Mi. performed the magnetic tweezer experiments. Y.Y.F. and C.H.Y performed the analysis of 3D simulation. N.Mi. and N.Mo. provided reagents and suggestions. H.F. and J.V. wrote the manuscript. All authors interpreted the data, contributed to data analysis, and discussed the manuscript. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. Transgenic lines are available upon request.

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abc6229
Materials and Methods
Figs. S1 to S11
Captions for Movies S1 to S8
References (*38*–*56*)
Movies S1 to S8
MDAR Reproducibility Checklist

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

4 May 2020; accepted 31 August 2021
10.1126/science.abc6229

MUSCLE REPAIR

# Muscle repair after physiological damage relies on nuclear migration for cellular reconstruction

William Roman[1,2], Helena Pinheiro[2], Mafalda R. Pimentel[2], Jessica Segalés[1], Luis M. Oliveira[2], Esther García-Domínguez[3], Mari Carmen Gómez-Cabrera[3], Antonio L. Serrano[1], Edgar R. Gomes[2], Pura Muñoz-Cánoves[1,4,5]

Regeneration of skeletal muscle is a highly synchronized process that requires muscle stem cells (satellite cells). We found that localized injuries, as experienced through exercise, activate a myofiber self-repair mechanism that is independent of satellite cells in mice and humans. Mouse muscle injury triggers a signaling cascade involving calcium, Cdc42, and phosphokinase C that attracts myonuclei to the damaged site via microtubules and dynein. These nuclear movements accelerate sarcomere repair and locally deliver messenger RNA (mRNA) for cellular reconstruction. Myofiber self-repair is a cell-autonomous protective mechanism and represents an alternative model for understanding the restoration of muscle architecture in health and disease.

The contractile function of skeletal muscle is reliant on multinucleated muscle cells (myofibers), which comprise sarcomeres linked by structural proteins. Despite inherent muscle strength and flexibility, mechanical strain and stress can result in muscle damage, which is repaired by muscle resident stem cells (satellite cells) (*1*). Although stem cell–dependent muscle repair has been widely researched, it is unknown whether myofibers are able to self-repair without satellite cell activation.

Physical exercise induces mild musculoskeletal injuries that need rapid and efficient repair to preserve muscle homeostasis. To monitor myofiber response to localized damage, we subjected adult mice to eccentric exercise (muscle lengthening during contraction) and stained the muscle for filamin C and Hsp27, two markers of sarcomere damage, used here as readouts of injury location (*2*). Myofibers isolated at 5 and 24 hours after exercise displayed filamin C–enriched regions (termed scars hereafter), whereas scarring was absent in unexercised myofibers or reduced at 48 hours after exercise, the latter suggesting that sarcomere damage is rapidly resolved (Fig. 1, A and B). Filamin C and Hsp27 scars appeared within 30 min after exercise (fig. S1, A to C), were diminished after training (fig. S1, D to F), and were promoted in eccentric (lengthening) versus concentric (shortening) contractions (fig. S1, G to I). Filamin C scars also occurred in human muscle fibers after a single eccentric exercise session (fig. S1K), indicating that this physiological mechanism is conserved across species.

Notably, myofibers isolated 24 hours after exercise, relative to 5 hours after exercise, displayed myonuclei aggregation at the injury site (Fig. 1A). To quantify this observation, we computationally projected the three-dimensional position of filamin C scars and nuclei onto a two-dimensional plane (fig. S2A), revealing an increased number of myonuclei within a 10-μm distance from scars in myofibers isolated at 24 hours (Fig. 1C). Minimum internuclear distances near injury sites 24 hours after exercise were shorter than distances 5 hours after exercise and in unexercised myofibers (fig. S1, A to C).

To determine whether satellite cells function in the repair of exercise-induced muscle damage, we labeled proliferating cells (including satellite cells) with ethynyl deoxyuridine (EdU) after exercise. No EdU-positive nuclei or embryonic myosin heavy chain–expressing myofibers were associated with filamin C scars in exercised muscles (Fig. 1D and fig. S3A). Moreover, we did not observe Pax7-positive cells or myonuclei expressing high levels of the myogenic regulatory factor MyoD in proximity to filamin C scars after exercise (fig. S3, B to E). Genetic depletion of satellite cells (*3*) did not alter the filamin C response after exercise (fig. S3, F and G), and no fusion of satellite cells was detected 24 hours after exercise (fig. S3H). These results demonstrate that exercise-induced muscle damage rapidly activates a myofiber self-repair response that is independent of satellite cells within 96 hours after exercise.

To determine how myonuclei move to injury sites, we laser-damaged in vitro myotubes (muscle cells differentiated from myoblasts) (*4*) in a region of interest ~50 μm from the nearest myonucleus to mimic eccentric exercise–induced muscle damage. This disrupted sarcomeres and created a small lump on the cell surface (Fig. 1E and fig. S4A). Live imaging revealed that ~73% of the myonuclei, anchored at the myotube periphery, migrated to the nearby injury within 5 hours (Fig. 1, E and F). These nuclear movements were accompanied by structural repair of myofibrils, with the reorganization of sarcomeres [yellow fluorescent protein (YFP)–α-actinin realignment]. Injuries performed at a 2:3 interval between adjacent myonuclei favored the migration of the closest myonucleus, suggesting a distance-dependent mechanism (fig. S4B). Nuclear migrations to injury sites also occurred in laser-damaged ex vivo myofibers from adult mice (fig. S4C and movie S1) (albeit slower than in myotubes). Correlative live imaging showed that migrating nuclei stained negative for Pax7, confirming that these are indeed myonuclei and not satellite cells.

Next, we investigated which signaling pathway attracts myonuclei to injury sites. Because plasma membrane repair is calcium-dependent (*5*, *6*), we tested the influence of calcium signaling on nuclear movement. By expressing the calcium reporter GCaMP6f in myotubes (*7*), we observed an accumulation of cytosolic calcium at injury sites. Chelating calcium with BAPTA-AM resulted in shorter-lived calcium foci (fig. S5, A and B) and impaired nuclear migrations (Fig. 2, A and B). Most of the accumulated cytosolic calcium originates from the extracellular space (fig. S5C), and clearance of excess cytosolic calcium at injuries by mitochondria is required for proper plasma membrane repair (*8*). After confirming that mitochondria locate to injury sites (fig. S5, D to F), we inhibited mitochondrial $Ca^{2+}$ uptake using the Ru360 calcium blocker. This delayed calcium foci clearance (fig. S5, A and B), resulting in fewer nuclear migration events (fig. S5G).

Because calcium activates Rho guanosine triphosphatases (GTPases; Rho, Rac, and Cdc42) in wound closure (*9*), we assessed nuclear movements in injured myotubes exposed to Rho GTPase inhibitors. Only Cdc42 inhibition (using the ML141 inhibitor) significantly reduced the number of myonuclei migrating to the injury site (Fig. 2C). To confirm this, we successfully attracted myonuclei in undamaged myotubes by locally activating Cdc42 using optogenetics (Fig. 2D) but not by activating RhoA or an empty Cry2 construct (fig. S6, A and B) (*10*). Finally, inhibition of phosphokinase C (PKC), a calcium-dependent kinase in the Cdc42 pathway (*11*), also reduced the number of migrating myonuclei (Fig. 2C).

We next investigated the machinery that displaces myonuclei from their anchored position to reach the injury site. We exposed laser-damaged myotubes to cytoskeleton-inhibiting drugs 3 hours after damage and recorded nuclear speed before and after treatment (fig. S7A). The microtubule-inhibiting drugs nocodazole, colchicine, and taxol markedly reduced the speed of nuclear migrations toward the injury

[1]Department of Experimental & Health Sciences, University Pompeu Fabra, CIBERNED, 08003 Barcelona, Spain. [2]Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, 1649-028 Lisboa, Portugal. [3]FreshAge Research Group, Department of Physiology, Faculty of Medicine, University of Valencia, CIBERFES, Fundación Investigación Hospital Clínico Universitario/INCLIVA, Valencia, Spain. [4]Centro Nacional de Investigaciones Cardiovasculares, 28019 Madrid, Spain. [5]ICREA, 08010 Barcelona, Spain.
*Corresponding author. Email: william.roman@upf.edu

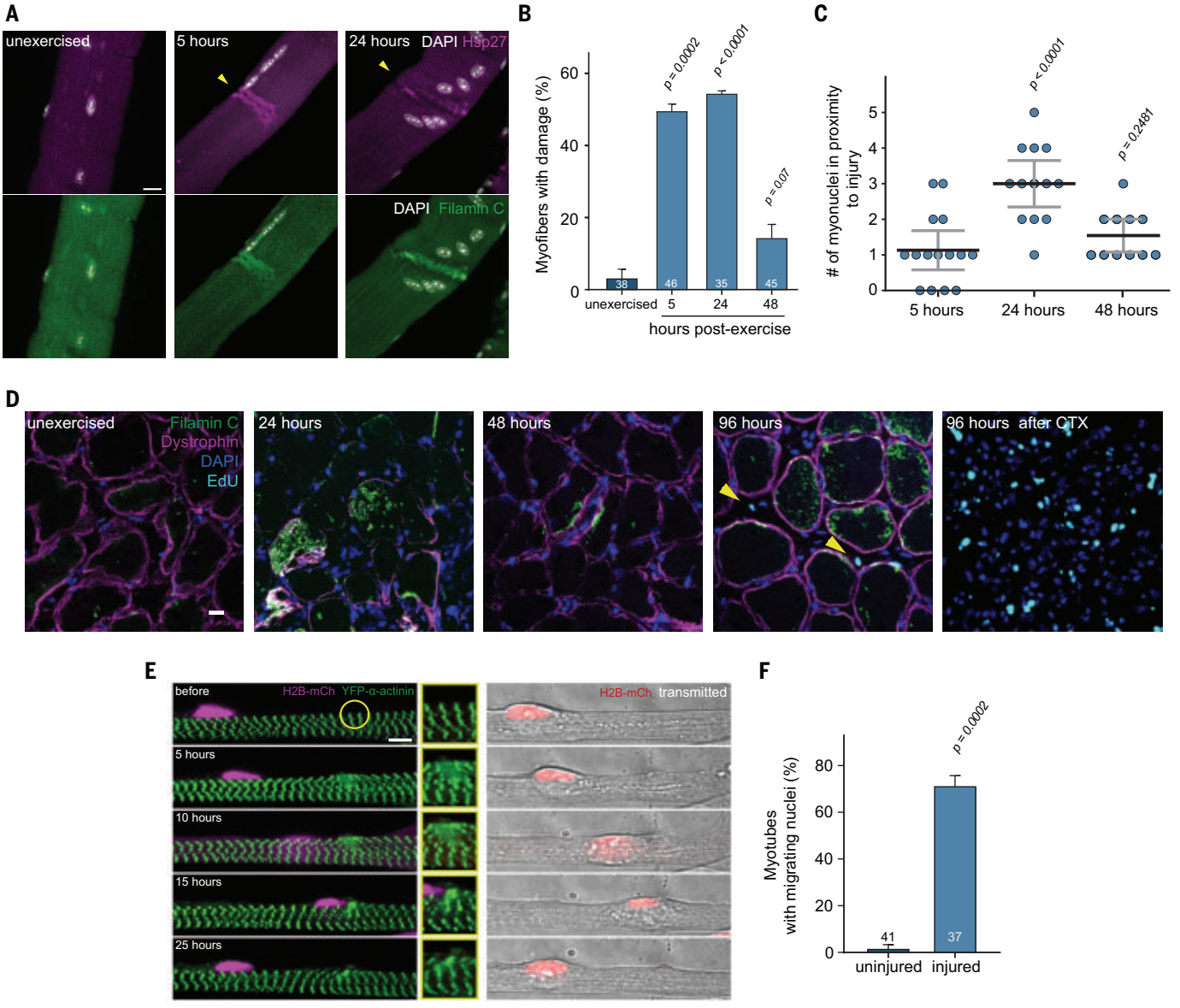



**Fig. 1. Myonuclei migrate to local injuries.** (**A**) Immunofluorescence of myofibers isolated from the soleus of unexercised mice and 5 and 24 hours after exercise. (**B**) Number of isolated myofibers with sarcomeric damage after exercise. (**C**) Number of myonuclei within 10 μm of sarcomeric damage after exercise. (**D**) Immunofluorescence of soleus cryosections stained for EdU in cardiotoxin-treated and exercised mice. (**E**) Time lapse of nuclear movement after laser-induced damage in a myotube expressing YFP–α-actinin (z-line) and mCherry-H2B (nucleus). (**F**) Nuclear movement quantification in undamaged or damaged myotubes. Yellow circles represent damage area; scale bar in (A), (D), and (E), 10 μm. *P* values are from unpaired two-tailed *t* tests [(B), (C), (F)]. Data are means ± SEM.

(Fig. 2, D and E). Ex vivo myofibers also had fewer myonuclei at injuries 24 hours after laser damage and nocodazole treatment (fig. S7, B and C). These experiments also revealed a microtubule network disruption at the injury site of untreated cells (fig. S7, C and D). To further explore these alterations, we tracked microtubule polymerization (EGFP-EB3 comets) before and after injury. Whereas microtubule growth speed remained constant after injury, the number of growing microtubules toward the nucleus was significantly decreased, whereas those moving toward the injury displayed no particular trend (fig. S7, E to H, and movies S2 and S3). This imbalance in the microtubule organization could favor nuclear migrations to the injury.

We next tested whether dynein, which is important for Cdc42-dependent nuclear centration early in myogenesis (*12*), participates in myonuclear movement toward the injury. Three hours after myotube injury, we inhibited dynein with ciliobrevin and observed a decrease in myonuclear speed, similar to the effect of microtubule inhibitors (Fig. 2E). Moreover, dynein accumulated at the injury site 3 hours after damage (fig. S7I). This suggests that microtubules and dynein are driving nuclear movement to the injury site, although off-target drug effects could be in play.

Interestingly, myonuclear speed increased in the presence of latrunculin A (Fig. 2E, S8A). We hypothesized that this was due to disruption of costameres, which link the plasma membrane to sarcomeres (*13*) and thus act as barriers restricting nuclear movements. To test this, we interfered with costamere turnover by inhibiting focal adhesion kinase (FAK) with PF-573228 (fig. S8B) (*14*, *15*). FAK inhibition reduced nuclear migrations toward the injury site without affecting microtubule reorganization (fig. S7, E and F), and nuclear migrations

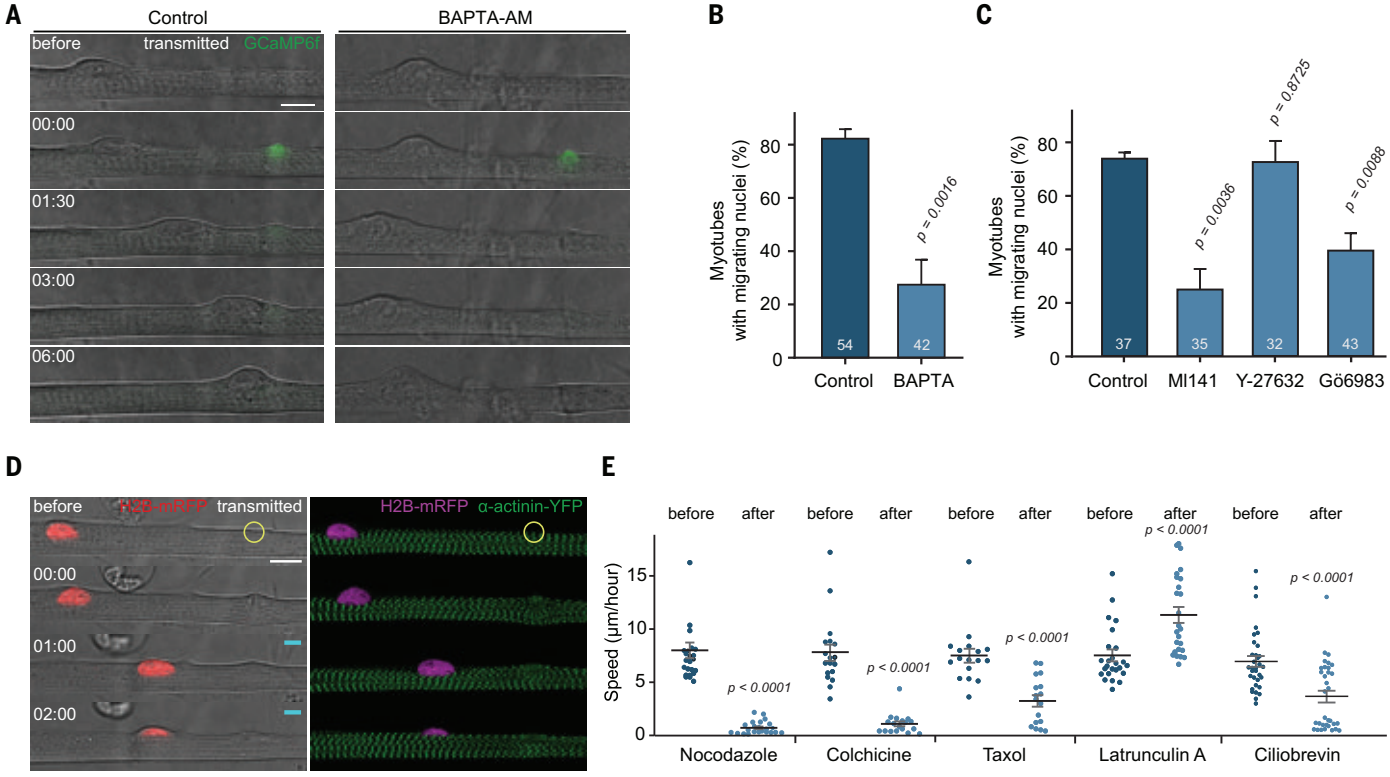



**Fig. 2. Calcium, PKC, and Cdc42 attract myonuclei to the injury; dynein and microtubules drive nuclear movement.** (**A**) Time lapse of nuclear movement after myotube damage from calcium chelation (BAPTA-AM). (**B** and **C**) Quantification of myotubes with myonuclei migrating to the injury. (**D**) Time lapse of nuclear movement (hours:min) in a damaged myotube expressing YFP–α-actinin and mCherry-H2B, exposed to nocodazole (cyan line) 1 hour after nuclear movement. Yellow circles represent pre-damage area; scale bar in (A) and (D), 10 μm. (**E**) Nuclear speed in damaged myotubes before and after treatment with cytoskeletal inhibitors. *P* values are from unpaired two-tailed *t* tests [(B) and (C)] and a paired two-tailed *t* test (E). Data are means ± SEM.

were restored when both FAK and actin were inhibited (fig. S8, C and D).

To assess whether nuclear migrations contribute to myofiber repair, we monitored sarcomere reconstruction in the presence of the PKC inhibitor Gö6983 and the FAK inhibitor PF-573228. We also inhibited transcription with triptolide to test whether repair depends on de novo mRNA synthesis. Untreated and treated myotubes were injured and stained at successive time intervals for sarcomerogenesis markers (*13*). Correlative microscopy revealed that within 24 hours after injury, sarcomere repair was almost complete in untreated myotubes, with titin reorganization at the injury site (Fig. 3, A and B, and fig. S9, A and B). In contrast, sarcomere repair was delayed in FAK- and PKC-inhibited myofibers and was blocked in transcription-inactivated myotubes. To confirm that nuclear proximity to injury affected myofiber self-repair, we performed injuries next to myonuclei in FAK-inhibited myotubes ("FAKi close") and observed a faster repair than in untreated fibers with distant injuries (Fig. 3, A and B). Two consecutive injuries resulted in a faster repair of the injury closest to myonuclei, whereas injuries on each side of myonuclei resulted in an overall delay of repair (fig. S9C). Actin-inhibited myotubes did not repair despite an accumulation of myonuclei at injury sites (fig. S9C).

To further confirm the involvement of nuclear movement in sarcomere reconstruction, we monitored myofibril (YFP–α-actinin) alignment in the same myotube before, immediately after, and 24 hours after injury using sinusoidal linear fitting. Sarcomeres almost fully realigned within 24 hours in untreated myotubes, whereas they remained misaligned in the FAK-, PKC-, and transcription-inhibited conditions (Fig. 3, C and D). Sarcomere realignment in FAK-inhibited myotubes improved when the injury was performed closer to myonuclei (FAKi close).

Our data show that transcription is required for sarcomere repair, which suggests that nuclear migration replenishes proteins at the injury site. We first monitored translation after damage and observed an accumulation of ribosomes at injuries of untreated and FAK-inhibited myotubes 6 hours after damage (fig. S10A). Using a puromycilation assay (*16*), we detected active translation hotspots enriched at myotube injury sites (Fig. 4, A and B). This enrichment also occurred in FAK- and PKC-inhibited myotubes, as well as in laser-damaged ex vivo myofibers 24 hours after injury (fig. S10, B and C).

Given the similar increase in translation hotspots at injury sites irrespective of nuclear positioning, we investigated the amount and distribution of mRNA by total and single-molecule fluorescence in situ hybridization (smFISH). In vitro myotubes were fixed 6 hours after injury and stained for total and specific mRNAs: *Actn2* (α-actinin), *Flnc* (Filamin C), *Hspb1* (Hsp27), and *Myh10* (NMIIb). These mRNAs encode proteins that accumulate at injury sites (*Flnc* and *Hspb1*; Fig. 1A) or are involved in repairing sarcomeres (*Actn2* and *Myh10*; Fig. 3). Total and *Actn2* mRNA concentrated between migrating myonuclei and injury sites (fig. S11, A to F). Reflecting the protein data (Figs. 1A and 3A), *Flnc*, *Hspb1*, and *Myh10* mRNAs increased sharply in injured myotubes (Fig. 4, C to E, and fig. S11, G to M). *Myh10* was also up-regulated in FAK- and PKC-inhibited myotubes, but mRNAs were more evenly dispersed between injury sites and anchored myonuclei. This suggests that although nuclear migration is not required for up-regulating the expression of specific mRNAs, it is important for locally delivering these mRNAs at injury sites. We observed a similar increase in total as well as *Actn2*, *Myh10*, *Flnc*, and *Hspb1* mRNAs at the injury site of ex vivo myofibers 24 hours after damage (Fig. 4E and fig. S11, B to M).

We next assessed whether exercise-induced muscle damage triggered translation and

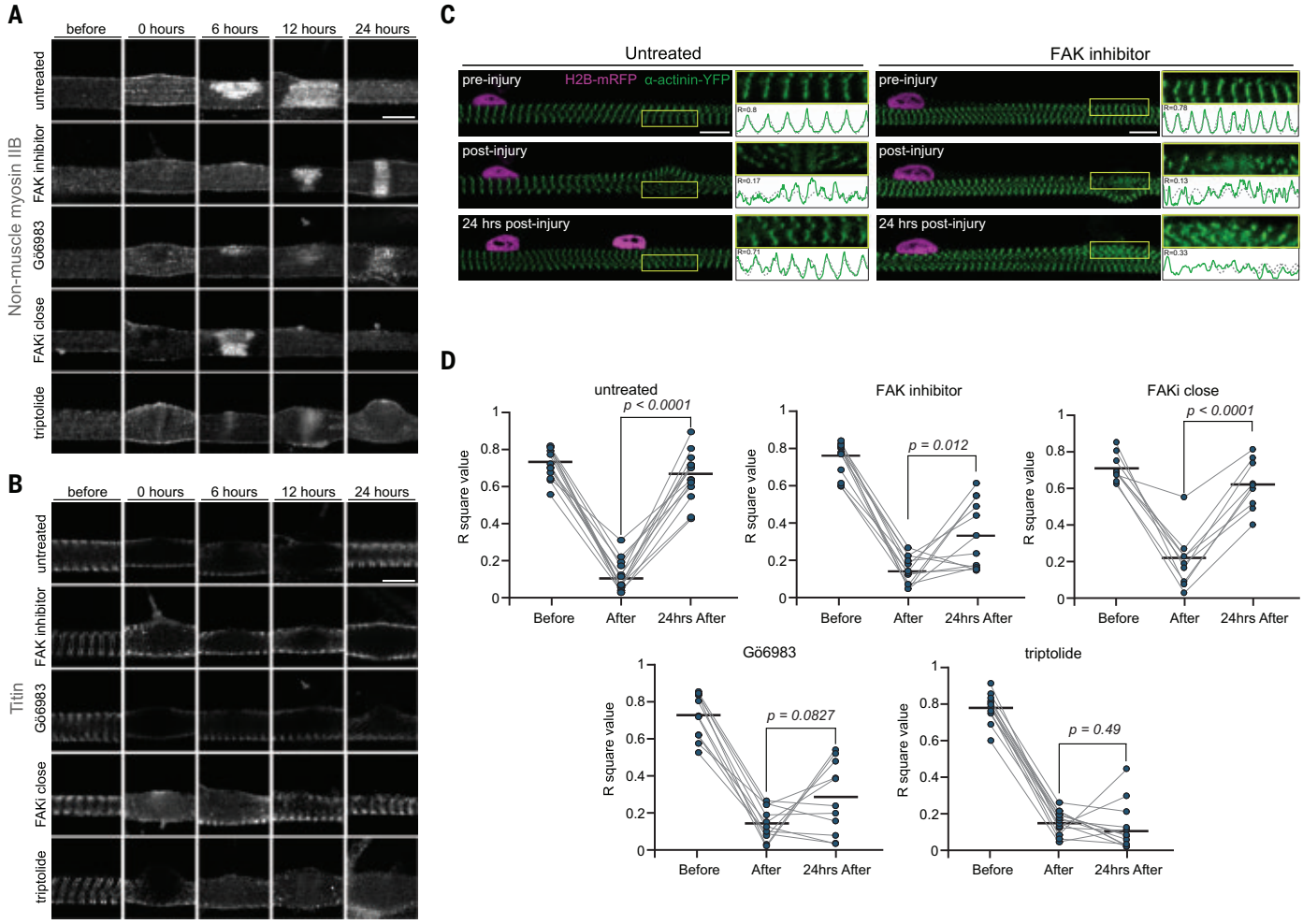



**Fig. 3. Myonuclear movement accelerates sarcomere repair.** (**A** and **B**) Time course of injuries stained for non-muscle myosin IIb and titin in FAK-inhibited (PF-573228), PKC-inhibited (Gö6983), and transcription (triptolide)–inhibited myotubes. (**C**) Time-lapse imaging of untreated and FAK-inhibited myotubes before, immediately after, and 24 hours after damage. Intensity line scans of sarcomeric alignment (green line) are used to derive an *R* value from the closest fitted sinusoidal (dashed gray line). (**D**) *R* values before, after, and 24 hours after damage in untreated and treated myotubes. Data are means; scale bar in (A), (B), and (C), 10 μm. *P* values are from a paired two-tailed *t* test (D).

transcription changes similar to those seen in laser-damaged myofibers. In myofibers isolated 5 and 24 hours after exercise, translation hotspots (Fig. 4, F and G) and mRNA (Fig. 4, H and I, and fig. S12, A to E) increased at injury sites relative to adjacent areas or unexercised myofibers. Consistent with the protein expression (Fig. 1A), *Hspb1* mRNA peaked at 5 hours and *Flnc* at 24 hours after exercise. This suggests that myonuclei adopt a temporally orchestrated transcriptional signature after injury. In agreement, expression of *Myh10* increased drastically 24 hours after exercise, coinciding with detection of this sarcomere repair protein and realignment of α-actinin at injuries (Fig. 4J).

Skeletal muscle homeostasis and regeneration are widely assumed to rely exclusively on satellite cells. Opposing this stem cell–centric view, we have shown that muscle integrity is also maintained through an alternative myofiber-autonomous repair mechanism (fig. S13). Using in vitro, ex vivo, and in vivo models of local muscle damage, we identified nuclear migration as a central mechanism for the local delivery of mRNAs required for rapid protein production and repair of the damaged sarcomeres.

## REFERENCES AND NOTES


1. S. B. P. Chargé, M. A. Rudnicki, *Physiol. Rev.* **84**, 209–238 (2004).
2. M. P. Collier *et al.*, *Sci. Adv.* **5**, eaav8421 (2019).
3. R. Sambasivan *et al.*, *Development* **138**, 3647–3656 (2011).
4. S. Falcone *et al.*, *EMBO Mol. Med.* **6**, 1455–1475 (2014).
5. A. R. Demonbreun *et al.*, *J. Clin. Invest.* **129**, 4657–4670 (2019).
6. A. Reddy, E. V. Caler, N. W. Andrews, *Cell* **106**, 157–169 (2001).
7. T.-W. Chen *et al.*, *Nature* **499**, 295–300 (2013).
8. V. Debattisti *et al.*, *Cell Rep.* **29**, 1274–1286.e6 (2019).
9. H. A. Benink, W. M. Bement, *J. Cell Biol.* **168**, 429–439 (2005).
10. L. Valon *et al.*, *Biophys. J.* **109**, 1785–1797 (2015).
11. S. J. Slater, J. L. Seiz, B. A. Stagliano, C. D. Stubbs, *Biochemistry* **40**, 4437–4445 (2001).
12. B. Cadot *et al.*, *EMBO Rep.* **13**, 741–749 (2012).
13. J. W. Sanger *et al.*, *Handb. Exp. Pharmacol.* **235**, 39–75 (2017).
14. N. L. Quach, T. A. Rando, *Dev. Biol.* **293**, 38–52 (2006).
15. J. K. Slack-Davis *et al.*, *J. Biol. Chem.* **282**, 14845–14852 (2007).
16. A. Bastide, J. W. Yewdell, A. David, *Biol. Protoc.* **8**, e2670 (2018).
17. J. Segalés *et al.*, *Nat. Commun.* **11**, 189 (2020).


## ACKNOWLEDGMENTS


We thank I. B. Martins for her contribution in the experimental work; V. Lukesova, E. Andrés, A. Navarro, A. Pollán, J. Ballestero, I. Ramirez-Pardo, and S. Campanario for their input and discussions; the bio-imaging facilities of the Barcelona Biomedical Research Park (PRBB) and the Instituto de Medicina Molecular (iMM); and V. Sebastiá (Clínica Ypsilon, Valencia, Spain) for advice on human exercise studies. **Funding:** Supported by the Association Francaise de Myopathie (#21450) (W.R.), the European Molecular Biology Organization (ALTF 1207-2017) (W.R.), the Human Frontier Science Program (T000133/2018-L) (W.R.), MINECO-Spain (RTI2018-096068) (P.M.-C.), ERC-2016-AdG-741966 (P.M.-C.), LaCaixa-HEALTH-HR17-00040 (P.M.-C.), UPGRADE-H2020-825825 (P.M.-C.), MWRF (P.M.-C.), Fundació La Marató TV3 (202021 to P.M.C. and 202033 to A.L.S.), María-de-Maeztu-Program for Units of Excellence to UPF (grant MDM-2014-0370), and the Severo-Ochoa-Program for Centers of Excellence to CNIC (grant SEV-2015-0505). Work in the Gomes laboratory was supported by the European Molecular Biology Organization (#2758) (E.R.G.), the H2020 and FP7 Ideas European Research Council (#810207 and 617676) (E.R.G.), the Fundação para a Ciência e a Tecnologia (PTDC/BIA-CEL/31075/2017) (E.R.G.), and the MEC | Fundação para a Ciência e a Tecnologia (PD/BD/128287/2017) (H.P.). Work in FreshAge laboratory was supported by PID2019-110906RB-I00/AEI 10.13039/501100011033 from MINECO, CB16/10/00435 (CIBERFES) from ISCIII, PROMETEO/2019/097 from GVA and ERDF funds, and DIABFRAIL-LATAM from EU-H2020 (825546).
**Author contributions:** W.R. carried out experiments and analyzed

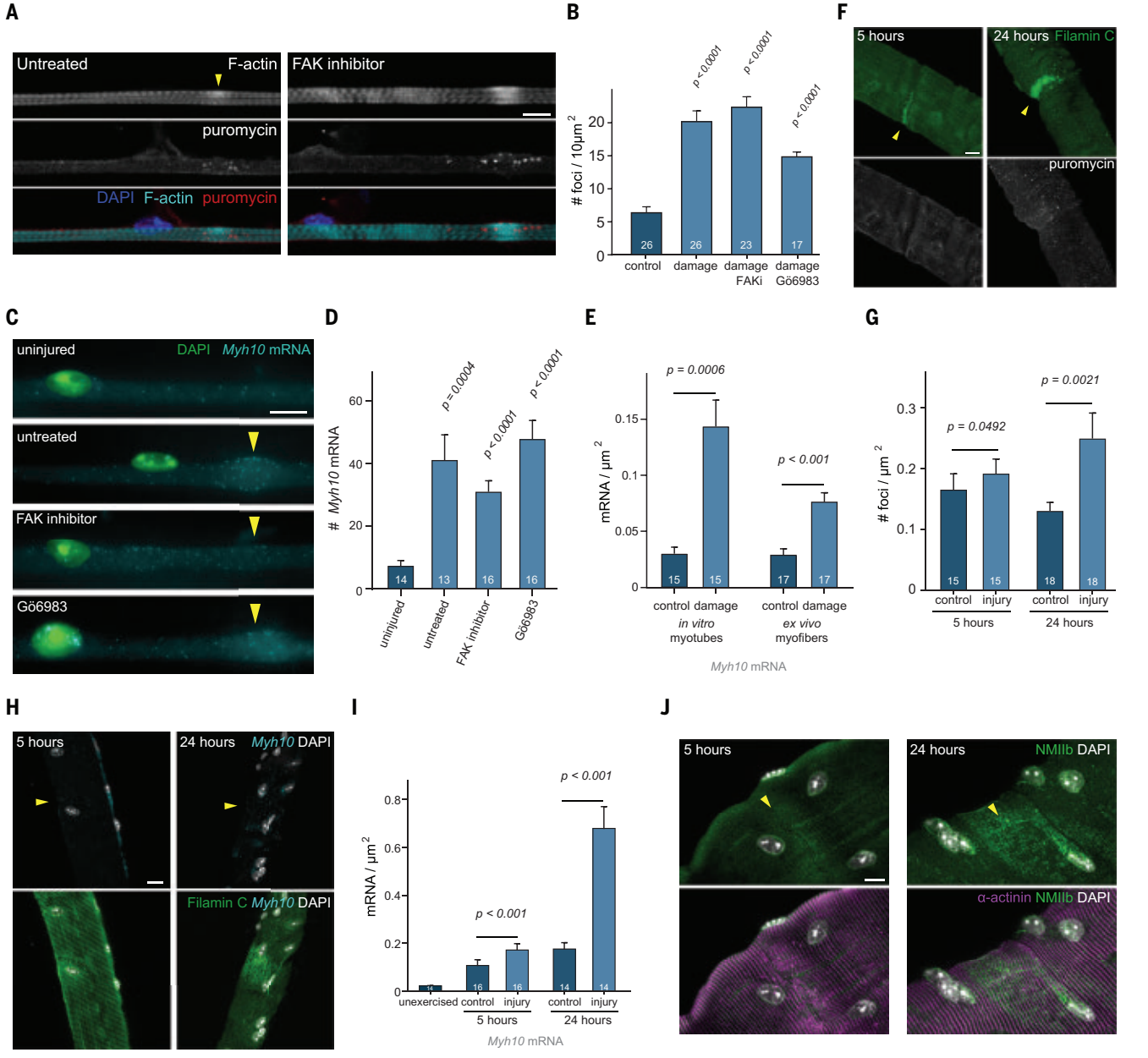



**Fig. 4. Migrating myonuclei locally deliver de novo mRNA for sarcomere reconstruction.** (**A**) Puromycilation assay of untreated and FAK-inhibited myotubes stained for puromycin (translation hotspots) and F-actin (sarcomeres) 6 hours after damage. (**B**) Translational hotspot density in untreated, damaged, FAK-inhibited, and PKC-inhibited myotubes 6 hours after damage. (**C**) *Myh10* probes in myotubes 6 hours after laser damage. (**D** and **E**) Quantification of *Myh10* mRNA in laser-damaged myotubes and myofibers. (**F**) Puromycilation assay in myofibers isolated 5 and 24 hours after exercise. (**G**) Quantification of translation hotspots at the injury in myofibers isolated 5 and 24 hours after exercise. (**H**) *Myh10* smFISH and filamin C immunofluorescence in exercised myofibers. (**I**) Quantification of *Myh10* at the injury and at an adjacent undamaged area in exercised myofibers. (**J**) Immunofluorescence of exercised myofibers stained for α-actinin and NMIIb. Yellow arrowhead designates injury location; scale bar in (A), (C), (F), (H), and (J), 10 μm. *P* values are from an unpaired two-tailed *t* test (B) and paired two-tailed *t* tests [(D), (E), (G), and (I)]. Data are means ± SEM.

data; H.P. and M.R.P. performed the FISH stainings; L.M.O. performed the computational analysis of nuclear positioning and the EB3 comet analysis; A.L.S., M.C.G.-C., E.G.-D., J.S., and W.R. designed, performed, and analyzed in vivo experiments in mice; M.C.G.-C., E.G.-D., and A.L.S. performed and analyzed human exercise experiments; W.R. and P.M.-C. conceived and designed experiments; W.R., P.M.-C., and E.R.G. wrote the manuscript with assistance from other authors; and all authors participated in the critical review and revision of the manuscript. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. Computer code can be found at: https://github.com/william-roman/Myofiber-self-repair.git. DOI:10.5281/zenodo.4974256.

**SUPPLEMENTARY MATERIALS**

science.org/doi/10.1126/science.abe5620
Materials and Methods
Figs. S1 to S13
Reference (*17*)
Movies S1 to S3

7 September 2020; resubmitted 6 May 2021
Accepted 20 August 2021
10.1126/science.abe5620

**Photochemistry LED Illuminator**
Uniqsis introduces the Photochemistry LED Illuminator—a new tool to accelerate reaction optimization and batch scale-up of photochemical reactions. The new illumination system incorporates LED arrays operating at six different wavelengths from 365 nm–525 nm. The LEDs are mounted on cassettes for easy interchangeability, in industry-standard 24- or 48-well configurations. The illuminator also has the option of using the six-multiwavelength screening plate for rapid reaction optimization, and is designed and built to meet the British Standards Institute BS EN 61010 safety requirements for measurement, control, and laboratory use of electrical equipment. The system employs interlocks to ensure the LEDs are not powered when uncovered, thereby meeting strict eye-safety requirements.
**Uniqsis**
For info: +44-(0)-845-864-7747
www.uniqsis.com

**Aptamers for Targeting SARS-CoV-2 Proteins**
AMS Biotechnology (AMSBIO) has further expanded its product range to help in the fight against COVID-19 by introducing a new range of aptamers targeting SARS-CoV-2 proteins, to aid development of new kits and assays. Aptamers have generated significant interest as promising tools that compete with monoclonal antibodies for detection and inhibition of specific markers of human diseases, because they offer high specificity for their target, no to low immunogenicity, and easy in vitro selection. Using an aptamer-based technique can help researchers to better understand the genetic architecture of host proteins interacting with SARS-CoV-2 or mediating host response to COVID-19, and to identify new drugs or repurpose existing drugs targeting those proteins. AMSBIO aptamers targeting SARS-CoV-2 S1 protein, S2 protein, receptor-binding domain (RBD) protein, or nucleocapsid protein can be used in a wide range of detection methods and for development of COVID-19 diagnostic assays.
**AMS Biotechnology**
For info: +1-617-945-5033
www.amsbio.com/aptamers-for-infectious-diseases

**4-Position High-Pressure Benchtop Reactor**
The Quadracell 4-position high-pressure reactor from Asynt is a new synthetic tool for scientists seeking the convenience of routine parallel pressure chemistry in a compact footprint. The robust Quadracell reactor is simple to set up and operate, with heating and agitation provided by a standard magnetic hotplate stirrer and magnetic flea. Constructed in durable 316 stainless steel, the Quadracell has four cells—each with a reaction volume of 10 mL—that can routinely operate at temperatures up to 250°C. For labs using aggressive or caustic media, the Quadracell is also available with Hastelloy C276 construction for enhanced chemical resistance. Different versions of the Quadracell are available to safely operate up to your chosen maximum working pressure of 50, 100, or 200 Bar. Each reactor is supplied as standard with a pressure release valve, pressure gauge, gas inlet, and gas outlet via two stainless-steel needle valves.
**Asynt**
For info: +44-(0)-1638-781709
www.asynt.com/product/quadracell-4-position-high-pressure-reactor

**Biomarkers for miRNA Cancer Detection and Treatment**
Explore BioGenex's comprehensive range of Super Sensitive Nucleic Acid (SSNA) miRNA probes for microRNA (miRNA) in situ hybridization (ISH) staining. Our unique miRNA probes enable detection of a single-nucleotide mismatch, allowing highly sensitive and specific detection. The higher sensitivity and specificity of SSNA miRNA probes make them a useful diagnostic and miRNA research tool for the identification of cancer types. Our ready-to-use miRNA cancer probes address the challenges in the identification of breast cancer cells (Hsa-miR-299-3p, Hsa-miR-4787-3p); pancreatic cancer cells (Hsa-miR-301A-3p); colorectal cancer cells (Hsa-miR-556); tumor/node/metastasis (TNM)-stage II colon cancer cells (Hsa-miR-656-3p, Hsa-miR-5010-3p, Hsa-miR-5100); melanoma cells (Hsa-miR-1537, Has-miR-4787-3p); and colon adenomas (Hsa-miR-671-3p, Hsa-miR-5010-3p, Hsa-miR-5100); as well as non-small-cell lung cancer and pancreatic cancer cells (Hsa-miR-5100). BioGenex offers 240 miRNA cancer probes that are validated and optimized for human tissues and are available in a ready-to-use format. The probes are compatible with BioGenex fully automated staining systems, ISH, and miRNA detection systems for clean and intense staining.
**BioGenex**
For info: +1-800-421-4149
biogenex.com

**Evaporator**
Porvair Sciences' Ultravap Mistral XT 150 blowdown evaporator is now able to operate at temperatures up to 80°C, significantly increasing the rate of evaporation of both organic solvents and aqueous solutions. Designed to provide drydown of samples in microplates and glass tubes up to 150 mm in length, the XT150 offers a low-maintenance, sample-safe alternative to using water bath–based evaporators. Nearly all common chromatography solvents, including acetonitrile, dichloromethane, hexane, methanol, and water can be evaporated with ease by the XT150. The availability of straight or spiral blowdown evaporator needles allows users to choose between faster drydown (spiral) or optimized final drying in V-well microplates (straight). Installation of the blowdown evaporator requires only connection to a gas supply and mains electricity. Safety of solvent removal operation is ensured as this compact, European Conformity (CE)-marked unit fits into all fume cupboards.
**Porvair Sciences**
For info: +1-800-552-3696
www.microplates.com/blowdown-evaporator-ultravap-mistral

**Ultra-Low Temperature Freezer**
Eppendorf continues to follow the path of sustainability by redesigning the 400-L ultra-low temperature (ULT) freezer class. The well-known standard-sized ULT freezer is now completely equipped with future-proof green cooling liquids and green insulation foam as well as better performance. Low energy consumption of down to 6.8 KWh/day (−80°C; 230 V) is combined with powerful pull-down to −80°C within 200 min and fast recovery back to −80°C after door opening. The Cryo F440 is the perfect solution for freezers that are used in the lab daily, keeping valuable samples safe at constant conditions. Compared to its precursor, the F440 saves up to 23% of power combined with 30% more space. The F440 series can be connected to the VisioNize Lab Suite for monitoring via the VisioNize box. The eLABInventory sample management software can support the smart Eppendorf storage solution of high-value samples.
**Eppendorf**
For info: +1-800-645-3050
www.eppendorf.com/freezers

Electronically submit your new product description or product literature information! Go to www.science.org/about/new-products-section for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

ScienceCareers online @sciencecareers.org

BOSTON UNIVERSITY

## FACULTY POSITION IN COMPUTATIONAL BIOLOGY

The Faculty of Computing & Data Sciences (CDS) at Boston University invites applications for a tenure-track **Assistant Professor** position in the area of basic research focused on the design, implementation, and application of computational and data science methodologies and platforms to tackle important problems in biology, including research in algorithm development, computational modeling, machine learning, data mining, AI systems, and tools for generating and analyzing large biological datasets across all fields of biology and medicine, broadly interpreted. One of a number of open tenure-track faculty positions available in CDS, this position is for a candidate who will also be affiliated with the Boston University Program in Bioinformatics.

Boston University expects excellence in teaching and in research and is committed to building a culturally, racially, and ethnically diverse scholarly community, which is essential to its mission as expressed at www.bu.edu/info/about/diversity. Qualifications required of all applicants include a PhD in any of the disciplines that span computing and data science; a strong record of research; a demonstrated capacity for interdisciplinary collaboration; and a commitment to innovative teaching at the undergraduate and graduate levels.

Consideration and review of applications will start on **November 1, 2021**, and will continue on a rolling basis until **April 15, 2022**.

Interested applicants should review the full position posting and apply at **https://academicjobsonline.org/ajo/jobs/19582**.

*BU is an equal opportunity employer and all qualified applicants will receive consideration for employment without regard to race, color, religion, sex, sexual orientation, gender identity, national origin, disability status, protected veteran status, or any other characteristic protected by law. We are a VEVRAA Federal Contractor.*

---

RUTGERS
School of Arts and Sciences

The Department of Cell Biology and Neuroscience invites applications for a tenure-track position at the Assistant Professor level to develop an innovative research program focused on cellular metabolism, broadly defined, using physiological and human disease models. Research approaches may encompass, but are not limited to, systems biology, proteomics, and metabolomics.

As part of the Division of Life Sciences, the Department of CBN is home to an interactive, collegial faculty with broad interests, including immunology, cellular signaling, stem cell biology, and neuroscience. Multiple opportunities for interdisciplinary collaboration exist within the Department and nearby labs at The Rutgers Robert Wood Johnson Medical School, the Cancer Institute of New Jersey, The Child Health Institute, and Princeton University. Rutgers offers excellent facilities, faculty mentoring, and competitive start-up packages. The campus is located within one hour of New York and Philadelphia.

Outstanding applicants will hold a Ph.D., M.D., or equivalent degree and have a minimum of three years of postdoctoral training in a relevant field. A strong track-record of achievement is required. The successful candidate will primarily teach undergraduate courses in the areas of cell biology, immunology, or neurobiology and maintain a productive, extramurally funded research program.

Rutgers University has one of the most diverse student bodies in the nation and boasts several programs established to nurture equity and inclusion. These programs enhance the achievement of excellence by both faculty members and students. We encourage applications from women and members of communities who are underrepresented in the sciences and will evaluate the potential of the applicant to mentor and empower our students.

Interested individuals are encouraged to apply online through the department website (**cbn.rutgers.edu**) with a curriculum vitae, brief statement of research plans, statement discussing approach to diversity, outreach, service and/or mentoring, statement of teaching outlook and experiences, and the names, addresses, and contact information of three references.

Applications should be submitted as soon as possible but not later than **December 15, 2021**. Late applications will be considered only if the position remains available.

*Rutgers University is an equal opportunity/affirmative action employer.*

---

Penn
University of Pennsylvania

### University of Pennsylvania

Assistant Professor in Experimental Particle Physics

The Department of Physics and Astronomy at the University of Pennsylvania seeks applications from outstanding candidates for an appointment as an **Assistant Professor** in experimental particle physics. In exceptional cases, an appointment at a higher rank may be considered. The successful candidate will develop an innovative research program that will join our existing effort in collider physics on the ATLAS experiment at the CERN Large Hadron Collider. The candidate will be expected to attract external research funding, to advise graduate students, and to excel in teaching undergraduate students.

Applicants must apply online at **http://apply.interfolio.com/94423**. Applicants should have a Ph.D. at the time of appointment. Required application materials include the following: curriculum vitae with a list of publications, a research statement, a teaching statement, and at least three confidential letters of recommendation. Review of applications will begin no later than **30 November 2021** and will continue until the position is filled. It is anticipated that the position will start on **1 July 2022**. For inquiries about this position, please contact Ms. Michelle Last (**michlast@physics.upenn.edu**)

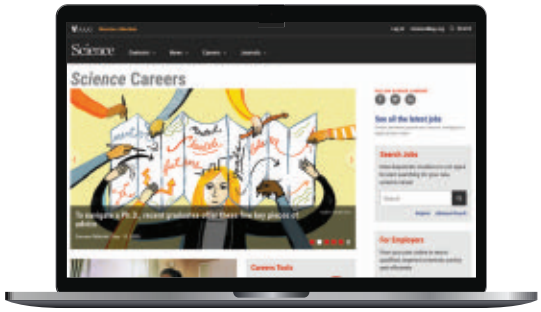
*The Department of Physics and Astronomy is strongly committed to Penn's long-standing Action Plan for Faculty Diversity and Excellence and to creating a more diverse faculty (for more information see: http://www.upenn.edu/almanac/volumes/v58/n02/diversityplan.html). The University of Pennsylvania is an equal opportunity employer. Minorities/Women/Individuals with disabilities/Protected Veterans are encouraged to apply.*

ScienceCareers online @sciencecareers.org

# *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.ORG

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

**U.S. POSTAL SERVICE**

Statement required by the Act of 12 August 1970, Section 3685, Title 39, United States Code, showing the ownership, management, and circulation of:
1–9. *Science*, Publication No. 0036-8075, is published weekly on Friday, except the last week in December, at 1200 New York Avenue, N.W., Washington, DC 20005. Date of filing: 29 September 2021. This is also the address of the publisher, the editor, and the managing editor, who are, respectively, Bill Moran, Holden Thorp, and Monica M. Bradford.
10. The owner is the American Association for the Advancement of Science, 1200 New York Avenue, N.W., Washington, DC 20005. Stockholders: None.
11. Known bondholders, mortgages, and other security holders owning or holding 1 percent or more of total amount of bonds, mortgages, or other securities: None.
12. The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes have not changed during the preceding 12 months.
13–15. The average number of copies of each issue during the preceding 12 months is (A) Total number of copies printed: 50,535; (B) Paid circulation: 46,380; (1) Paid/Requested outside-county mail subscriptions stated on form 3541: 40,977; (2) Paid/Requested in-county subscriptions stated on form 3541: 0; (3) Sales through dealers and carriers, street vendors, counter sales: 5,398. (4) Other classes mailed through USPS: 5; (C) Total paid circulation: 46,380; (D) Free distribution: samples, complimentary, and other free copies: 3,245; (1) Outside-county as stated on form 3541: 2,898; (2) In-county as stated on form 3541: 0; (3) Other classes mailed through the USPS: 2; (4) Free distribution outside of mail carrier or other means: 346; (E) Total free distribution: 3,245; (F) Total distribution: 49,625; (G) Copies not distributed: 910; (H) Total: 50,535; (I) Percent paid and/or Requested Circulation: 93.5%.
Actual number of copies of single issue (9/17/2021) published nearest to filing date are (A) Total number of copies printed: 47,804; (B) Paid circulation: 43,784; (1) Paid/Requested outside-county mail subscriptions stated on form 3541: 38,670; (2) Paid/Requested in-county subscriptions stated on form 3541: 0; (3) Sales through dealers and carriers, street vendors, counter sales: 5,109; (4) Other classes mailed through USPS: 5; (C) Total paid circulation: 43,784; (D) Free distribution: Samples, complimentary, and other free copies: 3,165; (1) Outside-county as stated on form 3541: 2,877; (2) In-county as stated on form 3541: 0; (3) Other classes mailed through the USPS: 2; (4) Free distribution outside of mail: Carrier or other means: 286; (E) Total free distribution: 3,165; (F) Total distribution: 46,949; (G) Copies not distributed: 855; (H) Total: 47,804; (I) Percent paid and/or Requested Circulation: 93.3%.
I certify that the statements made above are correct and complete. (signed) Bill Moran, Publisher.

**Search more jobs online**

Access hundreds of job postings on ScienceCareers.org.

**Expand your search today.**

## CHAIR, DEPARTMENT OF PHARMACEUTICAL SCIENCES

**Summary of Duties:** Washington State University College of Pharmacy and Pharmaceutical Sciences is seeking a Chair for the Department of Pharmaceutical Sciences. The Chair will serve as a catalyst to promote inter- and intra-departmental collaborations within the College and University. The Chair will support innovation and entrepreneurship within the research mission of the College and provide strategic and operational leadership to achieve the vision, mission, and goals of the Department, College, and University.

Reporting to the Dean and serving on the Executive Committee, the Chair will be responsible for leading a world-class, research-focused department and for maintaining an outstanding, externally funded research program that generates in excess of $10M annually. Faculty members within the department conduct basic and translational research in critical areas of pharmaceutical and toxicological sciences, addressing a variety of major health disorders including inflammatory, cardiovascular, and infectious diseases, and cancer using in vitro and in vivo models. Research approaches of our faculty include computational pharmacology and molecular modeling, interindividual variability in drug metabolism, disposition and toxicity, pediatric- and physiologically-based pharmacokinetic modeling, pharmacogenomics, and drug discovery and medicinal chemistry.

The Department supports the training of students in the Doctor of Pharmacy (PharmD) program, including teaching the concepts of the pharmaceutical sciences within the curriculum, and provides graduate student training in programs that lead to the PhD and PharmD/PhD degrees in Pharmaceutical Sciences and Molecular Medicine. The Department offers a popular summer undergraduate research program that is funded by the American Society of Pharmacology and Experimental Therapeutics. The Chair will be responsible for leading additional growth in the Department's signature research areas and for enhancing collaborative research throughout the WSU Health Sciences campus. The Chair will support academic programs for professional pharmacy students and graduate students, as well as teaching and training opportunities for postdoctoral fellows.

The position requires excellent interpersonal skills and the ability to unite diverse groups of scientists, students, and staff to advance the common goals of the Department, College, and University. The successful candidate will be a strong leader with a proven track record in an area of research within or related to the pharmaceutical sciences, a clear understanding of the complexities of leading an academic department, and the ability to lead growth and development of the department.

**The responsibilities and duties of the Chair include those described below:**

**Required Qualifications:** Applicants must possess a PhD, or equivalent degree, within the broad area of Pharmaceutical Sciences, or a related discipline, with scholarship commensurate with appointment at the rank of tenured Professor.

**Additional Information:** As part of the recruitment process, please be prepared to submit the following documents:

1) a current curriculum vitae; and
2) a cover letter addressing your professional goals

Applicants are required to include contact information for professional references within the application.

The screening of application materials will begin on November 15, 2021, and continue until a suitable candidate has been identified.

To apply and view a complete position description with all qualifications please go to:
**https://wsu.wd5.myworkdayjobs.com/en-US/WSU_Jobs/job/PHARMACEUTICAL--BIOMEDICAL-SCIENCES-BLDG-WSU-SPO/Chair--Department-of-Pharmaceutical-Sciences_R-2229**

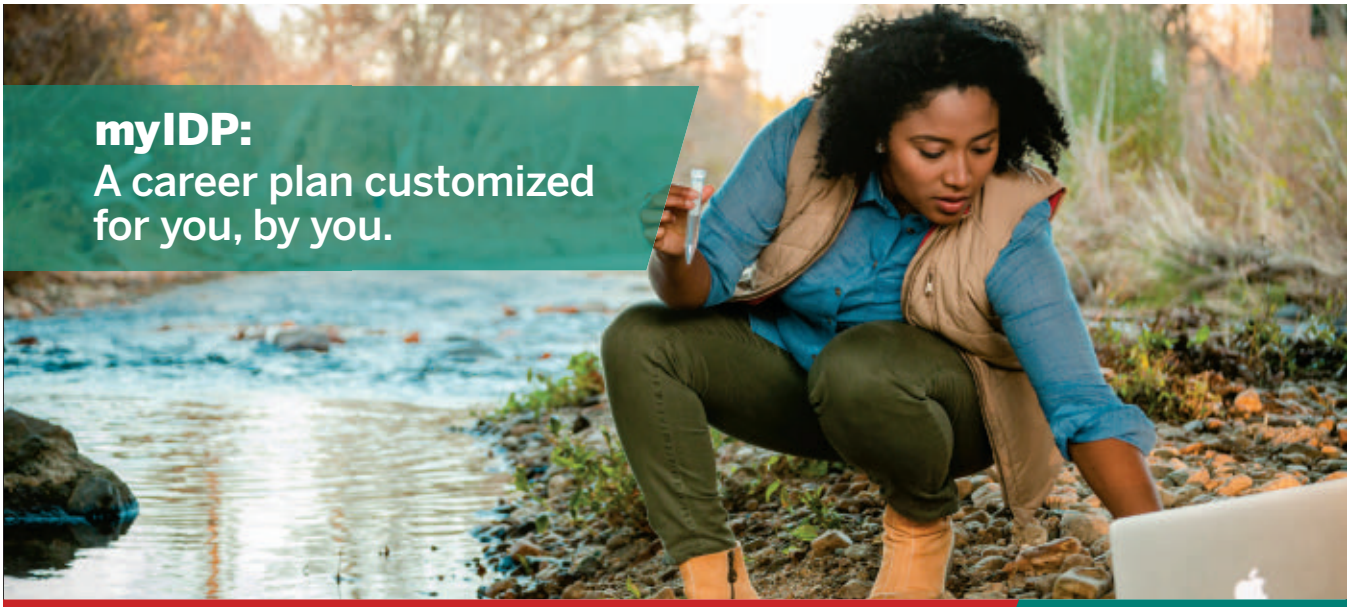
*WASHINGTON STATE UNIVERSITY IS AN EQUAL OPPORTUNITY/AFFIRMATIVE ACTION EDUCATOR AND EMPLOYER. Members of ethnic minorities, women, special disabled veterans, veterans of the Vietnam-era, recently separated veterans, and other protected veteran, persons of disability and/or persons age 40 and over are encouraged to apply.*

*WSU employs only U.S. citizens and lawfully authorized non-U.S. citizens. All new employees must show employment eligibility verification as required by the U.S. Citizenship and Immigration Services.*

*WSU is committed to providing access and reasonable accommodation in its services, programs, activities, education and employment for individuals with disabilities. To request disability accommodation in the application process, contact Human Resource Services:* 509-335-4521 (v), Washington State TDD Relay Service: Voice Callers: 1-800-833-6384; TDD Callers: 1-800-833-6388, 509-335-1259(f), or hrs@wsu.edu.

ScienceCareers online @sciencecareers.org

ScienceCareers online @sciencecareers.org




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.

myIDP

Visit the website and start planning today!
myIDP.sciencecareers.org

ScienceCareers AAAS In partnership with:

FASEB Federation of American Societies for Experimental Biology
University of Massachusetts Medical School
UCSF University of California San Francisco
UIC
BURROUGHS WELLCOME FUND

# 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

*By* **Sakshi Ghai**

# Getting personal

It was my first time leading a team of undergraduate research assistants (RAs) on one of my Ph.D. projects, and I thought everything had gone pretty well. But I didn't want to assume they felt the same. So, I asked for their candid feedback about my leadership and collaboration style. They were a bit taken aback; after all, conversations like this are not exactly the norm in academia. As trainees, we may get regular feedback about our science, but input about "soft" skills is rarer; the common ethos is that research output is the only thing that matters. Conversations about personal style can be uncomfortable, but my prior professional experience taught me that such feedback is crucial. I wasn't going to let academic culture get in the way.

Before I began my training in academic science, I worked in a corporate setting. Early on, my bosses advised that my diffidence would chip away at others' confidence in me. Hearing that hurt. But after the pain abated, I realized they were right—and it was the push I needed to become more assertive. I began to practice using more direct language with fewer apologies and planning my comments in advance. As I took charge of my personal growth, I stopped fearing feedback and started to seek it out. I also began to document comments—negative and positive—which kept me honest about what was said and helped me decide how I wanted to grow.

**"As I took charge of my personal growth, I stopped fearing feedback and started to seek it out."**

Later, when I began my scientific training, I hoped to continue to receive regular feedback about honing my soft skills. But that rarely happened—and when it did, it wasn't constructive. When I asked, the responses typically focused on the technical content of what I said rather than how I said it. Other times, I was told I needed to "fix" my South Asian identity or that I "didn't have it in me" to be a successful academic. Offering my feedback to others didn't go much better. When I tried to tell supervisors how their leadership affected me, their defensive, angry reactions forced me to beat a hasty retreat.

Yet science is built on feedback. I am surrounded by colleagues who see silver linings in rejections, share CVs of failures, and take reviewer criticisms in stride. I wasn't ready to give up.

That's what led me to initiate the discussion with my two RAs. I began by assuring them any negative comments would not affect my letters of recommendation for them or their chance of co-authorship on publications arising from our work. Once they overcame some initial hesitation, the conversation was a great mix of constructive criticism and affirmation of my mentoring. Both flagged that I could have done a better job communicating the uncertainty of the project timelines, but they really appreciated my efforts to create a collaborative space rather than just telling them what to do.

Emboldened by this experience, a few months later I asked my current supervisor for some interpersonal feedback. I had been working with her for almost a year at that point, and although my previous attempts with other academic advisers hadn't fared well, I wanted to keep trying my luck.

I was right to keep at it: She enthusiastically agreed to reflect on what she had observed and offer feedback. When the time came, she started by highlighting my strengths, including my grit, agility, communication skills, and team-player work ethic. She then observed that I needed to stand up for myself and not undermine my own authority. This made me travel back to my early 20s when I was told I was too self-effacing—clearly I am still a work in progress. The positive comments reminded me to be kind to myself. The conversation also inspired me to find new ways to learn how to take up my rightful space, including signing up for feminist theater and acting workshops.

As my current RAs and I near the end of our time working together, I am looking forward to asking them for their feedback. I am, of course, a little nervous about what they might say. But I know they will surprise me with something about myself and leave me the better for it. ■

Sakshi Ghai is a Ph.D. student at the University of Cambridge.
Send your career story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER

# 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

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

Science
JOURNALS AAAS

## Publish your research in the *Science* family of journals

The *Science* family of journals (*Science, Science Advances, Science Immunology, Science Robotics, Science Signaling*, and *Science Translational Medicine)* are among the most highly-regarded journals in the world for quality and selectivity. Our peer-reviewed journals are committed to publishing cutting-edge research, incisive scientific commentary, and insights on what's important to the scientific world at the highest standards.

**Submit your research today!**

Learn more at **Science.org/journals**