Investigation: Editors take bribes to publish dubious papers p. 252

Long Covid markers of immune dysfunction pp. 262 & 273

Mutual neutralization of isolated water ions p. 285

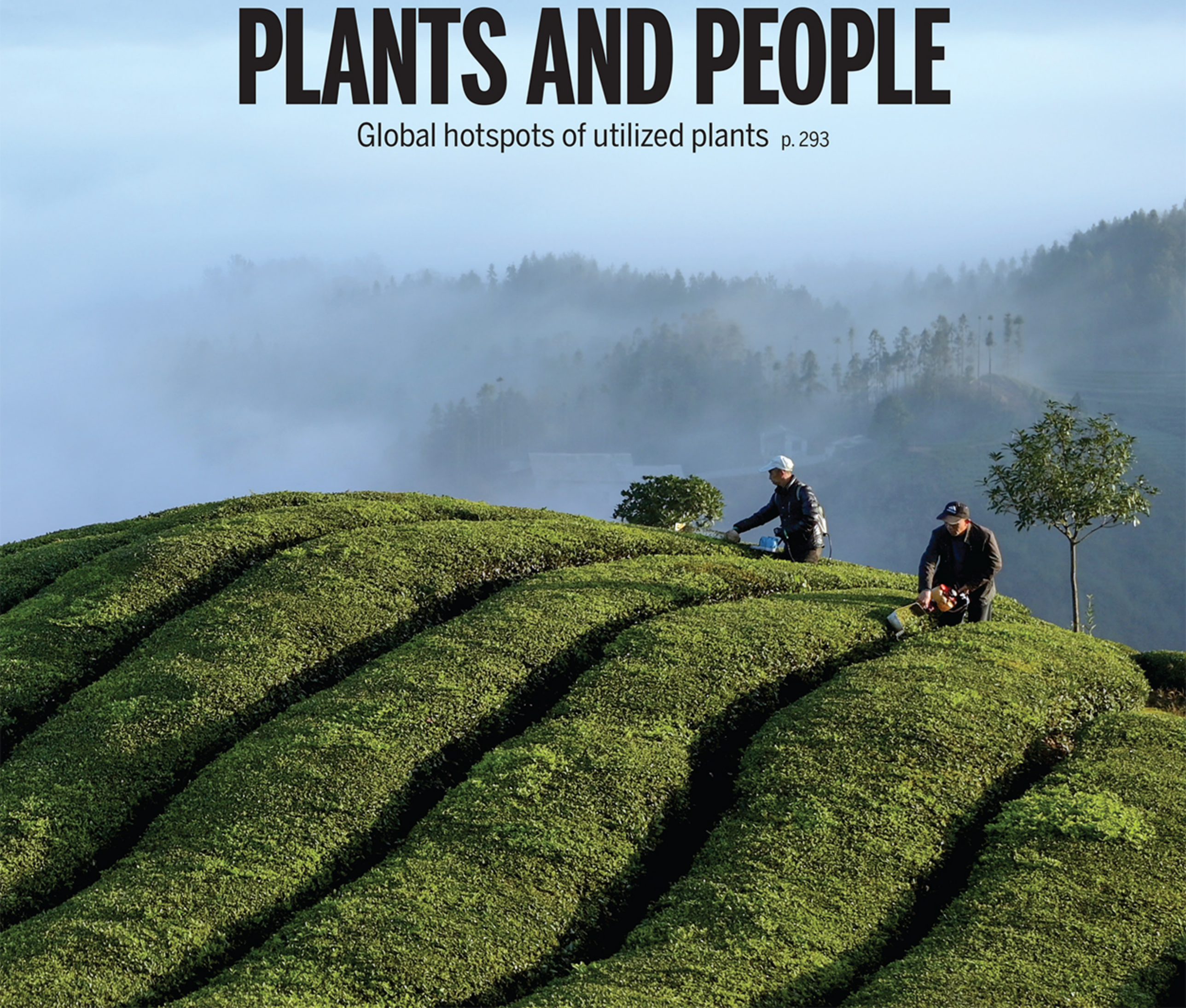
Science

$15
19 JANUARY 2024
science.org

AAAS

PLANTS AND PEOPLE

Global hotspots of utilized plants p. 293



EDITORIAL

Time to support Indigenous science

Faced with the profound challenges of a rapidly changing environment, society needs other ways of knowing to illuminate a different way forward. Thanks to the leadership of Indigenous scholars and allied collaborators, Indigenous knowledge is receiving long overdue recognition for its potential to provide solutions for the mutual thriving of lands and cultures. An urgent question is how institutions can appropriately support (and not hinder) Indigenous science's key role in creating a sustainable future.

After years of marginalization by Western science, regard for Indigenous knowledge is reaching high places. For example, in 2022, the White House called for elevating such knowledge in research, policy, and land management. This is extraordinary given the United States' track record of attempted erasure of Indigenous thought through policies of removal and forced assimilation.

There is a global groundswell of Indigenous-led research on stewardship of lands and waters, providing opportunities for Indigenous and Western knowledges to flourish together. A major step in this direction was announced last September by the US National Science Foundation, in its establishment of the Center for Braiding Indigenous Knowledges and Science (CBIKS). Led by a team of 54 predominantly Indigenous scholars and headquartered at the University of Massachusetts, Amherst, CBIKS aims to focus on complex issues at the nexus of nature and culture. The research teams, which span the globe, will address climate disruption, food insecurity, and cultural survival through learning from Indigenous community-based approaches. The goal is to identify and advance models of ethical and effective integration of Indigenous and Western sciences by creating mutually respectful and reciprocal relationships between them. CBIKS will develop generalizable approaches for a diversity of scientific communities.

CBIKS is a prime example of a model that supports research guided by the worldview and priorities of Indigenous peoples around the world. Similar initiatives in Australia, Canada, Aotearoa/New Zealand, and elsewhere are also leading the way. For too long, Indigenous peoples have been fighting for a voice in decisions regarding their lands, waters, and lives. Indigenous-led research efforts will point to different paths forward—those in which Indigenous peoples do not merely have a seat at Western science's table but are setting research agendas that reflect their priorities and protocols.

While celebrating these developments, the responsibility of institutions should not be overlooked. It is crucial that new enthusiasm not take the form of "knowledge mining," akin to a company suddenly recognizing the value of a previously overlooked mineral, rushing in to extract the ore for its own benefit, and leaving behind toxic tailings. Supporting and engaging Indigenous knowledge first and foremost involves supporting Indigenous communities. Attempts of outside actors to "incorporate" Indigenous knowledge into their own work without full consent of Indigenous communities is highly extractive and undermines the sovereignty of these communities over their own knowledge. Collaborators intent on supporting Indigenous knowledge systems might instead listen, learn, and, if requested, contribute their own knowledge or research resources to communities. They might support local governance sovereignty, the return of expropriated land, and the repatriation of ancestral remains and cultural treasures held in museums, universities, or private collections.

> "...a climate must be created that values pluralism while protecting sovereignty..."

Supporting Indigenous-led research also requires addressing the well-documented institutional barriers that limit full participation and visibility of Indigenous worldviews. Certain embedded protocols may be at odds with Indigenous ethics, values, and processes. For example, Indigenous-led research is supported by environments where the metrics of success not only include the number of scholarly papers published, but also recognize the enhanced well-being of land and culture. Countries must advance policies that support, rather than infringe upon, the wisdom, sovereignty, and rights of Indigenous peoples. To realize the transformative potential of this approach, a climate must be created that values pluralism while protecting sovereignty of diverse knowledges. In this way, solutions can emerge from the symbiosis between Western and Indigenous knowledges that benefit everyone.

For centuries, Indigenous scientists have had to adapt to, and develop fluency in, Western modes of knowledge making. It's now Western scientists' turn to learn from, and respect, Indigenous science.

–Robin Wall Kimmerer and Kyle A. Artelle

Robin Wall Kimmerer is a SUNY Distinguished Teaching Professor of Environmental Biology at the Center for Native Peoples and the Environment, State University of New York College of Environmental Science and Forestry (SUNY-ESF), Syracuse, NY, USA. rkimmer@esf.edu

Kyle A. Artelle is an assistant professor in the Department of Environmental Biology and the Center for Native Peoples and the Environment, SUNY-ESF, Syracuse, NY, USA. kaartell@esf.edu

10.1126/science.ado0684

IN BRIEF

Edited by **Jeffrey Brainard**

PALEONTOLOGY

Tyrannosaurus rex may have a new sibling

Scientists last week unveiled what they conclude is the closest relative of *Tyrannosaurus rex*—a newly discovered species that provides fresh clues about where these giant, fearsome carnivores evolved. The specimen of *T. rex*'s putative sister species, dubbed *Tyrannosaurus mcraeensis* (illustration above), consists only of skull fragments, but the research team estimates the full creature might have measured about 13 meters long, about the same as *T. rex*. The skull fragments had been hiding in plain sight for years; staff from the New Mexico Museum of Natural History and Science found them decades ago in 73-million-year-old rocks in the state's McRae Formation. They were originally thought to belong to a *T. rex*. But further analysis showed the specimen lacked *T. rex*'s characteristic ridges behind the eyes. The anatomical differences and *T. mcraeensis*'s age—up to 7 million years before *T. rex*—suggest a separate species, the researchers argued in *Scientific Reports*. They say the finding supports that a line of tyrannosaurs evolved in North America before spreading to Asia—although the leading theory suggests the reverse.

Is psychiatric guide tainted?

BIOMEDICINE | A majority of U.S.-based physicians who helped shape the latest version of a key manual of psychiatric disorders had recently received drug industry payments, a study has concluded. The findings, reported last week in *The BMJ*, raise questions about industry influence over the diagnosis and treatment recommendations. Based on financial data in the U.S. government database Open Payments, the analysis revealed that of the 92 U.S.-based physicians involved in revising the 2022 edition of the *Diagnostic and Statistical Manual of Mental Disorders*, 60% received drug industry payments in the 3 years leading up to the revision. The payments, which totaled $14 million, included consulting and speaking fees and coverage for expenses such as food and travel. The descriptions of disorders and diagnostic criteria in the manual influence which drugs get approved, prescribed, and covered by insurance. The American Psychiatric Association, which publishes the manual, told *STAT*

it managed financial conflicts in several ways, including by requiring participating physicians to disclose them.

COVID-19 booster cut strokes

PUBLIC HEALTH | Older Americans who received a COVID-19 booster vaccine were 47% less likely to develop a heart attack, stroke, or blood clot if they got sick with the disease within 6 months after inoculation compared with those who received only an original messenger RNA vaccine, a study has found. The analysis by researchers at the U.S. Centers for Disease Control and Prevention examined Medicare records of nearly 13 million people 65 and older, who are more at risk than others for these COVID-19 complications. The study followed patients who received a booster vaccine that debuted in September 2022 and attacked both the Omicron and original strains of SARS-CoV-2. Uptake of COVID-19 booster vaccines by Americans 65 and older has been poor, at less than 45%.

Hagfish reveal vertebrates' past

EVOLUTION | Hagfish are ugly, slimy, jawless vertebrates but play a starring role in recent genomic studies that raise new questions about how all vertebrates evolved. In a study last week in *Nature Ecology & Evolution* and another due out soon in *Nature*, two groups report the sequencing of a hagfish genome, filling in a key gap in vertebrate DNA data. The studies verify previous suggestions that genome sizes doubled in the branch of the vertebrate family tree leading up to cyclostomes, which include hagfishes and lampreys, another kind of jawless fish. Later, vertebrate genomes doubled again, and the new studies show this happened after cyclostomes split off from the rest of the tree. Researchers have long thought that the second doubling helped expand the diversity of animal species, but the hagfish genomics suggest other causes were at play, one of the teams concludes: Although cyclostome genomes expanded as well after the split, these animals never diversified physiologically or morphologically.

Common sense isn't common

COGNITIVE SCIENCE | Saying that something is "just common sense" may help you win an argument. But there is little consensus on what common sense actually is, a study has found, and people vastly overestimate how much the world agrees with their own personal definition of the term. A research team selected 4407 statements, including science-related facts such as "the Sun rises in the east," and personal beliefs such as "the sale of alcohol should be restricted at sporting events"; the team asked people whether they considered them to be "common sense" and whether they thought others would, too. (A total of 2046 people rated up to 50 statements each.) Few claims elicited general agreement on either point, the study's authors report this week in the *Proceedings of the National Academy of Sciences*. The paper also presents a way to measure the extent to which any statement should be labeled "common sense." That tool, the authors say, could help scientists improve artificial intelligence models; many users say their output often doesn't reflect what they see as common sense.

Cancer genomes can guide care

BIOMEDICINE | Whole-genome sequencing (WGS) of solid tumors may improve patient outcomes, researchers with the United Kingdom's ambitious 100,000 Genomes Project reported last week. In the largest study of its kind, published in *Nature Medicine*, the team analyzed WGS of 33 types of solid tumors in nearly 14,000 adult patients. Depending on type, between 3% and 95% of tumors had "actionable" mutations suggesting a specific treatment. The results support the U.K. National Health Service's plans to regularly use WGS as part of care for three types of cancer—ovarian cancers, sarcomas, and glioblastomas. Some experts noted, however, that lower cost genetic tests yield much of the same information.

NASA questions space solar power

ENERGY | Using satellites to collect solar power and beam it to Earth won't be feasible soon and even by 2050 would cost more than other forms of clean energy, NASA said in a report last week. The challenges include assembling the large solar panel arrays and developing efficient power beaming, it says. A 2022 study conducted for the European Space Agency was more optimistic about space-based solar power, concluding it could compete with other renewable energy sources as economies of scale drive down the price. Climate scientists have said the world must begin a rapid switch to clean energy sources well before 2050 to avert the worst consequences of climate change.

ILLUSTRATION: CHINA NEWS SERVICE/WIKIMEDIA COMMONS

ASTRONOMY

'Lobster eye' in space promises new look at x-rays

China last week launched an x-ray observatory with an unusual telescope inspired by the structure of lobster eyes to gather new data on gamma ray bursts, supernovae, and stars being swallowed by black holes. The Einstein Probe (illustration above)—a joint project of the Chinese Academy of Sciences, the European Space Agency, and the Max Planck Institute for Extraterrestrial Physics—will also capture x-rays from violent events that generate gravitational waves, such as two neutron stars colliding. The telescope features a survey instrument, modeled on a lobster's eye, with microscopic square tubes that can funnel x-rays from many directions to a single detector—the first time the approach has been incorporated into a major telescope. The instrument can survey the entire night sky in less than 5 hours, allowing it to discover transient x-ray–emitting sources and alert astronomers for follow-up observations.

IN DEPTH

Shelled by Russian troops in the spring of 2022, the main building of Mariupol State University stands empty.

SCIENTIFIC COMMUNITY

Uprooted Ukrainian academics reboot in exile

Many institutions and scientists displaced from occupied territories may never return

By **Richard Stone**, *in Kyiv, Ukraine*

When Ukrainian President Volodymyr Zelenskyy visited Mariupol State University (MSU) on 17 November 2023, he didn't have to slip behind enemy lines into the devastated city, occupied by Russia since May 2022. He simply had to drive a few minutes to a leafy neighborhood here in Ukraine's capital, where MSU has found a new home in exile.

Zelenskyy had come to pay tribute to the university's resilience. Since MSU relocated in April 2022, it has re-enrolled 3200 students, about 70% of the prewar number. The new facilities also include a humanitarian hub to distribute food and other essentials to some 25,000 Mariupol refugees in the Kyiv region. "This is a major achievement," says Mychailo Wynnyckyj, Ukraine's deputy minister of education and science.

MSU is one of 31 state universities uprooted from Russian-occupied territory since 2014, along with two dozen research institutes and scientific centers affiliated with the National Academy of Sciences of Ukraine (NAS). Tens of thousands of students and academics have been displaced. But not all attempts at relocation have been as successful as MSU's. With institutions losing access to key infrastructure and many displaced scholars finding havens elsewhere in Ukraine and abroad, a new challenge will loom when the war finally ends: restoring vitality to a fractured academic landscape.

MSU's own survival test began on 24 February 2022, when Russian troops marched on Mariupol, a port on the Sea of Azov. That day, MSU rector Mykola Trofymenko and colleagues filled a minibus with laptops and servers and dispatched it to the city of Dnipro, 300 kilometers to the north in central Ukraine. They followed in their own cars with vital MSU documents. Two days later, Trofymenko made a fateful decision. With many MSU staff and students still in Mariupol, he says, "I felt I should be back with them." He returned to the embattled city with his wife, their 5-year-old son, and his mother-in-law. "That was the biggest mistake of my life."

The Russian military encircled Mariupol and began bombing it relentlessly. Trofymenko and his family were trapped in their basement. During lulls, he would venture out to forage for food and for snow to melt for drinking water. After 3 weeks, regional police evacuated the family. They drove through a devastated cityscape, Trofymenko says, "but our house was standing. It was a miracle." He and his family made it safely back to Dnipro and then later Kyiv, to help set up MSU's new offices and lecture rooms in a derelict section of another university. But eight MSU staff and 10 students died during the siege of Mariupol, along with tens of thousands of others. Scores of students are still missing.

Two hundred kilometers to the west, in Melitopol, the Russian military began shelling near the Dmytro Motornyi Tavria State Agrotechnological University (TSATU), killing a student. By mid-April 2022 the situation had grown untenable. TSATU Rector Serhii Kiurchev and his colleagues scooped up official documents and drove 100 kilometers north to Zaporizhzhia, running a gauntlet of 20 Russian checkpoints where men of fighting age were often detained.

That May, occupiers seized TSATU's campus and renamed it Melitopol State University—a "pseudouniversity," Kiurchev says. He set up TSATU's administration in Zaporizhzhia, itself a frontline city where shelling killed a student. TSATU's courses are now entirely online, and it boasts 8000 students, including 2300 freshman—more than any other displaced university this year, Kiurchev says.

Other research centers in Ukraine were uprooted earlier, after Russia's 2014 invasion of Crimea and parts of Donetsk and Luhansk—the Donbas region. The Donetsk Institute for Physics and Engineering now has footholds in Kyiv and Kharkiv. But it had to abandon millions of dollars of equipment including a facility for synthesizing oxide nanopowders, a scanning electron microscope, and a cryogenic site for helium production.

At the Donetsk Botanical Garden (DBG), many scientists and nearly all technical staff chose to stay put when separatists took control a decade ago. Botanist Ganna Boiko decided to leave, fleeing to Kyiv with her husband and 12-year-old son, in part because of bad blood on the staff. "It wasn't just the shelling," she says. "It was the tension between people"—separatists versus those loyal to Ukraine or indifferent to the conflict.

If Ukraine can reclaim its occupied territories, officials plan to vet academics in rogue institutions to determine who deserves a second chance. "Some people were just victims of Russian propaganda," says Volodymyr Ustymenko, head of the Donetsk Scientific Center, which coordinates scientific activities of universities and NAS institutes in Donbas.

The Ukrainian government is also drawing up plans to lure academics back to the territories it intends to recapture. "Science must be an engine of the future economic growth in Donbas," says Ustymenko, a legal scholar. Incentives will be needed to attract high-tech investments, he says. "We cannot just restore the academic landscape as it was."

Still, the longer the war lasts, the harder it will be to coax back academics who have started over in Ukraine and abroad. Boiko, now vice president of the Ukrainian Botanical Society, landed a job at NAS's M.G. Kholodny Institute of Botany in the capital; her son studies at the Igor Sikorsky Kyiv Polytechnic Institute. "We aren't going back," she says. Population geneticist Sergiy Pryvalikhin, who also left DBG in 2014, thought he and his wife would soon return. Now, he has a plum job at NAS headquarters—and two young children who were born here. "They don't know Donetsk," he says.

Some displaced universities might also balk at returning. Before 2014, Donetsk National University (DNU) was one of the country's top universities; that year it set up shop in Vinnytsia, in central Ukraine. DNU has woven itself into that community, says Wynnyckyj, a sociologist by training. "It's no longer a question of relocation. If deoccupation happened tomorrow, we'd have to rebuild the university in Donetsk from the ground up." Still, he says several frailer Donbas universities failed to take root in their new environs, and "have effectively died."

MSU will return to its hometown after liberation, Trofymenko says. Until then, as Zelenskyy observed during his visit to MSU-in-exile, Ukraine must keep the educational flame alive for its displaced academics. "What I saw was more than a university," Zelenskyy said. "Without students, there will be no future for which our country is fighting." ■

Richard Stone's reporting was supported in part by the Richard Lounsbery Foundation.

SYNTHETIC BIOLOGY

Bacteria stitch exotic building blocks into novel proteins

Efficient method for creating proteins with unusual amino acids opens the door to new medicines and catalysts

By **Robert F. Service**

Biology students learn that DNA codes for just 20 amino acids, the building blocks of all of life's proteins. Researchers would love to build proteins from hundreds of other, exotic amino acids, but life's preference for standard ones has hindered the quest. Now, researchers report an efficient way to coax bacteria to add structurally unusual amino acids into proteins. The approach could lead to medicines that persist longer in the body and better catalysts for industry.

"It's a big accomplishment to get these new categories of amino acids into proteins," says Chang Liu, a chemist at the University of California, Irvine who was not part of the study, which was published last week in *Nature*.

The 20 common amino acids (plus two rare ones) that living things readily use are known as alpha amino acids, which refers to their common structure. Researchers have tinkered with cells' proteinmaking machinery to accept dozens of nonstandard but still closely related alpha amino acids. But hundreds of even more exotic amino acids exist, including beta and gamma varieties that have unique twists and turns in their molecular backbones. To date, these have only been added to proteins in test tubes by synthetic chemists.

Cells make proteins in two key steps: ferrying amino acids to the ribosome, the cell's protein assembler, and coaxing the ribosome to stitch them onto a growing protein. Delivering the amino acids is the job of short RNA snippets called transfer RNAs (tRNAs). Each tRNA has a three-letter genetic sequence that codes for a specific amino acid. Enzymes known as aminoacyl tRNA synthetases (aaRS) attach the proper amino acid to each tRNA. The loaded tRNAs then travel to the ribosome, where they bind to corresponding sequences on a long strand of so-called messenger RNA (mRNA), which contains a full-length copy of a gene to be translated into a protein. The ribosome goes down the strand of mRNA, plucking off each amino acid and linking them together in a growing chain.

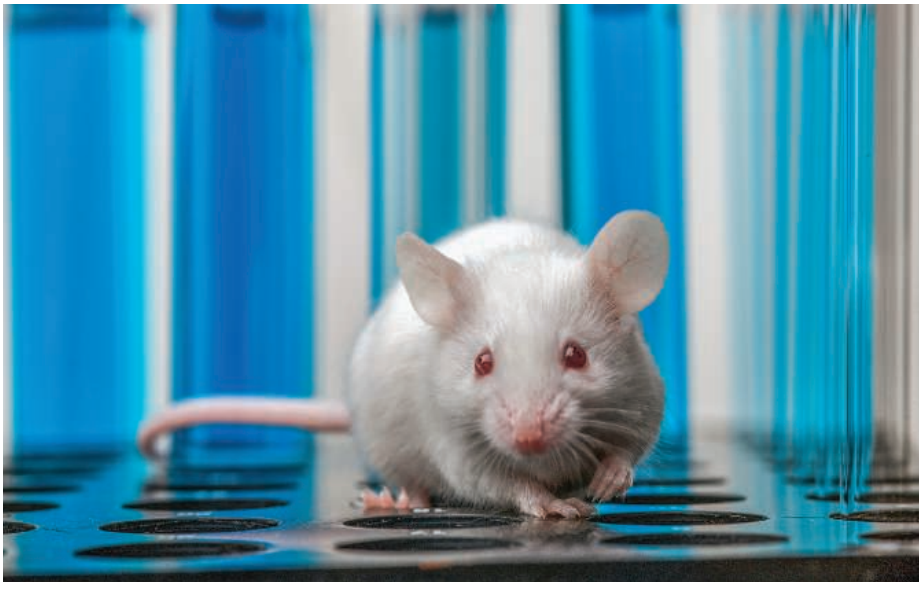
Alexandria Deliz Liang, a chemist at the University of Zurich, likens proteinmaking to assembling a train: First you must load the railway cars, and then you need to link those cars together. To produce novel proteins, researchers have to get both steps—the loading and the linking—to work at once. "If either of those don't work, the system fails," adds Jason Chin, a chemist at the U.K. Medical Research Council's Laboratory of Molecular Biology (LMB).

Chin and his LMB colleagues have now found a better way to do the first step. By mutating genes for the aaRS enzymes, they created millions of alternative versions that might bind to exotic amino acids. They also attached RNA tags to track which enzymes worked out of the mix of millions. They inserted genes for these enzymes into *Escherichia coli* bacteria and watched to see whether their ribosomes would incorporate the exotic amino acids into proteins.

The researchers found eight enzymes that successfully loaded exotic amino acids, and *E. coli*'s ribosomes were able to add four into growing protein chains: three beta amino acids and an equally exotic variety known as an alpha,alpha-disubstituted amino acid. "We broke the deadlock," Chin says.

The achievement will likely have big implications down the road, says Samuel Gellman, a chemist at the University of Wisconsin–Madison. For starters, he says, companies could design protein-based drugs resistant to bodily enzymes that readily degrade alpha amino acid–based proteins, reducing the need for repeat doses. And because exotic amino acids have different shapes than standard ones, the approach could also improve industrial catalysts involved in making everything from pharmaceuticals to fuels.

Chin's process still relies on the willingness of the ribosome to occasionally accept oddball amino acids. So his group is also trying to alter the ribosome itself, by making mutations that allow it to recognize tRNA codes not found in nature and tolerate amino acids with unusual shapes. As the loading and linking lines of research come together, Chin says researchers could also engineer polymer materials composed entirely of exotic amino acids, which could lend them novel properties. ■



ANIMAL RESEARCH

EPA scraps plan to end all testing in mammals by 2035

U.S. environmental agency's hard deadline had split scientific community

By **David Grimm**

The U.S. Environmental Protection Agency (EPA) has abandoned a controversial plan to phase out all use of mammals to test the safety of chemicals by 2035. The hard deadline—imposed in 2019 to accelerate a move toward nonanimal models such as computer programs and "organs on a chip"—made EPA unique among U.S. federal agencies. But it divided scientists, some of whom say animals remain the gold standard for assessing the safety of chemicals that could harm humans and wildlife.

Removing the deadline is a "good move," says Jennifer Sass, a senior scientist at the Natural Resources Defense Council who blasted the 2035 target. "I'm not comfortable with EPA calling a chemical safe based just on cells in a petri dish."

Former EPA Administrator Andrew Wheeler, who set the deadline, fears that without it, the agency won't get serious about reducing its reliance on animal experiments. "The status quo will continue," he says. "It's taking the easy way out."

EPA says the reversal was motivated by science. "We need to focus on what the science is telling us in order to advance methods that don't involve animal testing," says Chris Frey, the agency's assistant administrator for R&D—"and not focus so much on arbitrary dates."

EPA relies on studies of thousands of animals each year for assessing the safety of pesticides, determining dangerous levels of chemicals in drinking water, and other key toxicity studies. Much of this work is done in mammals such as rodents and rabbits, though some is also done in fish. A 2016 amendment to the federal Toxic Substances Control Act required the agency to move away from animal experiments, although it set no deadline.

In September 2019, Wheeler pledged that EPA would stop conducting or funding any studies on mammals by 2035. He also set an interim goal of 2025 for reducing these studies by 30%. Other federal agencies have long vowed to reduce their reliance on animal testing—a 2020 U.S. spending bill even compelled some to do so—but none has set a hard deadline for ending animal research.

At the time, critics accused Wheeler—who served under former President Donald Trump—of kowtowing to the chemical industry, which views animal testing as time consuming and expensive. (Chemical companies pay for much of this testing themselves to satisfy regulatory requirements. But Wheeler says he was motivated by his family's love of animals and his belief that new technologies would be able to fully replace animal testing within 16 years.

Now, both the 2025 and 2035 deadlines are gone. Last month, the White Coat Waste Project, an advocacy group pushing to end taxpayer-funded animal experiments, brought attention to agency documents indicating the change of plans—a development first reported by *The Washington Times*. An EPA work plan for reducing animal use published in June 2020 mentions both deadlines, but they're gone in a December 2021 version of the report. An internal agency email obtained by White Coat Waste via a public records request also refers to specific target dates being "stripped out."

"They made no effort to meet these benchmarks—they didn't even try," says Justin Goodman, White Coat Waste's senior vice president of advocacy and public policy. "The deadlines were the most important part of the work plan," he says; without them, any effort to reduce animal testing is "toothless."

Wheeler, now a volunteer adviser for White Coat Waste, says he's disappointed by the reversal. "I felt like things were moving in the right direction," he says. "Without a deadline, we're not going to make progress."

EPA tells *Science* that it remains committed to exploring alternatives to animal experiments. Just because the dates are gone, "that commitment has not changed," Frey says, although he doesn't know when—or even whether—his agency will completely transition to nonanimal models. "Fully phasing out animal testing is the goal, and we will always have that goal," he says. "But I don't want to get ahead of our scientists."

Sass applauds the move. In March 2023, her organization was one of dozens of environmental, health, and social justice groups that sent a letter to EPA urging it not to abandon animal testing. New technologies are useful as a supplement to animal experiments, Sass says, "but I'm not sure that they'll ever be able to tell us whether a child exposed to lead as a fetus is going to have trouble sitting still in a classroom."

Kim Boekelheide, a toxicologist at Brown University, expects animals to disappear from environmental testing soon, however. His lab is developing clusters of cells called organoids that mimic aspects of the heart, brain, and other organs for evaluating the impact of pollutants, and he believes that such approaches are finally beginning to rival the predictive power of animal studies. "The transition from animals to in vitro test systems will happen in the next decade, whether there's a deadline or not." ■

PALEONTOLOGY

A mammoth's life story, written in tusk

The travels of "Elma" show she faced twin pressures—climate change and human hunting

By **Michael Price**

First, the 13,000-year-old woolly mammoth whose tusks were found in 2009 near Fairbanks, Alaska, needed a name. "They wanted a name from the Indigenous community," recalls Evelynn Combs, an archaeologist and Mendas Cha'ag tribal member. The Healy Lake Village Council settled on Élmayųujey'eh, "an affectionate nickname for things that look funny" in the tribe's Dené language, Combs explains.

Then, Elma (for short) needed a life story, which a detailed analysis of the tusks has now provided. Her travels are giving Combs and colleagues a rare glimpse into the ways of her species at the end of the last ice age—and insight into how pressure from a changing climate as well as hunting by early humans may have helped spur mammoths' extinction.

"It's a lovely piece of work," says paleoenvironmental researcher Rachel Schwartz-Narbonne of Sheffield Hallam University. She praises the motley collection of disciplines—geochemistry, genetics, archaeology, and oral traditions—that went into the work, published this week in *Science Advances*. "It is a mammoth study, pun intended."

A 13,000-year-old mammoth's life history reveals stresses on the species as the ice age ended.

ILLUSTRATION: SCIENCE PHOTO LIBRARY/ALAMY

Like elephant tusks, those of mammoths were always growing, building up layers of ivory in growth rings a bit like those of trees. As mammoths roamed, ate, and drank, they ingested chemical markers called isotopes, which embedded themselves in those layers. The ratios of isotopes vary from place to place, depending on geology, so scientists can reconstruct a mammoth's lifetime travels based on the layered isotopic signatures in their tusks. In 2021, a team led by scientists from the University of Alaska Fairbanks (UAF) used the method to chart the life history of a 17,000-year-old male mammoth nicknamed Kik (*Science*, 13 August 2021, p. 806).

Kik ranged across the lowlands of the Alaskan interior, grazing on herbs and grasses during the frigid height of the last ice age. Elma, born 4000 years later, spent her 20 years of life in the ice age's waning days, as glaciers thawed and vegetation changed. Applying the same analytic methods, the researchers found her range significantly overlapped with Kik's but was much smaller. She favored the highlands, where mammoths' preferred vegetation still lingered.

Elma spent most of her time in two regions about 1000 kilometers apart, one in eastern Alaska, the other in western Yukon. Both are rich in signs of human presence. Artifacts at Elma's final resting place, Swan Point, date back 14,000 years. The site "is largely regarded as the earliest unequivocal evidence for human occupation in Beringia," the now partly submerged land between Alaska and Siberia, says archaeologist and co-author Ben Potter of UAF.

The fact that Kik and Elma largely traveled the same trails, despite being separated by 4000 years, suggests consistent, predictable movement patterns, the authors say. As with modern game animals such as elk, that may have allowed humans to plan hunting camps around their presence. Swan Point contains the same types of small, razorlike projectile blades known to have felled Siberian mammoths some 30,000 years earlier.

"The idea that humans would have been aware of and influenced by mammoth behavior … makes complete sense," says Larisa Grawe DeSantis, a paleontologist at Vanderbilt University. Existing Dené dialects may reflect that ancient awareness, says study co-author Gerad Smith, an anthropologist at the University of Alaska Anchorage. Several preserve a word, *negutih* and variations thereof, that roughly translates to "a creature carrying a singular object in front of its face"—perhaps a mammoth and its trunk.

DNA from Elma's tusks revealed she was closely related to previously excavated juvenile and infant mammoths at Swan Point. "At the very least, the juvenile and neonate were probably hunted by humans," says lead author and UAF graduate student Audrey Rowe. Elma herself might have been felled by humans: Isotopes in the outermost layers of her tusks suggest she was healthy and well-nourished when she died.

That makes sense to paleobiologist Hervé Bocherens at the University of Tübingen. "Humans are the only regular predators of mammoths, and probably the only ones able to kill prime adults," he says.

Except for a few island stragglers, woolly mammoths largely vanished about 10,000 years ago. For decades, scientists have argued over why. The new work suggests human hunting pressure, in concert with shrinking habitats and altering food availability, may have sealed the mammoths' fate, the authors say. That echoes a similar conclusion published last year that both shifting climate and human-caused wildfire sparked the extinction of megafauna in Southern California as the ice age ended (*Science*, 18 August 2023, p. 724).

"We do see coexistence" between humans and megafauna during the ice age, Schwartz-Narbonne says. "When we add in two harsh pressures—hunting and climate change—at the same time, that's what eventually leads to extinction."

Besides illuminating the past, the authors think their findings could help efforts to resurrect mammoths using their ancient DNA. Knowing how mammoths lived and where they roamed in their ice-age heyday would be critical if that long-shot project actually succeeds. ■

BIOSAFETY

Wisconsin bill would restrict pathogen studies

Efforts to ban "gain-of-function" research on viruses and bacteria worry scientists

By **Jocelyn Kaiser**

A hearing last week in Wisconsin on a proposal to bar research that may make human pathogens more dangerous has scientists worried that such state bans would hamper a broad range of microbiology studies. The laws could also conflict with or exceed new federal restrictions on so-called gain-of-function (GOF) research, due out any moment.

The Wisconsin bill is still early in the legislative process, and its prospects are uncertain. But the hearing gave a platform to a small, vocal group of scientists who want to forbid certain GOF studies. Several hold the controversial view that such research is to blame for the COVID-19 pandemic, claiming that the causative virus "leaked" from a Chinese lab that had, accidentally or intentionally, made a less harmful virus deadlier.

Yoshihiro Kawaoka (center) once modified a virus to spread more easily, now prompting bills to ban similar work.

Many virologists and others disagree, favoring a natural origin for the pandemic. Gigi Kwik Gronvall, a biosecurity expert at Johns Hopkins University, says the bill "will create uncertainty" that will deter the state's scientists from doing important work to prepare for future pandemics. "Their hearts are in the right place—I mean, who doesn't want to stop pandemics and prevent lab accidents?" she says of the bill's sponsors. "But this bill won't do that."

The statewide GOF bans, which have also been proposed in Texas and are law in Florida, reflect a debate sparked 13 years ago when two labs funded by the U.S. National Institutes of Health (NIH) modified the potent H5N1 bird flu virus to make it spread more easily between ferrets—and potentially among people. Although meant to help prepare for pandemics, the work raised unacceptable risks of a lab leak, some GOF research critics said. The uproar led to a 2017 U.S. policy requiring a high-level government review for such federally funded studies. But COVID-19 has stoked new controversy, with some scientists and lawmakers alleging that NIH-funded work on bat coronaviruses in Wuhan, China, sparked the pandemic.

Several proposals in Congress, including one approved last year by the House of Representatives, would replace the review with a ban on federally funded research on certain bacteria, viruses, and fungi, dubbed enhanced potential pandemic pathogens (ePPPs). It's not clear the Democratic Senate will endorse any of them.

Last year, Florida enacted the first state-level GOF ban. Its law uses the current, narrow federal definition of an ePPP, and analysts found it would have no impact because no such research is underway in Florida.

The University of Wisconsin–Madison (UW), however, is home to virologist Yoshihiro Kawaoka, who led one of the 2011 studies that set off the GOF debate. Kawaoka, who is also at the University of Tokyo, continued his H5N1 work under the 2017 rules but has since ended it. Last year, concerns about past safety incidents in his UW lab led to a House committee requesting a long list of documents related to Kawaoka's research.

The safety concerns also helped inspire Wisconsin bill AB413, which would effectively ban ePPP studies by barring state funding for institutions that allow this work. When scientists enhance pathogens, "You're playing with a dangerous, dangerous thing. I don't think the science is worth the risk," the bill's sponsor, Wisconsin State Assembly Representative Elijah Behnke (R), told *Science*. It would also require that any proposed study on natural, unmodified microbes that qualify as PPPs—the bill offers no list—be reported to the state. A UW spokesperson said Kawaoka declined to comment on the bill.

The bill is copied from a Texas proposal that died last year. Both bills use a broad definition of pathogens that could cause a pandemic. Critics say it could cover any that spreads easily among people, such as cold viruses, blocking routine studies that modify such microbes to understand their biology and develop vaccines and treatments.

It would "significantly hinder the ability of researchers in Wisconsin to conduct research of extreme importance to the state" on human, plant, and animal pathogens, UW and the private Medical College of Wisconsin warned in testimony submitted for the 10 January hearing. Last year, the UW system estimated that the reporting rule would affect thousands of studies and could jeopardize federal grants by delaying the work.

One witness at the Wisconsin Assembly committee hearing was Justin Kinney, a quantitative biologist at Cold Spring Harbor Laboratory and co-founder of Biosafety Now, a nonprofit that supports the lab-leak pandemic theory. Although the group has rallied behind the Wisconsin bill, Kinney said it should use a narrower ePPP definition proposed by a federal advisory panel. He said that definition would ban work in only one lab—Kawaoka's—and require reporting studies of just a handful of risky human pathogens, such as H5N1, SARS-CoV-2, Ebola, and mpox. "There will be no impact at all on the vast majority of faculty," Kinney said.

"We are more than open" to the change, Behnke says. Even if the bill passes, however, Wisconsin's Democratic governor is unlikely to sign it. Meanwhile, microbiologists are anxious to see the new federal ePPP policy. Congress had ordered that it be completed by the end of 2023. ■

The Greenland Telescope was one of many radio dishes that added to the picture of M87's black hole.

PHOTO: NIMESH PATEL/CENTER FOR ASTROPHYSICS | HARVARD & SMITHSONIAN

ASTRONOMY

Second image of 'shadow' confirms giant black hole is real

To zoom in farther, Event Horizon Telescope wants to add more radio dishes to its network—and go to space

By **Daniel Clery**

A familiar shadow looms in a fresh image of the heart of the nearby galaxy M87. It confirms that the galaxy harbors a gravitational sinkhole so powerful that light cannot escape, one generated by a black hole 6.5 billion times the mass of the Sun. But compared with a previous image from the network of radio dishes called the Event Horizon Telescope (EHT), the new one reveals a subtle shift in the bright ring surrounding the shadow, which could provide clues to how gases churn around the black hole.

"We can see that shift now," says team member Sera Markoff of the University of Amsterdam. "We can start to use that." The new detail has also whetted astronomers' desire for a proposed expansion of the EHT, which would deliver even sharper images of distant black holes.

The new picture, published this week in *Astronomy & Astrophysics*, comes from data collected 1 year after the observing campaign that led to the first-ever picture of a black hole, revealed in 2019. The dark center of the image is the same size as in the original image, confirming that the image depicts physical reality and is not an artifact. "It tells us it wasn't a fluke," says Martin Hardcastle, an astrophysicist at the University of Hertfordshire who was not involved in the study. The black hole's mass would not have grown appreciably in 1 year, so the comparison also supports the idea that a black hole's size is determined by its mass alone.

In the new image, however, the brightest part of a ring surrounding the black hole has shifted counterclockwise by about 30°. That could be because of random churning in the disk of material that swirls around the black hole's equator. It could also be associated with fluctuations in one of the jets launched from the black hole's poles—a sign that the jet isn't aligned with the black hole's spin axis, but precesses around it like a wobbling top. That would be "kind of exciting," Markoff says. "The only way to know is to keep taking pictures."

Although incredibly massive, black holes like M87 are relatively small: One could fit within the Solar System. To see it, the EHT relies on radio emissions, which penetrate the gas and dust shrouding the black hole. The array generates a sharp enough image by combining data from telescopes spaced as widely apart as possible, creating, in effect, an Earth-size dish. So far, the EHT team—some 300 researchers, also spread across the world—has observed for a couple of weeks each year, using up to a dozen observatories from the South Pole to Greenland and Hawaii to France.

Now, the team wants to add more telescopes to the network, which would further sharpen its images and enable it to see black holes in more distant galaxies. Last month, the team submitted a proposal to the National Science Foundation for a $73 million next-generation EHT (ngEHT). It calls for building 9-meter radio dishes in four locations—Wyoming, the Canary Islands, Chile, and Mexico—and adding the 37-meter dish at the Massachusetts Institute of Technology's Haystack Observatory. This "fills in the holes beautifully," says EHT founding director Shep Doeleman of the Center for Astrophysics | Harvard & Smithsonian (CfA).

The ngEHT project would also deploy new hardware and software to speed up data processing. The team could produce results in days instead of years, creating the opportunity for "black hole cinema," Doeleman says. Sharpening the resolution in both space and time will help researchers unravel one of astrophysics' most enduring mysteries: how a black hole's spin, magnetic field, and swirling disk of material conspire to launch the powerful jets of particles far out into space.

Theorists also predict that within the fuzzy bright ring in the EHT images are sharper circles of light emitted by photons trapped in orbits as close as it's possible to get to the black hole's boundary, or event horizon. To zoom in on these so-called photon rings will require a virtual radio dish extending well beyond Earth. "We are reaching the limits of what can be done from the ground," says CfA's Michael Johnson. He is leading a mission proposal called Black Hole Explorer that he aims to submit to NASA in 2025. To be launched in 10 years into geosynchronous orbit, the spacecraft's 4-meter dish would expand the baseline of ngEHT to about 35,000 kilometers, enough to see the photon rings. "With just one satellite we could pick out these orbits," Markoff says.

Theory predicts multiple, nested photon rings, and the size and shape of the innermost one would help pinpoint the black hole's mass, spin, and even the inclination of its swirling disk. By imaging multiple black holes with different life histories, researchers could find out how their spins are affected by mergers or a sudden feast of surrounding material.

The ngEHT and its orbiting outrigger should also be able to test Albert Einstein's theory of gravity, general relativity, in the most extreme conditions ever, Doeleman says. "[It will] bring us as close to the edge of a black hole event horizon as we are likely to be for many years to come." ■

PAPER TRAIL

In the latest twist of the publishing arms race, firms churning out fake papers have taken to bribing journal editors

By **Frederik Joelving,**
Retraction Watch

One evening in June 2023, Nicholas Wise, a fluid dynamics researcher at the University of Cambridge who moonlights as a scientific fraud buster, was digging around on shady Facebook groups when he came across something he had never seen before. Wise was all too familiar with offers to sell or buy author slots and reviews on scientific papers—the signs of a busy paper mill. Exploiting the growing pressure on scientists worldwide to amass publications even if they lack resources to undertake quality research, these furtive intermediaries by some accounts pump out tens or even hundreds of thousands of articles every year. Many contain made-up data; others are plagiarized or of low quality. Regardless, authors pay to have their names on them, and the mills can make tidy profits.

But what Wise was seeing this time was new. Rather than targeting potential authors and reviewers, someone who called himself Jack Ben, of a firm whose Chinese name translates to Olive Academic, was going for journal editors—offering large sums of cash to these gatekeepers in return for accepting papers for publication.

"Sure you will make money from us," Ben promised prospective collaborators in a document linked from the Facebook posts, along with screenshots showing transfers of up to $20,000 or more. In several cases, the recipient's name could be made out through sloppy blurring, as could the titles of two papers. More than 50 journal editors had already signed on, he wrote. There was even an online form for interested editors to fill out.

"Jackpot!" Wise thought, and then, "Oh geez, I'm going to have to report this."

AT LEAST TENS OF MILLIONS of dollars flow to the paper mill industry each year, estimates Matt Hodgkinson of the independent charity UK Research Integrity Office, which offers support to further good research practices, who is also a council member at the nonprofit Committee on Publication Ethics. Publishers and journals, recognizing the threat, have beefed up their research integrity teams and retracted papers, sometimes by the hundreds. They are investing in ways to better spot third-party involvement, such as screening tools meant to flag bogus papers.

So cash-rich paper mills have evidently adopted a new tactic: bribing editors and planting their own agents on editorial boards to ensure publication of their manuscripts. An investigation by *Science* and Retraction Watch, in partnership with Wise and other industry experts, identified several paper mills and more than 30 editors of reputable journals who appear to be involved in this type of activity. Many were guest editors of special issues, which have been flagged in the past as particularly vulnerable to abuse because they are edited separately from the regular journal. But several were regular editors or members of journal editorial boards. And this is likely just the tip of the iceberg.

Hodgkinson recalls hearing one publisher say it "had to sack 300 editors for manipulative behavior." He adds, "These are organized crime rings that are committing large-scale fraud."

Ben seemed to view co-opting editors as normal business procedure. Reached by phone, he appeared to believe he was being approached by a journal editor looking to collaborate, despite repeatedly being told he was talking to a journalist.

"I have many customers [who] want to publish," Ben said. He added that he needed partners to help get his papers into journals.

"First time we will pay like this: after accept, half, and after paper online, half," Ben explained, noting that the kickback's size would depend on the journal. "You can offer your price."

When he realized he was not speaking with a journal editor, Ben asked to switch to WhatsApp. In a written exchange he denied paying editors, claiming his company only offered advice about manuscripts, and most of the incriminating posts on his Facebook profile vanished.

But Olive Academic's relationship with an editor named Malik Alazzam belies Ben's claim. On LinkedIn, Alazzam describes himself as an "editor of Scopus and ISI journals," referring to journals included in two leading reputable databases, as well as a former researcher and assistant professor in Saudi Arabia, Malaysia, and Jordan. (He did not agree to be interviewed for this story.) Alazzam's connection to Olive Academic is apparent from the screenshots in Ben's Facebook posts recruiting new editors and advertising to authors. One of the two papers whose titles could be discerned, "Influencing Factors of Gastrointestinal Function Recovery after Gastrointestinal Malignant Tumor,"

ILLUSTRATION: (OPPOSITE PAGE) DAVIDE BONAZZI

was published in a special issue of Hindawi's *Journal of Healthcare Engineering* in 2021—and edited by Alazzam. Three days after the article was accepted, the screenshots show Olive Academic paid $840 to Tamjeed Publishing; the company's website lists Alazzam as the sole member of the team, and Alazzam's LinkedIn profile says he is an editor there. Other payments, of up to $16,300, showed the first and last letters of the recipient's name: "M" and "ZZAM."

Wise believes Tamjeed's activity goes beyond Alazzam and that the company acts as a broker, sharing payments from the paper mills with multiple editors—including Omar Cheikhrouhou of Taif University in Saudi Arabia and the University of Sfax in Tunisia. Cheikhrouhou was the editor for the other identifiable paper from Ben's Facebook posts, "Relationship between Business Administration Ability and Innovation Ability Formation of University Students Based on Data Mining and Empirical Research," which brought in $1050 for Tamjeed 2 days after acceptance in a special issue of Hindawi's *Mobile Information Systems*. (Cheikhrouhou stopped responding to messages after *Science* requested to interview him.) Cheikhrouhou and Alazzam have both edited other Hindawi special issues and are currently guest editors for several journals published by the Multidisciplinary Digital Publishing Institute (MDPI) and IMR Press.

The two identified papers were retracted on 1 November 2023, when Hindawi and its parent company, Wiley, pulled thousands of papers in special issues because of compromised peer review. (In December, Wiley announced it will "sunset the Hindawi brand.") "Over the past year, we have identified hundreds of bad actors, present in our portfolio and others, some of whom held guest editorial roles," a Wiley spokesperson told *Science* by email. "These individuals have since been removed from our systems."

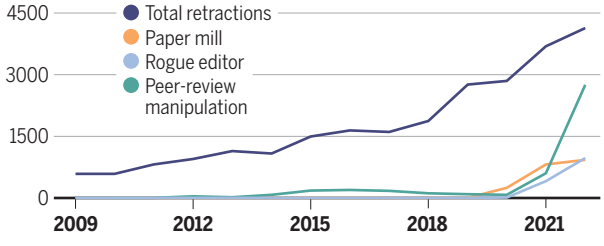
OLIVE ACADEMIC AND TAMJEED are far from the only firms employing editors with questionable credentials, or even made up from whole cloth. A Ukrainian paper mill dubbed Tanu.pro, for example, appears to have planted an editor who was either still a student or had just obtained her master's degree, leveraging journals' sometimes lax vetting process for editors, according to Anna Abalkina, a social scientist at the Free University of Berlin who identified and described the scheme in a recent preprint.

The editor, Liudmyla Mashtaler, accepted several papers linked to the paper mill through the email addresses used for a 2022 special issue of *Review of Education*, a title copublished by Wiley and the nonprofit British Educational Research Association (BERA). (The papers were retracted on 5 November 2023, after Abalkina's preprint appeared.) Mashtaler went on to become a member of the journal's editorial board. Abalkina found no evidence that Mashtaler has a doctorate, even though she was listed on the editorial board website as "Dr."; a 2020 Ukrainian government document refers to her as a first-year master's student. "This is a scandal," Abalkina says.

Mashtaler, who disappeared from the journal's editorial board after *Science* contacted the publisher for this story, continues to edit special issues, sometimes under the last name Obek. She did not respond to repeated emails. BERA said it was working "to tighten procedures for identifying fraudulent activity, including paper mills, following this experience."

Alarming trend

Retractions linked to questionable publishing practices have grown disproportionately, according to Retraction Watch's database. "Rogue editor" and "peer-review manipulation" can both signal paper mill involvement. (Multiple reasons can be assigned to a single retraction.)



In another case, the editors of a special issue in Hindawi's *Scientific Programming* identified via Olive Academic's ads did not appear to correspond to real people at all. Wise believes the paper mill itself organized the special issue from start to finish—a tactic also described by a scientist who graduated from a medical school in China and tracks paper mills in that country. In such cases the paper mills handle all the correspondence with the journal, including proposing the issue in the first place, either through a real academic colluding with it or by inventing a fake identity for the occasion. "The latest generation paper mill, they're like the entire production line," says the researcher, who requested anonymity for fear of retaliation against family members in China.

The problem goes beyond special issues. Of nearly a dozen editors of special issues linked to Olive Academic through ads posted by the company on Chinese social media sites, the majority have also held regular editor posts at journals published by Wiley, Elsevier, and others. These include Oveis Abedinia, an electrical engineer at Nazarbayev University in Kazakhstan who until last year was a regular editor of *Complexity*, published by Hindawi in partnership with Wiley. (Abedinia did not respond to interview requests via phone or email.) Tamjeed Publishing also appears to have targeted *Complexity*; on social media, Alazzam listed it as one of the journals his company has "contracted" and invited researchers to publish there.

COLUMBIA UNIVERSITY Ph.D. student Siddhesh Zadey has firsthand experience of such marketing. While he was visiting his parents in India last summer, a Dr. Sarath of iTrilon reached out to him on WhatsApp, offering authorship of "ready-made papers" with "100% Acceptance Guarantee." Angling for more information, Zadey—who is also co-founder of the India-based think tank ASAR, which addresses social problems through research—pretended to be a clueless medical student. "Is the article already accepted?" he asked Sarath. "This says 100% acceptance."

"Means we have network with Journal editors," Sarath replied. "So we can guarantee Acceptance."

One of the journals Sarath claimed to be working with was *Health Science Reports*, published by Wiley. A spokesperson for the publisher said it had recently issued retractions in the journal "due to peer review manipulation, and there are additional investigations ongoing."

In an interview, Sarath acknowledged selling authorship but denied iTrilon colluded with editors. "Just we rely on the work," he said.

However, papers linked to the company reveal likely editor involvement. In Sarath's pitch to Zadey, he touted five author slots available on an already-accepted "original research article." The paper went on to be published in the journal *Life Neuroscience* just 14 days after the ad was posted, with six total authors—two of whom are also high-level editors at the journal.

One of them was the paper's corresponding and final author—Nasrollah Moradikor, director of the International Center for Neuroscience Research (ICNR) in Georgia, where Sarath told Zadey the work was conducted. (Other authors on the paper are based in India, South Korea, and Spain.) Moradikor did not agree to be interviewed. But as corresponding author, he must

have been aware of the postacceptance author additions.

The other author-editor, Indranath Chatterjee, a professor of computer science at Tongmyong University in South Korea, told *Science* he did not know what kind of services iTrilon provides nor that his paper had been advertised by the company. But he acknowledged there had been authorship changes on the paper because "some expertise of some other people" had been required. In September 2023, he gave a talk on scientific publishing organized by iTrilon and ICNR. Both Moradikor and Chatterjee are also editors at other journals, Chatterjee as a section chief editor at *Neuroscience Research Notes* and Moradikor as a guest editor for publishers such as MDPI, De Gruyter, and AIMS Press.

PUBLISHERS ARE QUICK to point out that most of the tens of thousands of editors they work with are honest and professional. But they also say they are under siege. A spokesperson for Elsevier said every week its editors are offered cash in return for accepting manuscripts. Sabina Alam, director of publishing ethics and integrity at Taylor & Francis, said bribery attempts have also been directed at journal editors there and are "a very real area of concern."

Jean-François Nierengarten of the University of Strasbourg, co-chair of the editorial board of *Chemistry–A European Journal*, published by Wiley, was targeted in June 2023. He received an email from someone claiming to be working with "young scholars" in China and offering to pay him $3000 for each paper he helped publish in his journal.

But Xiaotian Chen, a librarian at Bradley University who has studied paper mills in China, says publishers are not blameless. Chen points out that publishing houses have shown no sign of cutting back on the tens of thousands of special issues they put out every year in open-access journals—reportedly the preferred target for paper mills. Such issues generate hefty profits from the publication fees paid by authors. "Some of the for-profit publishers, they're just as greedy as a paper mill," Chen says. "And they count heavily on the contribution from Chinese authors to survive."

China is a major market for fake papers, and critics say measures to rein in paper mills there have been largely ineffectual. According to a new preprint, more than half of Chinese medical residents say they have engaged in research misconduct such as buying papers or fabricating results. One reason is that publications, though no longer always a strict requirement for career advancement, are still the easiest path to promotion in a range of professions, including doctors, nurses, and teachers at vocational schools, according to sources in China. Yet these groups may have neither the time nor the training to do serious research, Chen says. In such a setting, paying a few hundred or even thousand dollars to see one's name in print may seem a worthwhile investment, he says.

The towering demand for academic articles is not unique to China. In Russia and several ex-Soviet countries, for example, policies focused on publication metrics, coupled with a culture of corruption and the transition to market economy, have contributed to a similar situation, according to Abalkina. Research output is also gaining importance in India as universities there strive to climb rankings and junior doctors and scientists vie for prestigious jobs at home and abroad. Some universities even require undergraduates to publish papers as part of their curricula, a trend academics say is spreading.

"Students are really desperate to get research papers in whichever way possible," Zadey says. "No one really cares about the outcomes," he adds. "It's all about outputs."

Although publishers have ramped up their efforts against fraud, including establishing a hub for information sharing, critics say it's too little and too late. "They were too naïve, the real editors, the real people running these journals," says Elisabeth Bik, a microbiologist who spends her time scanning scientific papers for signs of fraud. At a meeting for journal editors she attended last year, "people were saying, 'Yeah, we've been asleep at the wheel,'" Bik recalls. "And now we need to sort of deal with that damage."

Zadey agrees with the need to tackle paper mills, but he worries about the implications for global research inequities. "There is going to be a whole lot of added scrutiny for people with my face and my name when we try to publish."

IN JULY 2023, Wise reported his findings about Olive Academic to several major publishers. Most promised to investigate and said they would circle back to him once they knew more or if they needed further information. So far, he hasn't heard back. "Whilst these investigations do certainly take time, I am a bit disheartened, if not surprised," he says.

Editors trying to safeguard their journals can also get discouraged. When Jer-Shing Huang of the Leibniz Institute of Photonic Technology in Germany joined Elsevier's journal *Optik* as editor-in-chief 1 year ago, his hope was to help junior scientists, particularly those in the Global South, improve their manuscripts. Instead, Huang says he ended up trying "to clean up the mess."

"These are organized crime rings that are committing large-scale fraud."

Matt Hodgkinson,
UK Research Integrity Office

It turned out that *Optik*, which was delisted from Web of Science in 2023, had a massive paper mill problem. Olive Academic was among its attackers. With Elsevier's blessing, Huang says, he started "rejecting a lot of really bad papers every day," as well as proposals for special issues. He also introduced policies requiring supervision of guest editors of special issues, which he said had been major drivers of the journal's growth. And he set about combing through hundreds of suspect papers that had already appeared.

Before he went on vacation last summer, Huang says, he had retracted more than 20 papers. But it was grueling work, and he had no idea how many more papers were left to check. "I'm really killed by this," he says.

Last fall, Huang told Elsevier he would resign as editor-in-chief. Not only was he spending his time fighting fires instead of doing science, he had also been attacked on the online forum PubPeer in what he believed was an act of revenge by paper mills rattled by his efforts. The publisher eventually convinced him to stay, but Huang remains conflicted. "This is not at all what I had imagined." ■

Frederik Joelving, based in Denmark, is an editor for Retraction Watch. This story was produced in collaboration with Retraction Watch.

ILLUSTRATION: DAVIDE BONAZZI

ASTRONOMY

Mystery in the "mass gap"

The identity of a compact astronomical object in the Milky Way is unclear

By **Maya Fishbach**

Astrophysical compact objects come in two varieties: neutron stars and black holes. Created when massive stars die and their cores collapse, neutron stars consist of protons and neutrons, whereas black holes are too heavy and dense to be described by anything other than gravity. These entities have uncertain mass limits. The neutron star maximum mass is between 2.2 and 2.5 solar masses (the Sun's mass as a unit of measure) (*1*). Black holes of less than 5 solar masses rarely have been observed. These limits suggest a "mass gap" between the most massive neutron stars and least massive black holes (*2*, *3*). On page 275 of this issue, Barr *et al.* (*4*) report the discovery of a compact object of ~2.35 solar masses that sits at the lower edge of this mass gap. It could be either the most massive neutron star or the least massive black hole ever observed.

Barr *et al.* identified this mass gap compact object as a partner to a millisecond pulsar known as PSR J0514–4002E. The pulsar (a rapidly rotating neutron star) was identified by a survey using the MeerKAT radio telescope array in South Africa. PSR J0514–4002E was located within a globular cluster (which contains hundreds of thousands of stars). By monitoring PSR J0514–4002E and accounting for relativistic effects, Barr *et al.* measured the total mass of the binary system (the pulsar and its companion) to be 3.887 solar masses. They then inferred the companion's mass to be around 2.35 solar masses, or between 2.09 and 2.71 solar masses at 95% credibility (see the figure). Because the companion is too faint to be a living star and too heavy to be a white dwarf (the dense core that remains after a less massive star dies), Barr *et al.* concluded that it is either a neutron star or a black hole.

The compact object in question is more massive than any known pulsar, which are all less than 2.1 solar masses (*5*). Its mass, between 2.09 and 2.71 solar masses, most resembles some compact objects observed in association with gravitational waves. For example, the gravitational wave merger of two neutron stars, GW170817, created an object of 2.46 solar masses (*6*). Another gravitational-wave event, GW190814, involved an object of 2.6 solar masses (*7*). These compact objects all sit in the murky region of mass space where, because of uncertainty in the maximum neutron star mass, they cannot unambiguously be classified as neutron stars or black holes (*8*).

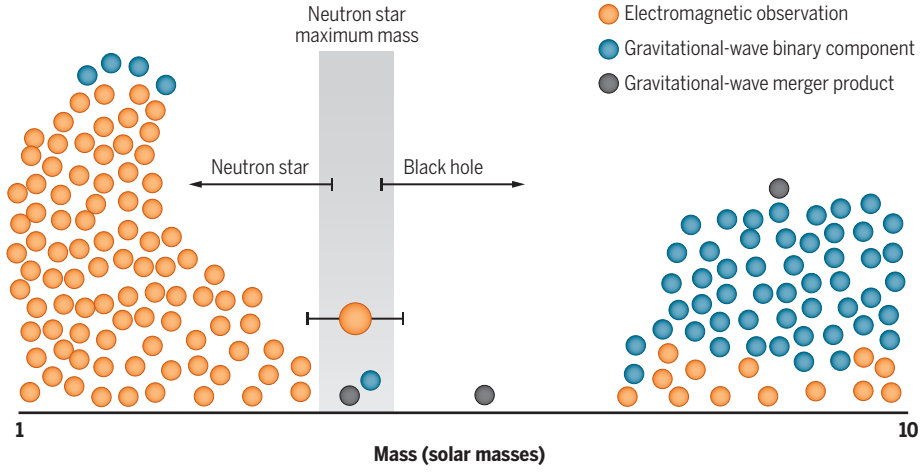
Before gravitational wave observations, it was unknown whether any compact objects exist with masses between around 2 and 5 solar masses. Even among gravitational wave events, there appears to be a dearth of binary systems with component masses in this mass range, with the 2.6–solar mass component of GW190814 serving as an exception (*9*). The origin of this putative mass gap is theoretically uncertain. Proposed explanations include the supernova explosion mechanism (*10*), which may prevent stars from producing black hole remnants in the mass gap. As well, stellar binary interactions may prevent black holes in the mass gap from participating in binary mergers (*11*). Observational selection effects also may prevent the observation of objects in the mass gap (*12*). Even if stars do not collapse into black holes in the mass gap, such objects are produced in the merger of two neutron stars, as observed with the gravitational-wave events GW170817 and GW190425 (*6*, *13*).

In addition to the unusual companion mass of pulsar PSR J0514–4002E, the binary system's total mass of 3.887 solar masses is remarkable. It is heavier than any known binary neutron star system, including Galactic double–neutron star systems (the heaviest of which is only 2.8887 solar masses) and neutron star mergers (the heaviest product of which is around 3.4 solar masses) that produce gravitational waves (*13*, *14*). Additionally, the binary is eccentric (with an orbital eccentricity of 0.71), and the pulsar is spinning rapidly (spin period of 5.6 ms).

This binary system identified by Barr *et al.* is likely the result of its extreme formation environment in a dense globular cluster. Stars at the center of a globular cluster constantly tug on each other gravitation-

Masses of neutron stars and black holes

Neutron stars and stellar-mass black holes have been observed in electromagnetic events (pulsars, x-ray binaries, and isolated black holes), as well as with certain gravitational-wave events. The gray band represents the uncertain mass limit separating neutron stars and black holes. The companion to pulsar PSR J0514–4002E is represented by the larger circle, with its 95% uncertainty interval.



CREDITS: (GRAPHIC) K. HOLOSKI/*SCIENCE*; (DATA) W. M. FARR, K. CHATZIOANNOU, *RES. NOTES AAS* **4**, 65 (2020); M. FISHBACH, V. KALOGERA, *ASTROPHYS. J. LETT.* **929**, L26 (2022); K. BADRY *ET AL.*, *MON. NOT. R. ASTRON. SOC.* **521**, 4323 (2023); R. ABBOTT *ET AL.*, *PHYS. REV. X* **13**, 011048 (2023)

Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, ON, Canada.
Email: fishbach@cita.utoronto.ca

ally, breaking up weaker binaries and assembling new, tighter binaries. Globular clusters can therefore give rise to exciting and puzzling stellar populations, including millisecond pulsars, gravitational-wave events, and fast radio bursts (transient radio pulses as fast as a fraction of a millisecond, whose sources remain unclear).

Barr *et al.* point out that the eccentric orbit, fast pulsar spin, and high total mass of the PSR J0514–4002E binary system suggest that the pulsar and its companion did not start off in a stellar binary. Instead, the pulsar exchanged its initial binary companion for the present, higher-mass object in a dynamical encounter. It is even possible that the companion in question is itself the product of a previous binary system inside the globular cluster—perhaps the merger of two neutron stars, even though mergers involving neutron stars are thought to be rare inside globular clusters (*15*).

Regardless of its origin—whether a stellar remnant, the merger product of two neutron stars, or the result of another type of merger or mass transfer event—the discovery of a compact object with a mass between 2.09 and 2.71 solar masses in a globular cluster has fascinating implications. If a neutron star, it is probably the heaviest one known to date, with lessons for the uncertain physics of extremely dense nuclear matter. If a black hole, it may be the lightest known, which could affect the understanding of supernova explosions or dynamical interactions such as neutron star mergers inside globular clusters. With upcoming electromagnetic and gravitational-wave observations, the growing population of compact objects between 2 and 5 solar masses will allow further resolution of the mass distribution and reveal the details of their formation. ■

REFERENCES AND NOTES

1. I. Legred, K. Chatziioannou, R. Essick, S. Han, P. Landry, *Phys. Rev. D* **104**, 063003 (2021).
2. F. Özel, D. Psaltis, R. Narayan, J. E. McClintock, *Astrophys. J.* **725**, 1918 (2010).
3. M. Fishbach, R. Essick, D. E. Holz, *Astrophys. J. Lett.* **899**, L8 (2020).
4. E. D. Barr *et al.*, *Science* **383**, 275 (2024).
5. W. M. Farr, K. Chatziioannou, *Res. Notes AAS* **4**, 65 (2020).
6. B. P. Abbott *et al.*, *Phys. Rev. Lett.* **119**, 161101 (2017).
7. R. Abbott *et al.*, *Astrophys. J. Lett.* **896**, L44 (2020).
8. R. Essick, P. Landry, *Astrophys. J.* **904**, 80 (2020).
9. A. M. Farah, M. Fishbach, R. Essick, D. E. Holz, S. Galaudage, *Astrophys. J.* **931**, 108 (2022).
10. C. L. Fryer, A. Olejak, K. Belczynski, *Astrophys. J.* **931**, 94 (2022).
11. L. A. C. van Son *et al.*, *Astrophys. J.* **940**, 184 (2022).
12. J. C. Siegel *et al.*, *Astrophys. J.* **954**, 212 (2023).
13. B. P. Abbott *et al.*, *Astrophys. J. Lett.* **892**, L3 (2020).
14. R. D. Ferdman *et al.*, *Nature* **583**, 211 (2020).
15. C. S. Ye *et al.*, *Astrophys. J. Lett.* **888**, L10 (2020).

10.1126/science.adn1869

MEDICINE

Surpassing sensitivity limits in liquid biopsy

Attenuation of cell-free DNA clearance in vivo is an alternative strategy to maximize recovery

By **Tina Moser**[1,2] *and* **Ellen Heitzer**[1,2]

Cell-free DNA (cfDNA) can be found in the blood and contains tumor-derived fragments [circulating tumor DNA (ctDNA)] when extracted from cancer patients. In advanced-stage tumors, ctDNA analyses to detect drug resistance mechanisms or actionable targets have already been translated into clinical practice (*1*). In recent years, ctDNA approaches have continuously been expanded to earlier disease stages. However, a persistent challenge of cancer screening and the molecular detection of minimal residual disease (MRD) has been the need to increase sensitivity, allowing for the detection of minute amounts of ctDNA with unprecedented accuracy. Most attempts to enhance sensitivity have focused on ex vivo strategies, such as sampling, analytical parts (e.g., library preparation), and bioinformatics. On page 274 of this issue, Martin-Alonso *et al.* (*2*) report confronting this challenge by transiently modulating the natural clearance mechanisms of cfDNA in vivo to increase its concentration in blood samples.

The amounts of cfDNA (and ctDNA) in the blood are determined by the interplay of its release—mainly driven by cell turnover—and its degradation and clearance, which are mediated through nuclease digestion, renal excretion, and uptake by macrophages of the mononuclear-phagocyte system (MPS) in the liver (*3*). Given the rapid clearance of cfDNA, with an estimated half-life of 30 to 120 min, a standard blood draw of 10 ml typically yields limited quantities of cfDNA. From 10 ml of blood, ~5 ml of plasma can be isolated, providing an average of 10 ng of cfDNA per milliliter, corresponding to roughly 15,000 haploid genome equivalents (GEs)—i.e., the amount of DNA in one copy of a genome (*4*). On average, ctDNA may make up as much as 10% (sometimes much higher) of the overall cfDNA pool in patients with advanced-stage cancer. However, the ctDNA fractions drop substantially to 0.1 to 1% in locally advanced disease and to <0.1% in early-stage disease or after curative-intent treatment. Therefore, when ctDNA fractions are as low as 0.1 or 0.01%, this translates to only 15 or 1.5 GEs, respectively, derived from the tumor. Such a blood sample may not contain sufficient ctDNA fragments for effective sequencing or detection.

Increasing sample volumes through plasmapheresis is one option to recover greater numbers of GEs and overcome sampling biases (*5*). However, this requires expensive instrumentation, is time-consuming, and might not be feasible for critically ill patients. To capture ctDNA more efficiently, the concept of proximal sampling, which refers to taking samples and body fluids more proximally to the tumor, has been introduced (*6*). Promising results were reported for the analysis of urinary cfDNA in bladder cancer (*7*). More recently, in vivo interventions, such as focused ultrasound (*8*) or radiation (*9*), have been proposed to increase ctDNA shedding temporarily. Yet, both proximal sampling and enhanced ctDNA release require a priori knowledge of tumor location, making the approach unamenable for screening approaches to detect cancers early.

Martin-Alonso *et al.* present a new proof-of-concept study using a preclinical model including healthy mice and mice bearing bilateral grafts of colorectal carcinoma cells to maximize the amount of ctDNA recovery by using intravenous priming agents that transiently delay cfDNA clearance in vivo (see the figure). In tumor-bearing mice, they demonstrated that liposomal nanoparticles mirroring the size of native cfDNA—that is mostly bound to histone proteins and circulates as mononucleosomes (~11 nm in size)—can attenuate phagocytic clearance by competing with uptake by the MPS. To further enhance the abundance of cfDNA and ctDNA, the authors used monoclonal antibodies (mAbs) with abrogated FcγR binding to protect cfDNA from enzymatic digestion by circulating deoxyribonucleases (DNases). Administered 1 to 2 hours before a blood draw, both agents (the liposomes and DNA binding mAbs) enabled a >10-fold increased recovery of cfDNA and ctDNA molecules. Application

of the liposomes improved the sensitivity of their ctDNA test up to 60-fold, particularly in mice with the lowest tumor burden. Although this approach enabled an improved recovery of tumor genomes and may substantially enhance the sensitivity of ctDNA diagnostic assays, it remains to be determined how such strategies would translate into humans in terms of formulation testing and tolerability.

Currently, the peak of sensitivity and specificity in ctDNA diagnostics is achieved using a bespoke approach, meaning that known mutations from the tumor tissue are tracked in plasma (*10*). However, when screening for a limited number of mutations, the number of available GEs dictates the analytical sensitivity. One effective strategy to capture rare ctDNA fragments is to sequence relevant target regions to exhaustion (up to 15,000 times). To this end, an increase of signal-to-noise ratios to reliably distinguish true tumor-associated mutations is necessary. A seminal breakthrough in reducing the impact of sequencing errors and increasing confidence in mutation detection was the use of molecular barcodes combined with sophisticated bioinformatics analyses that enable error suppression (*11*). For instance, in duplex sequencing, the two strands of DNA are tagged and sequenced independently, resulting in a 1000-fold error reduction compared with standard sequencing, but at a considerable expense (*12*). To reduce the required number of sequencing reads per target, the MAESTRO technique was developed, which uses short probes to enrich for patient-specific mutant alleles and uncovers the same mutant duplexes using up to 100-fold fewer reads (*13*). Enrichment of specific cfDNA fragment subpopulations during library preparation and the selective recovery of specific fragment sizes in silico might also boost ctDNA detection and could complement or provide an alternative to deep sequencing cfDNA (*14*).

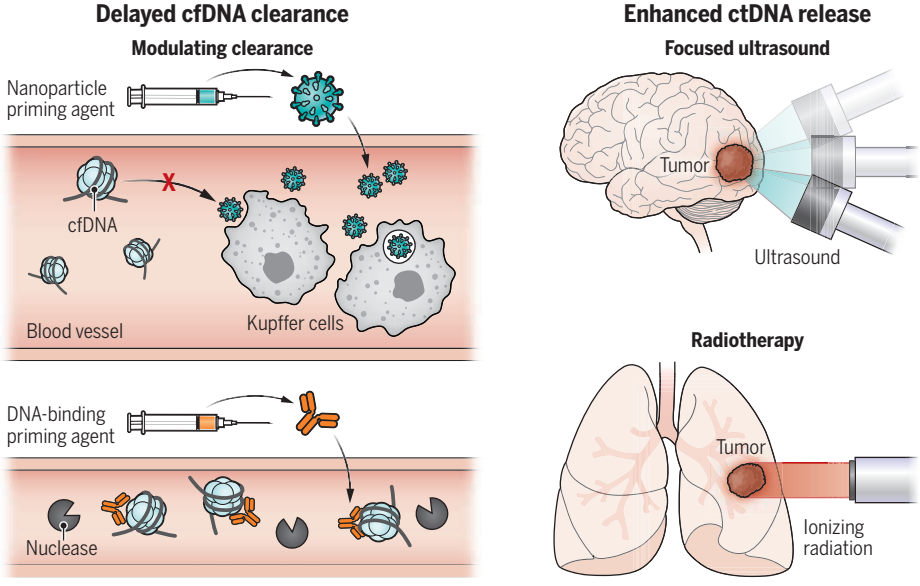
Another strategy to overcome the limited availability of GEs is to broaden the scope of mutation tracking for MRD. Instead of focusing on a single or a few mutations at high depth, this approach involves screening for a much more extensive array of mutations—potentially in the hundreds or thousands—which may greatly increase the odds of detecting ctDNA, even with limited amounts of GEs (*15*). Some groups showed that if multiple markers are simultaneously evaluated from the same plasma sample, the overall lower limit of detection of ctDNA is inversely correlated with the number of markers.

For the widespread use of ctDNA to detect early-stage cancer, more generalizable, tissue-agnostic ctDNA assays are required. These assays should offer broader patient coverage, faster turnaround times, and potential cost-effectiveness because they are not limited by tissue type. In this context, the integration of biological cfDNA features extending beyond mutations, such as fragment profiles, nucleosome features, and methylation patterns, harbor complementary information and may thereby improve the sensitivity boundaries (*3*). Regardless of which strategies (or a combination thereof) will prevail, increasing the sensitivity of ctDNA analyses remains pivotal for improving the clinical utility of liquid biopsies in cancer diagnostics and monitoring. Looking ahead, integrating multianalyte data—including RNA; exosomes; and circulating tumor cells, proteins, and metabolites—and combining different molecular features—e.g., genomics and epigenetics—holds promise for further improving sensitivity and specificity in liquid biopsy.

In addition to the requirement of amplifying sensitivity, other challenges persist, including the need for standardization of liquid biopsy protocols, addressing interlaboratory variability, and navigating ethical and regulatory considerations (e.g., handling false negative and false positive results and cost coverage). Additionally, evaluation of the cost-effectiveness of high-sensitivity technologies warrants attention to enable broad accessibility in clinical settings. Ongoing research and collaboration among academia, industry, and regulatory bodies are essential in refining existing methods, overcoming these challenges, and unlocking the full potential of liquid biopsy to revolutionize cancer diagnostics and treatment. The journey to redefine liquid biopsy continues, and with each breakthrough, a future where early cancer detection is a reality comes closer. ■

Increasing liquid biopsy sensitivity in vivo

In vivo approaches to increase the amount of cell-free DNA (cfDNA) that can be extracted from the blood in a liquid biopsy can increase the sensitivity of detection. For example, cfDNA clearance can be delayed when mice are treated with nanoparticles to compete with native cfDNA (usually in the form of mononucleosomes) for uptake by phagocytes. DNA binding monoclonal antibodies can also protect cfDNA in mice from degradation by nucleases. Other strategies include increasing circulating tumor DNA (ctDNA) release from tumors through ultrasound or ionizing radiation.



GRAPHIC: K. HOLOSKI/*SCIENCE*

[1]Institute of Human Genetics, Diagnostic and Research Center for Molecular BioMedicine, Medical University of Graz, Graz, Austria. [2]Christian Doppler Laboratory for Liquid Biopsies for Early Detection of Cancer, Medical University of Graz, Graz, Austria. Email: ellen.heitzer@medunigraz.at

REFERENCES AND NOTES

1. J. Pascual *et al.*, *Ann. Oncol.* **33**, 750 (2022).
2. C. Martin-Alonso *et al.*, *Science* **383**, eadf2341 (2024).
3. T. Moser, S. Kühberger, I. Lazzeri, G. Vlachos, E. Heitzer, *Trends Genet.* **39**, 285 (2023).
4. E. Heitzer, I. S. Haque, C. E. S. Roberts, M. R. Speicher, *Nat. Rev. Genet.* **20**, 71 (2019).
5. S. Weber *et al.*, *Cancers* **12**, 1588 (2020).
6. A. Tivey, M. Church, D. Rothwell, C. Dive, N. Cook, *Nat. Rev. Clin. Oncol.* **19**, 600 (2022).
7. P. S. Chauhan *et al.*, *npj Precis. Oncol.* **7**, 6 (2023).
8. L. Zhu, A. Nazeri, C. P. Pacia, Y. Yue, H. Chen, *PLOS ONE* **15**, e0234182 (2020).
9. J. M. Noh *et al.*, *Cancer Res. Treat.* **54**, 40 (2022).
10. N. P. Semenkovich *et al.*, *J. Immunother. Cancer* **11**, e006284 (2023).
11. A. M. Newman *et al.*, *Nat. Biotechnol.* **34**, 547 (2016).
12. S. R. Kennedy *et al.*, *Nat. Protoc.* **9**, 2586 (2014).
13. G. Gydush *et al.*, *Nat. Biomed. Eng.* **6**, 257 (2022).
14. F. Mouliere *et al.*, *Sci. Transl. Med.* **10**, eaat4921 (2018).
15. A. Zviran *et al.*, *Nat. Med.* **26**, 1114 (2020).

10.1126/science.adn1886

MEDICINE

Immune damage in Long Covid

Links between the complement and coagulation systems could lead to Long Covid therapies

By **Wolfram Ruf**[1,2]

Acute infections with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) cause a respiratory illness that can be associated with systemic immune cell activation and inflammation, widespread multiorgan dysfunction, and thrombosis. Not everyone fully recovers from COVID-19, leading to Long Covid, the treatment of which is a major unmet clinical need (*1*). Long Covid can affect people of all ages, follows severe as well as mild disease, and involves multiple organs. The persistence of lingering symptoms after acute disease creates a considerable challenge for understanding the specific pathophysiology and risk factors underlying Long Covid. On page 273 of this issue, Cervia-Hasler *et al.* (*2*) report a multicenter, longitudinal study of 113 patients who either fully recovered from COVID-19 or developed Long Covid, identifying localized activation of the innate immune defense complement system as a likely culprit that induces thromboinflammation and prevents the restoration of fitness after acute COVID-19.

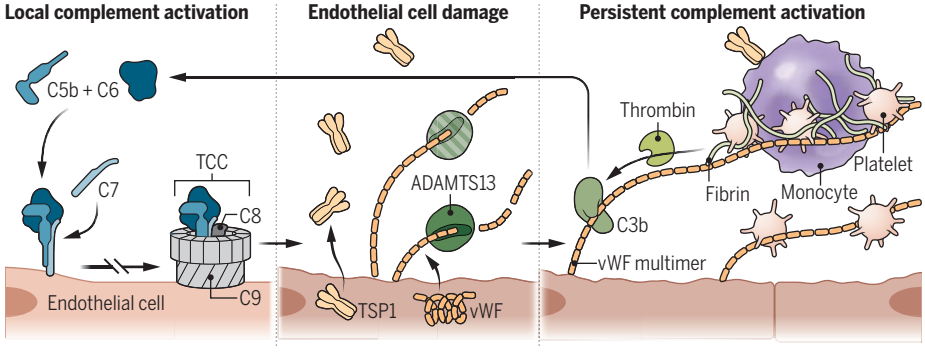
Patients with Long Covid display signs of immune dysfunction and exhaustion (*1*), persistent immune cell activation (*3*), and autoimmune antibody production (*1*), which are also pathological features of acute COVID-19. Cervia-Hasler *et al.* undertook a proteomic screen measuring serum levels of 6596 human proteins that are recognized by 7289 epitope-specific DNA oligonucleotide aptamer probes. The patients with severe or mild acute COVID-19 were analyzed during the acute infection and 6 months later. Comparison of the 40 Long Covid patients, 73 recovered patients, and 39 healthy controls revealed that most serum biomarkers that were elevated in patients with Long Covid at 6 months overlapped with those that were altered in the subgroup of the cohort with severe acute COVID-19.

In particular, the blood antimicrobial defense systems of complement and pentraxin 3 stood out as being significantly associated with Long Covid development. Complement components and pentraxins, which include serum amyloid proteins, serve humoral immunity by destroying and opsonizing pathogens for rapid clearance by innate immune cells. These proteins are up-regulated in hepatocytes as part of the inflammation-induced systemic acute phase response (*4*). However, the up-regulated markers of this response in Long Covid—pentraxin 3 and certain complement factors—are produced mainly by immune and other tissue-resident cells and not the liver (*5*), indicating that persistent inflammation in Long Covid patients is local rather than systemic.

[1]Center for Thrombosis and Hemostasis, Johannes-Gutenberg-University Medical Center Mainz, Mainz, Germany. [2]Scripps Research, La Jolla, CA, USA. Email: ruf@uni-mainz.de

Complement-coagulation cross-talk at the endothelial interface

During acute COVID-19, the complement and coagulation systems can become activated, and remain locally activated in various tissues in Long Covid patients. Endothelial cell damage by membrane insertion of the complement C5b-C7 complex results in vWF and TSP1 release. This is highly prothrombotic because the resulting vWF multimers recruit platelets and promote thrombin generation, and TSP1 promotes monocyte interactions with platelets. Additionally, reduced ADAMTS13 in Long Covid patients promotes the accumulation of ultralarge vWF multimers, which induces C3b binding and stimulates the alternative pathway of complement activation. Together, these feedback loops can maintain local complement activation and inflammation, as well as generating microclots that could underlie some of the features of Long Covid.



ADAMTS13, a disintegrin and metalloproteinase with thrombospondin motifs 13; TCC, terminal complement complex; TSP1, thrombospondin 1; vWF, von Willebrand factor.

The complement system is crucial for innate immune defense by effecting lytic destruction of invading microorganisms, but when uncontrolled, it causes cell and vascular damage. The complement cascade is activated by antigen–antibody complexes in the classical pathways or in the lectin pathway by multimeric proteins (lectins) that recognize specific carbohydrate structures, which are also found on the SARS-CoV-2 spike protein that facilitates host cell entry. Both pathways may contribute to the pronounced complement activation in acute COVID-19 (*6*). Independent of these pathogen recognition–triggered mechanisms, the so-called alternative pathway, in which C3b recruits complement factor B, leading to its proteolysis, can amplify complement activation on cell surfaces. The consecutive proteolytic cleavages in the complement cascade ultimately generate C5b that binds C6 and C7. C7 serves as the cell membrane anchor for the C5b-C7 complex and enables subsequent recruitment of C8 and C9 in the terminal complement complex (TCC), which mediates lysis of pathogens and host cells.

Focusing on Long Covid–specific proteome changes and taking into consideration age, sex, and hospitalization, Cervia-Hasler *et al.* detected increased C5b-C6 levels, supporting excessive complement activation. However, an aptamer probe measuring C7 showed surprisingly decreased levels. Careful validation of the aptamer target specificity revealed recognition of C7 in complex with other TCC components, but not of free C7. The reduced levels of circulating C7-containing complexes indicated that assembly of C7 with C5b-C6 resulted in increased membrane insertion and consequently cell damage in Long Covid. Consistently, a decreased ratio of complexed C7/C5b-C6 was a strong predictor for developing Long Covid (see the figure).

Cervia-Hasler *et al.* found that Long Covid patients also showed a marked up-regulation of serum von Willebrand factor (vWF) and thrombospondin 1, which are both released from damaged or activated endothelial cells as well as from platelets. Conversely, levels of a disintegrin and metalloproteinase with thrombospondin motif 13 (ADAMTS13), which cleaves vWF, were reduced in Long

Covid patients, indicating an imbalanced regulation of vWF in the circulation. ADAMTS13 plays a crucial role in controlling ultralarge vWF multimers that are released from activated endothelial cells and are highly prothrombotic by facilitating platelet adhesion. Moreover, vWF multimers bind C3b and thereby promote alternative complement pathway activation, which was also evident in the Long Covid patients based on measurement of activated factor B fragments. By contrast, processing of ultralarge vWF by ADAMTS13 increases the susceptibility of C3b to degradation and thereby attenuates local complement amplification (*7*). In addition to inefficient vWF degradation, Cervia-Hasler *et al.* found evidence for red blood cell lysis and release of heme, which may further contribute to the persistent local complement activation in Long Covid patients.

Coagulation activation leads to blood clots consisting of platelets and fibrin, which is lysed by fibrinolytic proteases. Acute, severe COVID-19 is associated with elevated levels of the fibrin degradation product D-dimer and by multiorgan thrombosis typically without reduction in platelet counts. Increased D-dimer and fibrin levels in acute COVID-19 predict the development of cognitive dysfunction ("brain fog") in Long Covid patients (*8*). D-dimer and fibrin elevations did not persist in Long Covid patients, but—apart from vWF—Cervia-Hasler *et al.* found that levels of coagulation factor 11 (F11) were elevated as a potential additional procoagulant mechanism. F11 serves in a coagulation amplification loop that generates thrombin on platelets that are recruited to endothelial vWF, and F11 can promote vascular inflammation without causing overt thrombosis (*9*). The lingering thromboinflammatory dysregulation in Long Covid occurs with apparently otherwise normal endothelial function. In particular, the biomarker profiles in Long Covid patients provided no evidence for an impaired endothelium-protective thrombomodulin pathway that not only regulates fibrinolysis and complement activation but also supports the degradation of inflammation-inducing anaphylatoxins C3a and C5a (*10*), which were unchanged in Long Covid patients. Furthermore, ongoing limited coagulation in Long Covid is indicated by the detection of circulating fibrinolysis-resistant microclots that are enriched in vWF, C7, acute-phase serum amyloid proteins, and fibrin (*11*), proteins identified in the proteomics screen by Cervia-Hasler *et al.* as contributors to Long Covid.

The formation of neutrophil extracellular traps (NETs), prothrombotic DNA "webs" that are ejected from neutrophils to immobilize pathogens, and other signs of immunothrombosis are frequent in acute COVID-19 but absent in Long Covid. Instead, platelets were found by Cervia-Hasler *et al.* to be associated with classical monocytes in the blood of patients with persistent Long Covid symptoms, in line with the demonstrated increased levels of thrombospondin, which stimulates platelet interactions with monocytes in addition to other platelet-derived proteins. Although single-cell sequencing of a small subset of Long Covid patients did not detect procoagulant or proinflammatory gene induction in classical monocytes, these showed down-regulation of nuclear receptor subfamily 4 group A member 1 (NR4A1) expression, the key transcription factor required for monocyte conversion into endothelial-protective, patrolling monocytes (*12*). In addition, the transcriptome of $CD16^+$ patrolling monocytes showed an upregulation in the expression of interferon-regulated genes, consistent with previous studies implicating persistent interferon signaling in Long Covid (*3*). Because monocytes can rapidly change their procoagulant properties in the context of complement activation without transcriptional induction of the coagulation initiator tissue factor, contributions of altered monocyte surveillance to endothelial dysfunction and local generation of thrombin, which also activates C3, should be considered in future studies as additional mechanisms for Long Covid–associated thromboinflammation.

"...the connections between viral reactivation, persistent interferon signaling, and autoimmune pathologies promise to yield new insights..."

Long Covid symptoms include a postexertional exhaustion reminiscent of other postviral illnesses, such as myalgic encephalomyelitis–chronic fatigue syndrome (ME-CFS) with suspected latent viral reactivation (*1*). Antibody titer changes in Long Covid patients indicate an association of fatigue with reactivation of latent Epstein-Barr virus (EBV) infections (*13*), and Cervia-Hasler *et al.* found that the severity of Long Covid symptoms is associated with cytomegalovirus (CMV) reactivation. In addition, various autoantibody alterations have been observed in Long Covid with unclear contributions to pathophysiology (*1*). Cervia-Hasler *et al.* found that autoantibodies to the B cell–stimulating C-X-C motif chemokine 13 (CXCL13) were decreased in patients with persistent Long Covid symptoms, which together with increased interferon signaling may contribute to persistent autoimmune pathologies. Intriguingly, complement-dependent prothrombotic antiphospholipid antibodies, which are observed in acute COVID-19, intersect with coagulation signaling to induce interferon responses in monocytes and promote autoimmunity by type I interferon production in dendritic cells (*14*). A better understanding of the connections between viral reactivation, persistent interferon signaling, and autoimmune pathologies promises to yield new insights into the thromboinflammation associated with Long Covid.

Although therapeutic interventions with coagulation and complement inhibitors in acute COVID-19 produced mixed results, the pathological features specific for Long Covid suggest potential interventions for clinical testing. Microclots are also observed in ME-CFS patients (*1*), indicating crucial interactions between complement, vWF, and coagulation-mediated fibrin formation in postviral syndromes. Target-specific coagulation inhibitors can reprogram innate immune phenotypes and interrupt the microangiopathic roles of vWF (*15*). Rebalancing complement activation can be achieved by blocking distinct amplification points with monoclonal antibodies that are clinically approved or in development. Complement and coagulation systems are not only an integral part of the innate immune response but also are connected at several levels in feedforward amplification loops. A better definition of these interactions in preclinical and clinical settings will be crucial for the translation of new therapeutic concepts in chronic thromboinflammatory diseases. ■

REFERENCES AND NOTES

1. H. E. Davis, L. McCorkell, J. M. Vogel, E. J. Topol, *Nat. Rev. Microbiol.* **21**, 133 (2023).
2. C. Cervia-Hasler *et al.*, *Science* **383**, eadg7942 (2024).
3. C. Phetsouphanh *et al.*, *Nat. Immunol.* **23**, 210 (2022).
4. A. Mantovani, C. Garlanda, *N. Engl. J. Med.* **388**, 439 (2023).
5. R. Lubbers, M. F. van Essen, C. van Kooten, L. A. Trouw, *Clin. Exp. Immunol.* **188**, 183 (2017).
6. M. Bosmann, *Sci. Immunol.* **6**, eabj1014 (2021).
7. S. Feng, X. Liang, M. H. Kroll, D. W. Chung, V. Afshar-Kharghan, *Blood* **125**, 1034 (2015).
8. M. Taquet *et al.*, *Nat. Med.* **29**, 2498 (2023).
9. S. Kossmann *et al.*, *Sci. Transl. Med.* **9**, eaah4923 (2017).
10. H. Loghmani, E. M. Conway, *Blood* **132**, 148 (2018).
11. E. Pretorius *et al.*, *Cardiovasc. Diabetol.* **20**, 172 (2021).
12. R. N. Hanna *et al.*, *Nat. Immunol.* **12**, 778 (2011).
13. M. J. Peluso *et al.*, *J. Clin. Invest.* **133**, e163669 (2023).
14. N. Müller-Calleja *et al.*, *Science* **371**, eabc0956 (2021).
15. P. Coppo, A. Cuker, J. N. George, *Res. Pract. Thromb. Haemost.* **3**, 26 (2018).

ACKNOWLEDGMENTS

W.R. receives grant support from the Deutsche Forschungsgemeinschaft (DFG 318346496/SFB1292/TP02). The author holds patents and consults for ARCA biopharma and Endpoint Health.

10.1126/science.adn1077

QUANTUM COMPUTING

Qubits without qubits

Light-based platforms for quantum computing do not require physical qubits

By **Olivier Pfister**

The realization of a quantum computer of practical use is fraught with challenges, including achieving performance that is fault tolerant (has robust error correction) and having a scalable architecture. Many physical platforms are competing to fit the bill, which makes for a vibrant field: quantum information (or qubit) processing by superconducting electronic circuits, by trapped ion or trapped atom systems, or by light fields. For the latter, quantum information is encoded in the amplitudes of light fields (qumodes), which can take a continuous range of values rather than just two distinct values as in qubits. Although qumode-based platforms for quantum information processing allow outstanding scalability compared with that of qubit platforms, they also make error correction much less intuitive than over qubits. On page 289 of this issue, Konno *et al.* (*1*) report the generation of a qubit state encoded in a qumode, providing an important step toward quantum computation with light.

Information handling by a classical computer follows binary rules over bits of information, which are either a 0 or 1. Qubits, on the other hand, can be both a 0 and 1 at the same time, which is called superposition, and superposition also extends to multiqubit states, which can lead to entanglement. Harnessing entanglement seems to be the key to unlocking the power of quantum computing. However, qubits are also eager to entangle with their intractably vast physical surroundings, and this leads to the washout of any quantum computation into a thermal bath. To prevent this, robust quantum error correction is required.

In 2001, Gottesman, Kitaev, and Preskill (GKP) posited that logical qubits could be encoded in physical qumodes—in particular, of light (*2*). Nine years later, it was proposed that an optical GKP state might be experimentally realized (*3*). Such qubit-encoding of information in qumodes (also known as bosonic codes) combines the two worlds of continuous and discrete variables, which have remained decidedly separated in classical information applications, largely because of the inability to correct for small drifts of continuous variables. Although this spelled the limits (and thus demise) of classical analog computing, the quantum situation is very different. GKP-based codes encode and error-correct qubits and, at the same time, correct against small drifts of boson (light) field amplitudes (typically, the position and momentum of a quantum harmonic oscillator). This has led to widespread interest in developing quantum computing platforms that feature bosonic fields alongside qubits.

There are two salient examples of quantum computing platforms that feature boson fields in addition to qubits. One involves the phonon field (mechanical quantum oscillator) that embodies the vibration of a trapped-ion qubit (*4*). The other platform is based on the microwave cavity field coupled to a superconducting qubit (of the transmon kind) (*5*). In both cases, GKP state generation in either a phonon or microwave photon field was enabled by a controlled gate that entangles the field to the qubit.

> **"...there is no fundamental reason that stands in the way of...a fully fledged photonic quantum computer."**

There is yet another way to generate bosonic qubit encodings that does not require a physical qubit and is therefore perfectly suited to "pure field" implementations of quantum information such as the quantum optics of propagating light fields. In such implementations, the physical qubits are replaced with physical qumodes—that is, the physical modes of the light defined by frequency, polarization, spatial profile, and propagation direction. In the majority of cases, physical qumodes are defined with exquisite precision by the resonant modes of an optical resonator that also contains a stimulated two-photon emitter. The whole assembly is called an optical parametric oscillator, as opposed to a laser that is based on single-photon stimulated emission. It has been shown theoretically (*6*) and experimentally (*7*–*11*) that one or two optical parametric oscillators can generate all the entanglement needed for quantum computing tasks. There are several reasons for this (*12*). Such oscillators emit a multitude of qumodes, either in the frequency or the temporal domain (or both). This forms the foundation for massively scalable quantum architectures. In addition, qumode entanglement can be engineered to take the form of cluster states that enable measurement-based quantum computing. And there exists a fault-tolerant threshold for qumode-based, measurement-based quantum computing if all qumodes are encoded as GKP qubits (*13*).

Konno *et al.* describe the first experimental generation of GKP states directly over optical qumodes. To achieve this, the authors achieved the comb structure of the GKP wave functions by leveraging the interference in qumode phase space that is present in exotic quantum states of light called "cat" states. In reference to Schrödinger's famous feline caught in a quantum superposition of "alive" and "dead," optical cat states can be viewed as a regular laser beam being both in and out of phase with itself at the same time. By classical physics, one could only understand this as destructive interference (no light present at all). Quantum mechanical light, however, can be placed in such a state and have nonzero intensity. Konno *et al.* generated cat states experimentally by subtracting photons one at a time from the light emitted by an optical parametric oscillator (*14*). They then created the interference of two beams of light in cat states with controlled relative phases and measured one interfering beam. This placed the other beam in a GKP state, which was precisely and unambiguously characterized by use of quantum state tomography.

Although Konno *et al.*'s experiment broke new ground, their GKP state does not yet have the quality required to reach fault tolerance. But there is no fundamental reason that stands in the way of experimental progress beyond this promising start toward a fully fledged photonic quantum computer. ■

Department of Physics, University of Virginia, Charlottesville, VA, USA. Email: olivier.pfister@gmail.com

REFERENCES AND NOTES

1. S. Konno *et al.*, *Science* **383**, 289 (2023).
2. D. Gottesman *et al.*, *Phys. Rev. A* **64**, 012310 (2001).
3. H. M. Vasconcelos *et al.*, *Opt. Lett.* **35**, 3261 (2010).
4. C. Flühmann *et al.*, *Nature* **566**, 513 (2019).
5. P. Campagne-Ibarcq *et al.*, *Nature* **584**, 368 (2020).
6. N. C. Menicucci *et al.*, *Phys. Rev. Lett.* **101**, 130501 (2008).
7. M. Chen *et al.*, *Phys. Rev. Lett.* **112**, 120505 (2014).
8. S. Yokoyama *et al.*, *Nat. Photonics* **7**, 982 (2013).
9. J.-I. Yoshikawa *et al.*, *APL Photonics* **1**, 060801 (2016).
10. W. Asavanant *et al.*, *Science* **366**, 373 (2019).
11. M. V. Larsen *et al.*, *Science* **366**, 369 (2019).
12. O. Pfister, *J. Phys. At. Mol. Opt. Phys.* **53**, 012001 (2020).
13. N. C. Menicucci, *Phys. Rev. Lett.* **112**, 120504 (2014).
14. A. Ourjoumtsev *et al.*, *Science* **312**, 83 (2006).

10.1126/science.adm9946

POLICY FORUM

AGRICULTURE AND ENVIRONMENT

Risks from solar-powered groundwater irrigation

Emissions reductions may not meet expectations, and groundwater use will likely increase

By **Soumya Balasubramanya[1], Dustin Garrick[2,3], Nicholas Brozović[4], Claudia Ringler[5], Esha Zaveri[6], Aude-Sophie Rodella[6], Marie-Charlotte Buisson[7], Petra Schmitter[7], Neha Durga[8,9], Avinash Kishore[10], Thai Thi Minh[11], Kashi Kafle[12], David Stifel[13], Sahana Balasubramanya[13,14], Ankit Chandra[4], Lesley Hope[15]**

Solar-powered groundwater irrigation is expanding exponentially in low- and middle-income countries (LMICs), creating opportunities and risks. In South Asia, more than 500,000 small stand-alone pumps have already been installed (see the figure). In Sub-Saharan Africa, solar pumps are gaining traction to expand food production and alleviate poverty. There is optimism about solar-powered irrigation helping LMICs meet their climate-change mitigation obligations, but insights from behavioral sciences, and early evidence, suggest that such emissions reductions are complex to calculate and likely lower than assumed. Groundwater pumping is likely to increase. Moving from siloed accounting of land, water, and energy use to integrated assessment frameworks can help manage unintended risks to land and water resources (*1*) and prevent lock-ins. By assessing social costs and benefits of solar-powered groundwater pumping, policy-makers can navigate trade-offs where irrigation expands food production and alleviates poverty but has unintended or unaccounted consequences for groundwater depletion and carbon emissions.

Emissions from agriculture, forestry, and other land uses accounted for 22% of carbon emissions globally in 2019 (*2*). Solar-powered groundwater irrigation allows for expanding energy use in agriculture, making it attractive to use in LMICs where poor farmers have growing energy needs.

There are three broad approaches to solar-powered irrigation in LMICs. The first is using a stand-alone pump powered by a solar panel, an option especially suited to areas not serviced by a reliable electric grid, and where irrigation is either uncommon or relies on diesel-powered pumps. Small stand-alone solar pumps are typically less than 10 hp (7.5 kW) and meant for individual use. Community stand-alone solar pumps deliver irrigation services to a community of farmers. A second approach consists of installing grid-connected (small) solar panels on the fields of farmers who use electric pumps, an option for agricultural areas where electric pumps are served by well-maintained grids. Typically, the solar panels installed on farms generate power and evacuate it into the same electric grid that powers the electric pumps. Energy use on the farm is "net-metered," with farmers charged or credited the difference between what the electric pump used and what the solar panel evacuated. The feasibility of this approach depends on the magnitude of administrative and engineering costs of installing small, distributed on-farm solar panels on scores of fields. The third approach consists of generating solar energy through large solar power projects and inputting it into the grid. Similar to any grid-connected energy generation project, these are typically implemented through

Solar panels power irrigation of a watermelon patch on the outskirts of Al-Hasakah, Syria, on 24 September 2023.

a contract and power-purchase agreement (that specifies the per unit price at which generated energy is purchased by the utility from the project developer). The feasibility of this approach depends on the ability of the developer to acquire land to install hundreds of large solar panels. Both grid-connected approaches also depend on the grid being able to receive varying mixes of renewable and nonrenewable energy.

Current and projected use of solar pumps in LMICs has mostly focused on small stand-alone ones, as these can be deployed irrespective of the coverage and condition of the grid. In South Asia, governments are providing sizable capital cost subsidies on pumps, and sometimes facilitating financing. For instance, India intends to install 2 million small stand-alone pumps and solarize 1.5 million existing electric pumps by 2026 (*3*). Bangladesh has installed 2829 solar pumps (*4*), mostly community stand-alone ones, while Nepal has installed around 2000 small stand-alone pumps. In Sub-Saharan Africa, governments have focused on providing an enabling environment, especially by reducing import and other taxes on small stand-alone pumps, while encouraging private vendors to sell and install them. Reporting of sales by vendors to industry initiatives (such as the Global Off-Grid Lighting Association) is voluntary, thus sales are likely underreported; consequently, there is considerable uncertainty around the number of small stand-alone pumps in use across different countries.

Early estimates of the impacts of stand-alone solar pumps show that farmers experience improvements in food production, food security, and livelihoods, necessary conditions for poverty reduction (*1*, *5*).

COMPLEX TO ASSESS

In India, carbon emissions associated with pumping groundwater for irrigation are estimated at 2 to 7% of annual national total emissions (*6*). India's experience with groundwater development for increasing agricultural productivity has provided a template for Sub-Saharan Africa, where emissions are likely to increase with groundwater development. A complete replacement of electric or diesel pumps with solar pumps would reduce greenhouse gas emissions. However, complete replacement is not guaranteed, and these benefits may come with risks, particularly for groundwater depletion. Some farmers may continue to use their former pumps along with the new solar pump if irrigation needs have previously been unmet and expanding energy use is profitable (*5*). Some farmers might opt to stack multiple energy sources—for example, combine using a diesel pump with a solar pump if the capacity of the solar pump prevents them from abstracting enough groundwater to meet their irrigation needs. Or, they may prefer to retain a device that works irrespective of weather conditions (storage batteries may be attached to solar pumps, but they add to equipment costs that are already substantial without government support). If farmers are irrigating multiple noncontiguous plots, then complete replacement of diesel with solar pumping is only possible if each plot has its own solar pump, an expensive proposition.

Additionally, land use and input choices might change. For example, farmers may change the cultivated area, cropping intensity, crop choice, and agrochemical use, especially if they increase net energy use (*5*). Other indirect changes such as fuel for transporting inputs and outputs may also be linked with the adoption of solar pumps. The probability of these occurrences and their intensity will vary by baseline energy use choices, cropping system, existing agricultural practices, and by farmer (*5*). However, such shifts in resource use will need to be considered, and this is complex to monitor and measure, especially in the case of stand-alone solar pumps. If all potential changes linked to the adoption of solar irrigation are considered, emissions reductions might well be lower than originally assumed.

Even if solar-powered irrigation leads to net-zero emissions, which our analysis calls into question, it may exacerbate groundwater depletion and threaten the viability of millions of wells already at risk of running dry (*6*, *7*). Solar irrigation is likely to increase groundwater use as its adoption expands, because it has zero pumping costs, meaning that cheap, clean energy can accelerate the groundwater boom in LMICs and higher-income settings alike (*1*, *5*, *7*). In the case of stand-alone pumps, there is limited incentive for farmers to reduce groundwater use, irrespective of whether they switched from diesel or electric pumps, and irrespective of whether irrigation needs were previously met or not (*5*), and existing evidence suggests the same (*1*, *5*). Increases in groundwater pumping may be a desirable outcome for improving food security and livelihoods if groundwater hydrogeology supports increased abstraction, such as in the case of local shallow aquifers that deplete and replenish seasonally.

In the case of installing grid-connected (small) solar panels on the fields of farmers using electric pumps—an option receiving considerable attention in South Asia—net metering has been proposed to potentially limit or reduce groundwater abstraction. It is argued that farmers would be able to earn money or credit from the utility if they generate more electricity than the pumps use (*8*). Estimates from scenario-based studies in India, with varying assumptions and based in different Indian states, have yielded differing conclusions on whether net metering could induce substantial reductions in groundwater use (*8*, *9*). Pilot studies from different Indian states examining similar financial incentives offered to farmers for reducing energy use have found modest to no reductions in pumping behaviors (*10*, *11*). Understanding the potential for net metering to reduce groundwater abstraction is important for scaling this approach in South Asia, because the technical logistics and financial costs are substantial, and the scope of utilities to offer farms with small solar panels large enough incentives to guarantee sufficient reductions in pumping is limited owing to the subsidized prices at which electricity is supplied to farmers (*8*).

Emerging evidence also suggests that ownership of solar pumps, both stand-alone and those with net metering, and whether supplied by government-coordinated programs or private vendors, usually accrues to economically and socially better-off farmers who can afford to access and purchase them (*1*). It is important to identify practical ways to extend the benefits of solar irrigation to those who cannot purchase pumps.

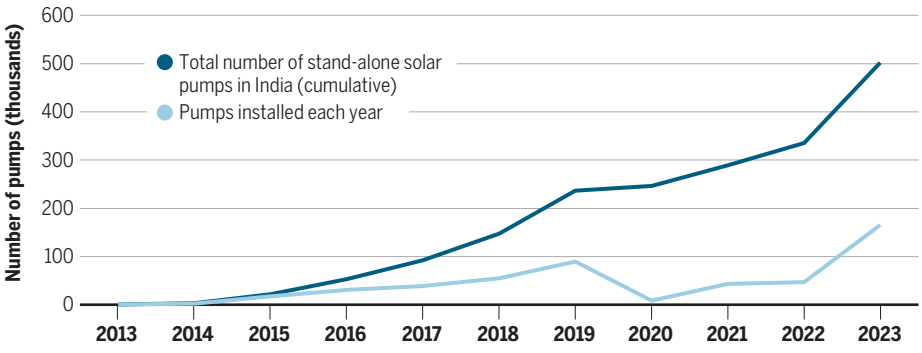
PROGRESS FOR POLICY

Linking solar irrigation to carbon credit programs will require integrated assessment and/or accounting frameworks for estimating emissions reductions. Such a framework would have guidelines to establish a reference baseline level of emissions; specify safeguards that have to be met, addressed, and reported periodically against (such as the risks of reversals and displacement of emissions); demonstrate additionality (reductions in emissions would not have occurred without the credit program in place); and implement a monitoring system that periodically reports on water–energy–land use changes. Such a framework may also need to take into

[1]Environment Global Practice, The World Bank, Washington, DC, USA. [2]Water Institute, University of Waterloo, Ontario, Canada. [3]University of Oxford, Oxford, UK. [4]Daugherty Water for Food Global Institute at the University of Nebraska, Lincoln, NE, USA. [5]International Food Policy Research Institute, Washington, DC, USA. [6]Water Practice, The World Bank, Washington, DC, USA. [7]International Water Management Institute, Colombo, Sri Lanka. [8]Indian Institute of Technology-Bombay, Maharashtra, India. [9]International Water Management Institute, Anand, Gujarat, India. [10]International Food Policy Research Institute, New Delhi, India. [11]International Water Management Institute, Accra, Ghana. [12]Texas A&M University, College Station, TX, USA. [13]Lafayette College, Easton, PA, USA. [14]State University of New York, Buffalo, NY, USA. [15]University of Energy and Natural Resources, Sunyani, Ghana. Email: sbalasubramanya@worldbank.org

Growth in stand-alone solar pumps in India

India plans to install 2 million small stand-alone pumps and solarize 1.5 million existing electric pumps by 2026 to improve livelihoods for the poor and boost incomes from agriculture.



account the "carbon footprint" of solar photovoltaic panels over their lifetime. There is an opportunity to learn lessons from the mixed experiences of REDD+ projects that aim to reduce emissions by creating carbon credits through improving forest extent and conditions. The complexities of implementing such programs highlight the need for a multiscale assessment of impacts along with realistic expectations regarding achievements on the ground (*12*).

Improving the scope and quality of information on water, energy, and land use changes is needed, with a shift from siloed to integrated approaches (*13*). On the biophysical side, remotely sensed data on actual evapotranspiration, irrigated area, and crop choice could introduce reporting of confidence levels to account for uncertainties associated with such data and thus support better accounting across scales. To ground truth data, especially in Africa where major knowledge gaps exist, ground-based observation networks are developing a multi-university consortium to assess rainfall-recharge relationships and improve public information on groundwater availability and variability (*14*). Because groundwater abstraction in LMICs is rarely metered, most estimates of changes are indirectly inferred, based on reported hours of pumping, liters of diesel use, and on the rare occasion of metered electricity use. Improvements in measuring abstraction directly would be important and also help better calibrate indirect measurements.

On the socioeconomic side, a practical approach may be to leverage and integrate existing data collection initiatives. For example, national statistics offices in LMICs could communicate planned data collection activities for the financial year in advance so that agriculture, energy, and water ministries have an opportunity to explore potential collaborations, by expanding the scope of established and periodic surveys to include prioritized indicators on water, land, and energy use, which would help a wide range of users. Remote and crowd-sourced approaches to collecting data could help decision-makers informally monitor emerging problems, especially when high accuracy is not needed.

Emerging technologies could solve some data challenges, especially where national statistics offices are capacity constrained. Solar pumps with inbuilt sensors could be deployed to measure water abstraction; if telemetry were enabled and data were shared, it would help governments monitor and manage groundwater resources over larger areas. A complementary effort to track regulatory interventions and governance mechanisms for setting and enforcing limits on emissions and pumping would build our understanding of the fit and effectiveness of different tools (*1*).

Insights from entrepreneurship and market design could improve policy responses for how irrigation using stand-alone solar pumps may reach the poorest (*15*). In recent years, there has been an emergence of small enterprises and foundations in Asia and Sub-Saharan Africa that sell on-demand services using portable solar or diesel pumps to communities who require irrigation but whose land holdings are too small to justify installing their own individual pump, and who cannot be serviced through the grid. Coordinated partnerships between government extension services, entrepreneurial initiatives, and companies that manufacture pumps would reduce transaction costs and help identify and scale appropriate services to less wealthy farmers that are informed by hydrological, agronomic and socioeconomic opportunities and constraints. In agricultural areas served by the electric grid (such as in South Asia), improving and upgrading infrastructure to allow grid solarization would help extend equity in access to energy over the longer term.

Regulation, planning, and enforcement capacity for energy and water need improvement in both LMICs and higher-income settings. The clean-energy boom will likely make groundwater problems worse regardless of income levels and regulatory capacity, in part because groundwater is so poorly managed globally (with some small exceptions) (*7*). The rapid adoption of solar irrigation makes these activities more challenging by giving a short time for governments and institutions to adapt. Cumulative limits for carbon emissions and groundwater depletion, nested at multiple levels from farms to districts and regions, need to be established, which increases regulatory and enforcement responsibilities and requires overcoming deeply entrenched interests benefiting from cheap electricity and free water. Public finance should be targeted where returns are likely to be high, including monitoring networks and targeted, time-limited subsidies to spur behavioral shifts for consumers (food choices) and producers (input choices). Advances in data, behavioral sciences, and entrepreneurship can equip policy-makers to set realistic expectations, equitably share benefits, and manage risks as solar irrigation expands. ■

REFERENCES AND NOTES

1. A. Closas, E. Rap, *Energy Policy* **104**, 33 (2017).
2. IPCC, "Summary for policymakers" in *Climate Change 2023: Synthesis Report Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* [Core Writing Team, H. Lee, J. Romero, Eds.] (IPCC, 2023), pp. 1.
3. Lighting Global/ESMAP, *IFC, Efficiency For Access Coalition, GOGLA, Open Capital Advisors, "Off-Grid Solar Market Trends Report 2022: State of the Sector"* (World Bank, 2022).
4. National Database of Renewable Energy, Sustainable and Renewable Energy Development Authority, "Statistics of Solar Irrigation Pumps"; https://ndre.sreda.gov.bd/index.php?id=15 (accessed 19 May 2023).
5. E. Gupta, *Energy Policy* **129**, 598 (2019).
6. V. Mishra, A. Asoka, K. Vatta, U. Lall, *Earths Futur.* **6**, 1672 (2018).
7. S. Jasechko, D. Perrone, *Science* **372**, 418 (2021).
8. N. Bassi, *Int. J. Water Resour. Dev.* **34**, 132 (2018).
9. T. Shah *et al.*, *Econ. Polit. Wkly.* **49**, 10 (2014).
10. R. Fishman, U. Lall, V. Modi, N. Parekh, *J. Assoc. Environ. Resour. Econ.* **3**, 819 (2016).
11. A. Mitra, S. Balasubramanya, R. Brouwer, *Am. J. Agric. Econ.* **105**, 861 (2022).
12. T. A. P. West *et al.*, *Science* **381**, 873 (2023).
13. A. S. Rodella *et al.*, *The Hidden Wealth of Nations: The Economics of Groundwater in Times of Climate Change* (World Bank, 2023).
14. M. O. Cuthbert *et al.*, *Nature* **572**, 230 (2019).
15. G. Falchetta *et al.*, *Environ. Res. Lett.* **18**, 094044 (2023).

ACKNOWLEDGMENTS

S.B. and D.G. contributed equally to this work. This work was funded by the World Bank's Global Water Security and Sanitation Partnership and Global Program on Sustainability, the CGIAR Initiative on NEXUS Gains, and the Swiss Agency for Development and Cooperation. The views expressed in this paper belong to the authors, and do not reflect the views of the World Bank, or of any of the institutions that any of the authors are affiliated with. All errors belong to the authors.

10.1126/science.adi9497

BOOKS *et al.*

A prescribed burn is ignited near Santa Cruz, California, in October 2023.

PHOTO: AFP VIA GETTY IMAGES

ANTHROPOCENE

Searching for meaning in a world aflame

A metaphor-filled memoir bears witness to our fraught existence in the age of wildfires

By **Stephen Pyne**

For Sonoma County, California, 2020 was a hard year. COVID-19 pushed people outdoors, smoke drove them indoors, and fires sent them fleeing to places of refuge. For Sonoma resident Manjula Martin, there was also a fiercely personal burden of pain from a botched operation. As she writes in her new book, *The Last Fire Season*, the outside world and the inside world combined to form a "set of concentric circles—a hurt body inside a thriving garden in the middle of a neglected forest, on an overheating planet surrounded by an unknowable universe."

The usual narratives do not apply to Martin's story. Well-worn conceits cannot contain the ceaseless, sometimes eruptive, chaos of the interpenetrating worlds. But fire might. She approaches her narrative from the perspective of one raised in a hippie-lite community outside Santa Cruz—think simple back-to-the-landers, not radical seekers of spiritual enlightenment—who distrusts traditional power structures. Instead, she seeks out the knowledge of Indigenes, small-scale practitioners, and women. Along the way, she encounters fire ("fire followed us wherever we went").

The Last Fire Season opens with a dry lightning storm that kindled 650 named fires that massed into complexes, one a million-acre gigafire. Tellingly, no date is given; this is a memoir, which holds to "remembered truth." The fires send Martin and her partner into serial evacuations, while the flames serve to illuminate the interior struggle she endures, described as an "underlying pulse of chronic pain."

It is a common rhetorical device to symbolically link external and internal worlds. But Martin goes further and makes the relationship reciprocal: The outside world of redwoods, megafires, smoke palls, social injustice, gardens, and climate change is internalized. "The forest exhaled and my body took it in," she writes, becoming a cipher for the insults and trials of the Anthropocene.

The book's core event centers on a broken IUD, a fragment of which gets absorbed into Martin's uterus and leads to a bungled hysterectomy. A fragment of the outside world has intruded, unwelcome and harmful, a harbinger of what is to come. The surgery leaves her with crippling pain that prevents her from engaging with the world as she would like; even tending a garden is difficult. Instead, she must bear witness.

Martin's prose is light, precise, wordy, and frequently figurative; sentences flow with the viscosity of air. I enjoyed particularly some of the sketches and vignettes that sprinkle the text on such topics as smoke, flowers, and Big Basin Redwoods State Park, which spring up from the pages like fire poppies. The overall organization replaces narrative and subplots with a daisy chain of topics that can double as symbols. The 4 months of fire season proper (August through November) offer a trellis onto which the many seeds she plants can loosely attach.

Martin offers no clear structure because she finds no clear resolution. Eventually, she experiences real fire personally and finds an awkward reconciliation. Learning to live with fire bears similarities with learning to live with pain. Both require ceaseless tending, both shuffle decks of messy metaphors, and both have no end point. This is not the last fire season.

Readers will learn little about fire behavior, fire institutions, fire policy, or fire history. There are anecdotes drawn from fire ecology, some delightful, but they come with scant explanation. When Martin asks a guide at Big Basin about a postburn botanical quirk, she is told opaquely that it is "fire-adapted," which to her seemed "like alchemy."

Readers may find themselves in the same state. Those who enjoy confessional memoirs, share Martin's worldview, and savor her sensibilities will likely relish the book. Those who don't, probably less so.

The Last Fire Season: A Personal and Pyronatural History
Manjula Martin
Pantheon, 2024. 352 pp.

Amitav Ghosh has written that the absence of climate change as an informing theme for contemporary literature bespeaks a crisis of cultural imagination (*1*). *The Last Fire Season* offers a reply to his challenge. It shows how a confused, compounding barrage of phenomena and experiences can be transmuted into personal meaning.

We have a lot of fire ahead of us—good, bad, surprising, mean, playful, existentially wretched. The fires will touch everything. To survive, we will need metaphors as well as models and artful memoirs as much as artificial intelligence–assisted simulations. ■

The reviewer is an emeritus professor at Arizona State University, Tempe, AZ, USA, and the author of *The Pyrocene: How We Created an Age of Fire, and What Happens Next* (Univ. of California Press, 2022) and *Fire: A Brief History* (Univ. of Washington Press, 2001). Email: stephen.pyne@asu.edu

REFERENCES AND NOTES

1. A. Ghosh, *The Great Derangement: Climate Change and the Unthinkable* (Univ. of Chicago Press, 2016).

10.1126/science.adm9750

PSYCHEDELICS

When the CIA met LSD

A pair of anthropologists' personal papers shed light on US psychedelic science in the mid-20th century

By **Alexis Turner**

Tripping on Utopia: Margaret Mead, the Cold War, and the Troubled Birth of Psychedelic Science
Benjamin Breen
Grand Central Publishing, 2024. 384 pp.

In 1945, Electronic Numerical Integrator and Computer (ENIAC)—the first programmable, general-purpose electronic digital computer—was put into operation at the US Army's Ballistic Research Laboratory. When ENIAC was announced to the public soon after, scientists of all stripes were infatuated with the machine, not just because of its intended use (calculating missile trajectories for improved accuracy) but also because of what they hoped it promised more broadly. If you could reprogram a computer, they wondered, could you also reprogram a human mind? The question was a seductive one in the aftermath of World War II, when the world was preoccupied with humanity's self-destructiveness. Not only did innovations such as ENIAC promise something better, they also allowed scientists to imagine themselves as the ones best positioned to deliver it.

This is the idealism at the heart of Benjamin Breen's new history, *Tripping on Utopia: Margaret Mead, the Cold War, and the Troubled Birth of Psychedelic Science*. While the scientific optimism of the mid-20th century easily attached itself to numerous technologies, Breen's book explores one of its more colorful manifestations: US research on the possible uses of lysergic acid diethylamide (LSD).

Well before hippies got their hands on the psychedelic compound, scientists were investigating the drug's potential to deprogram and reprogram minds in the hopes of creating a more stable world order. The task would prove easier said than done—and not just because optimism around the compound traveled easily across highly varied domains. Legally sanctioned LSD research took place in fields such as psychology, biology, anthropology, chemistry, and child development, and the compound was investigated at disparate institutions for very different purposes.

Despite this proliferation of research, much evidence of it has been lost, thanks in no small part to the Central Intelligence Agency (CIA), an institution that was heavily involved in early LSD experiments. Upon learning that its efforts would be subject to congressional oversight, the agency destroyed most of its documentation on the compound. Where, then, does the historian begin?

While *Tripping on Utopia*—a surefire crowd-pleaser that lets readers peep through a window on science meant to remain hidden—will be neither the first nor the last history to answer that question by turning to the CIA itself, Breen takes a novel approach. By plumbing the personal papers of anthropologists Margaret Mead and Gregory Bateson—researchers known to have worked with the agency on both cybernetics and propaganda, but the history of whose work with psychedelics has been largely neglected—Breen circumvents the missing evidence. The result is a methodologically clever book that sheds new light on the scientific history of LSD through an engaging, highly readable story that weaves a narrative far richer in detail than a typical academic history.

Margaret Mead's lesser-known work on LSD offers insight into secretive CIA projects.

The book begins with background on Mead's and Bateson's intellectual formation before moving into the work they did during World War II with the CIA's precursor agency, the Office of Strategic Services. Breen traces the slowly widening divide between the two researchers after the CIA's introduction to LSD in 1949, both personally and in their understanding of the role scientists ought to play in politics. Other scientists make appearances as well, from flamboyant acid guru Timothy Leary to the sober and serious Roy Grinker. By the end of the story, LSD has been classed as a Schedule I substance, putting it off-limits even to medical researchers and leaving its proponents searching in vain for a satisfactory answer to how things ended up so far from where they hoped they would be.

Although Breen never entirely answers this question, his choice to follow a small handful of actors—with occasional overemphasis on just how flawed they were—can leave readers tempted to imagine that LSD's downfall was a consequence of the actions of bad scientists from a bygone era. But the book's larger argument is that technical and political consensus within an open, large, profoundly complex system are questionable goals in and of themselves. I think he gets closest to the cause of LSD's demise in a throwaway point he makes but does not explore in great depth, in which he observes that the ambitious actors researching the drug intentionally undermined one another along the way, resulting in profoundly mixed messages about it competing on the public stage.

Ultimately, *Tripping on Utopia* is enough of a romp to be of widespread interest to casual readers, but it is also professionally timely, given the recent resurgence of medical interest in psychedelics. It will also be of interest to those who take seriously the difficulties of the public communication of science. ■

The reviewer is at the Department of the History of Science, Harvard University, Cambridge, MA 02138, USA.
Email: txturner@g.harvard.edu

10.1126/science.adm9748

African swine fever has devastated Borneo's bear pig population, previously a major source of food.

PHOTO: JURGEN FREUND

LETTERS

Edited by **Jennifer Sills**

Pig virus imperils food security in Borneo

African swine fever has devastated pig populations in Asia since 2018 (*1*). On the island of Borneo, which includes the country of Brunei, the Malaysian states of Sarawak and Sabah, and Indonesian Kalimantan, bearded pigs (*Sus barbatus*) were once the most numerous large mammal species (*2*), but African swine fever has led to population declines of 90 to 100% (*3*, *4*). The substantial drop may warrant a conservation status uplisting from Vulnerable to Critically Endangered (*5*). The loss of pigs disrupts food security and ecosystems and threatens other endangered wildlife.

In the past, bearded pigs constituted 81% of hunted wildlife weight in villages in East Kalimantan (*6*). The state of Sarawak alone once harvested a million bearded pigs per year (*2*), and Sabah's annual hunted pig weight was estimated at 8.6 million kg (*7*). Although the Muslim population of Borneo does not eat pork, the pig population collapse affects the livelihoods and cultural traditions of millions of non-Muslim people.

The decline of Borneo's pig population also poses unknown ecological effects. Wild pigs are important seed predators that play a substantial role as ecosystem engineers (*1*). In addition, in the absence of pigs, local people are likely to shift their focus to hunting endangered species, such as pig-tailed macaques (*Macaca nemestrina*) (*8*).

Although African swine fever has garnered substantial attention in countries with major pork industries (*9*), its effects in Borneo have been largely overlooked. Resistance to the disease in domestic pigs in southern Africa has been identified, but the basis for resistance remains unknown, especially in wild pigs (*10*). There is no evidence indicating that wild pig populations can fully recover in Borneo or on other islands in Southeast Asia where the disease has taken a toll, including Java, Sumatra, Timor-Leste, and the Philippines (*11*, *12*). Urgent research and interventions, with the participation of rural communities, should focus on preventing the spread of African swine fever to other regions where people depend on pigs, such as the island of New Guinea.

Erik Meijaard[1]*, Andi Erman[1], Marc Ancrenaz[1,2], Benoit Goossens[3]
[1]Borneo Futures, Jalan Sultan, Bandar Seri Begawan, BS8111, Brunei. [2]HUTAN–Kinabatangan Orangutan Conservation Programme, Kota Kinabalu, Malaysia. [3]Organisms and Environment Division, Cardiff School of Biosciences, Cardiff University, Cardiff CF10 3AX, UK.
*Corresponding author.
Email: emeijaard@borneofutures.org

REFERENCES AND NOTES

1. M. S. Luskin *et al.*, *Wildl. Lett.* **1**, 8 (2023).
2. J. Caldecott, "Hunting and wildlife management in Sarawak" (International Union for Conservation of Nature, Gland, Switzerland, 1988).
3. K. Chu, S. L. Wong, A. Chang, "Where are all the Sabah pigs?," *Macaranga* (2022); https://www.macaranga.org/where-are-all-the-sabah-pigs.
4. D. J. Kurz *et al.*, in *Wildlife Atlas of Sabah*, G. Davies, Ed. (WWF-Malaysia, Kota Kinabalu, 2022), pp. 123–132.
5. M. Luskin, A. Ke, E. Meijaard, M. Gumal, K. Kawanishi, "*Sus barbatus* (errata version published in 2018)" (IUCN Red List of Threatened Species, 2017); https://www.iucnredlist.org/ja/species/41772/123793370.
6. M. Schagen, "A review of hunting off-take rates in the Kelay Sub-District, East Kalimantan, Indonesia" (The Nature Conservancy—Indonesia Program, Samarinda, Indonesia, 2007).
7. A. Angelsen, "Economic Benefits of Wildlife in Sabah" (Sabah Wildlife Department, Kota Kinabalu, Malaysia, 2003).
8. N. Ruppert *et al.*, "*Macaca nemestrina* (errata version published in 2023)" (IUCN Red List of Threatened Species, 2022); https://www.iucnredlist.org/species/12555/223433999.
9. World Organization for Animal Health, "African swine fever" (2023); https://www.woah.org/en/disease/african-swine-fever/.
10. M. Penrith *et al.*, *Rev. Sci. Technol.* **23**, 965 (2004).
11. N. I. Dharmayanti *et al.*, *Transbound. Emerg. Dis.* **68**, 2890 (2021).
12. C.-H. Hsu *et al.*, *Pathogens* **12**, 1068 (2023).

10.1126/science.adn3857

Environmental risks and infertility in China

In 2015, the Chinese government terminated the "one-child policy," which had curtailed overpopulation by reducing births by approximately 400 million but led to an aging population, a shrinking labor pool, and a lopsided sex ratio (*1*). China's 2015 "two-child policy" (*1*) and 2021 "three-child policy" (*2*) have had some success in increasing birth rates (*3*), but the number of newborns remains surprisingly low (*4*). Birth rate declines are partly attributable to socioeconomic pressures, which cause some potential parents to opt for fewer children (*5*). However, the declines are also likely due to the prevalence of infertility (*6*), which has increased markedly in China in the past two decades (*1*, *2*). Infertility has been associated with environmental factors such as air quality (*5*, *7*, *8*). To increase birth rates, China needs to address the underlying issues driving the decline.

In 2022, about 46%, 39%, and 15% of newborns in China were the first, the second, and the third or beyond child of their parents, respectively (*3*). However, more second and third children did not mean more total births. China's birthrate dropped from 23.33 births per thousand people in 1987 (25.29 million newborns) to 6.77 births per thousand people in 2022 (9.56 million newborns) (*4*).

The infertility rate among the Chinese population of childbearing age increased from 11.9% in 2007 to 15.5% in 2010 (*1*). In 2020, 17.6% of the Chinese population of childbearing age—about 33 million couples—suffered from infertility (*2*). Environmental and climate risks may

contribute to impairment in reproductive health. For instance, air pollution is associated with sperm morphology during spermatogenesis (*7*), lowering semen quality (*5*, *8*). Heatwaves are linked to complications that lead to gestational diabetes and hypertensive disorders of pregnancy, which adversely affect the fetus (*9*). Endocrine-disrupting chemicals in drinking water and food are associated with decreases in sperm parameters, longer time to conceive, and compromised oocyte and embryo quality (*10*). In China, environmental factors have been correlated with damage to reproductive health for both men (*5*, *7*) and women (*8*). Environmental risks may also contribute to preterm birth (*8*, *9*), which has increased in China since 2012 (*11*).

To address China's falling birth rates, scientists need to establish the causality between declining reproductive health and each environmental risk. Researchers should also explore the underlying biological mechanisms; thoroughly understanding these interactions may become increasingly urgent given that, in mouse models, the progeny of mothers facing current risks during pregnancy show impaired fertility (*12*). In the short term, China needs to further promote assisted reproductive technology. Finally, China should mitigate environmental pollution and protect people of childbearing age from exposure to environmental risks.

Rui Feng[1,2] and Bin Chen[3]*
[1]College of Environmental and Resource Sciences, Zhejiang University, Hangzhou, China. [2]State Key Laboratory of Clean Energy Utilization, Zhejiang University, Hangzhou, China. [3]Assisted Reproduction Unit, Department of Obstetrics and Gynecology, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Key Laboratory of Reproductive Dysfunction Management of Zhejiang Province, Hangzhou, China.
*Corresponding author.
Email: hichenbin@zju.edu.cn

REFERENCES AND NOTES

1. J. Qiao *et al.*, *Lancet* **397**, 2497 (2021).
2. J. Qiao, *Chin J. Reprod. Contracept.* **43**, 7 (2023); https://doi.org/10.3760/cma.j.cn101441-20220906-00382 [in Chinese].
3. National Health Commission of the People's Republic of China, "Statistical bulletin on the development of health care in China in 2022" (2023), p/ 26; http://www.nhc.gov.cn/guihuaxxs/s3585u/202309/6707c48f2a2b420fbfb739c393fcca92/files/8a3994e41d944f589d914c589a702592.pdf [in Chinese].
4. National Bureau of Statistics of China, *China Statistical Yearbook 2023* (2023), Sections 2.1 and 2.2, pp. 31–32; https://www.stats.gov.cn/sj/ndsj/2023/indexch.htm [in Chinese].
5. N. Skakkebæk *et al.*, *Nat. Rev. Endocrinol.* **18**, 139 (2022).
6. World Health Organization (WHO), "Infertility prevalence estimates, 1990–2021" (2023); https://www.who.int/publications/i/item/978920068315.
7. Y. Guo *et al.*, *Sci. Tot. Environ.* **885**, 163532 (2023).
8. Q. Zhang *et al.*, *Innovation* **3**, 100312 (2022).
9. WHO, "Protecting maternal, newborn, and child health from the impacts of climate change: Call for action" (2023); https://www.who.int/publications/i/item/9789240085350.
10. M. P. Green *et al.*, *Environ. Res.* **194**, 110694. https://doi.org/10.1016/j.envres.2020.110694
11. K. Deng *et al.*, *Lancet Glob. Health* **9**, e1226 (2021).
12. W. Song *et al.*, *Sci. Tot. Environ.* **859**, 160431 (2023).

10.1126/science.adn3214

To reflect China's updated policies, the two children on the right were recently added to the original family of three in Guo Xue's "A beautiful future" scupture in Wuhan.

India's coal mining plan undermines climate goals

Carbon neutrality depends on the reduction of fossil fuel usage and the development of renewable energy (*1*, *2*). At the 2021 United Nations Climate Change Conference (COP26), India signed an agreement to reduce coal and enhance its renewable energy capacity to 500 GW by 2030, with the goal of achieving net-zero carbon emissions by 2070 (*3*, *4*). As the world's third-largest producer of coal and fourth-largest greenhouse gas emitter, India was expected to play an important role in global mitigation efforts (*5*). However, in November 2023, the country approved a plan to increase coal production (*6*), undermining national and global environmental goals.

To satisfy rapidly growing energy demands (*6*), India plans to increase annual output from its underground coal mines from 26 million tons in 2023 to 100 million tons by 2030 (*7*). The extensive development of underground coal mining will have profoundly negative and potentially irreversible effects on the India's ecosystems, including increased surface subsidence, soil erosion, land fissures, vegetation death, and threats to food security and infrastructure (*8*). Underground coal mining disrupts the flow of underground aquifers, harming water resources (*9*). The process brings substantial quantities of waste soil and rocks to the surface, which release toxic substances upon contact with air and water and contaminate air, soil, and aquatic ecosystems (*10*). Coal mining also poses the risk of gas explosions, resulting in casualties and exacerbating the greenhouse effect (*11*).

The Global Decarbonization Accelerator, adopted at the 2023 United Nations Climate Change Conference (COP28), calls for the acceleration of decarbonization in the energy industry (*12*). To protect the environment and foster a healthier planet, the Indian government should abandon its underground coal mining plan. India must uphold its COP26 commitments, actively formulate emission reduction plans, and set emission targets and actions at both national and local levels.

Shanshan Yang[1,2], Yifan Ruan[3], Jun Sun[1,2], Qiying Yang[1]*
[1]Institute for Advanced Marine Research, China University of Geosciences, Guangzhou 511462, China. [2]College of Marine Science and Technology, China University of Geosciences, Wuhan 430074, China. [3]School of Marxism, China University of Geosciences, Wuhan 430074, China.
*Corresponding author.
Email: yangqiying@cug.edu.cn

REFERENCES AND NOTES

1. G. Lin, Y. Zhao, J. Fu, D. Jiang, *Science* **380**, 699 (2022).
2. X. Zhao *et al.*, *Resourc. Conserv. Recycl.* **176**, 105959 (2022).
3. K. Singh *et al.*, *Biomass Bioenerg.* **177**, 106944 (2023).
4. N. Singh, *Energies* **15**, 5852 (2022).
5. A. Gupta *et al.*, *Energ. Procedia* **160**, 848 (2019).
6. B. Parkin, J. Singh, "India plans to triple underground coal mining to meet energy demand," *Financial Times* (2023).
7. R. Baruah, S. Narayan, "Govt weighs relaxations for underground coal miners" *Mint* (2023).
8. K. Zhang *et al.*, *Ecosyst. Health Sustain.* **9**, 0122 (2023).
9. D. Bedoya-Gonzalez *et al.*, *Hydrogeol. J.* **30**, 1717 (2022).
10. S. Kalisz *et al.*, *J. Environ. Manag.* **304**, 114239 (2022).
11. F. Li *et al.*, *Process Safe. Environ. Protect.* **167**, 274 (2022).
12. COP28 UAE, "COP28 Presidency launches landmark initiatives accelerating the energy transition" (2023); https://www.cop28.com/en/news/2023/12/COP28-Presidency-launches-landmark-initiatives-accelerating-the-energy-transition.

10.1126/science.adn3642

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

The 2023 earthquake in eastern Tu caused considerable surface deformation and damage to infrastructure.

PHOTO: JIANNAN MENG

NATURAL HAZARDS

Supersheer and surface faults

The 2023 earthquake sequence in Türkiye and Syria caused a great deal of damage and loss of life. Meng *et al.* captured surface features with a comprehensive set of field and drone observations before these features could be eroded or destroyed. The surface deformation caused from the underlying faults was complex but provides insight into fault displacement, rupture propagation, and slip transfer between fault systems. These observations should help with future seismic hazards and with interpreting events from the historic record. Ren et al. used an array of geophysical observations to determine how rupture propagated along the complex series of faults. The rupture speed was faster than the shear wave velocity at times, a condition called supershear. The authors pieced together some of the underlying physics behind this complex earthquake, helping us better understand the rupture process and subsequent ground shaking. —BG

Science p. 298, 10.1126/science.adj3770, p. 305, 10.1126/science.adi1519

WATER CHEMISTRY

Competitive chemistry between water ions

The mutual neutralization of the hydronium cation and hydroxide anion to form neutral water molecules is one of the most basic chemical processes. It therefore has attracted considerable interest, but direct experimental probing of the underlying reaction mechanisms has been lacking. By realizing the interaction in merged beams of two ionic species with near-zero relative velocity, Bogot *et al*. directly visualized the neutral products of these reactions and reported three different product channels. Two channels were attributed to a predominant electron-transfer mechanism, and a smaller channel was associated with proton transfer. The two-beam collision experiment is an important step toward understanding the quantum dynamics of this fundamental reaction. —YS

Science p. 285, 10.1126/science.adk1950

BIOGEOGRAPHY

Peoples' use of plants

A diverse array of plants provides the base materials for human sustenance and livelihoods. Although some of these plant species have become ubiquitous, others are endemic to certain locations, where they have specific cultural uses. Pironon *et al*. studied where plants used by people occur and how their hotspots of diversity and endemism overlap with broader patterns of plant diversity and the locations of protected areas. They found that utilized species diversity is correlated with overall plant diversity, suggesting that prioritizing biodiversity hotspots for conservation could protect many of these species. However, current protected areas are inadequate for conserving these species and their associated biocultural diversity. —BEL

Science p. 293, 10.1126/science.adg8028

ELECTROPHOTOCATALYSIS

Radicals from silver

Introducing trifluoromethyl groups to aryl rings is an important step in the preparation of many pharmaceuticals and agrochemicals, but the chemistry often requires specialized, expensive reagents. Campbell *et al*. now report

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

CORONAVIRUS

Blood alterations in Long Covid

Some individuals can endure persistent, debilitating symptoms for many months after an initial severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. However, the factors underpinning these health issues, called Long Covid, are poorly understood. Comparing the blood of patients with confirmed SARS-CoV-2 infection with that of uninfected controls, Cervia-Hasler *et al*. found that patients experiencing Long COVID exhibited changes to blood serum proteins indicating activation of the immune system's complement cascade, altered coagulation, and tissue injury (see the Perspective by Ruf). At the cellular level, Long Covid was linked to aggregates comprising monocytes and platelets. These findings provide a resource of potential biomarkers for diagnosis and may inform directions for treatments. —SHR

Science p. 273, 10.1126/science.adg7942; see also p. 262, 10.1126/science.adn1077

LIQUID BIOPSY

DNA hide-and-seek

Liquid biopsy for tumor analysis offers the potential for noninvasive access using a blood draw instead of a surgical procedure. In addition, sampling the blood can detect tumor DNA even when the location of a tumor is unknown. However, circulating tumor DNA is usually scarce, and it can be difficult to collect enough blood for adequate detection, especially in cases where tumors are small. Martin-Alonso *et al*. addressed this difficulty by developing two different types of priming agents that protect circulating DNA from destruction (see the Perspective by Moser and Heitzer). With these agents, more DNA remains in the bloodstream and is easier to detect even in small blood volumes. In mouse models of cancer, both approaches greatly increased the sensitivity of liquid biopsies. —YN

Science p. 274, 10.1126/science.adf2341; see also p. 260, 10.1126/science.adn1886

RADIO ASTRONOMY

A compact object in the mass gap

There is a substantial gap between the masses of the heaviest measured neutron star and the lightest measured black hole. This mass gap carries information about how both types of object form during supernovae (see the Perspective by Fishbach). Barr *et al*. used pulsar timing to investigate a binary system in a dense star cluster. They found that the pulsar's binary companion is an object in the mass gap, but could not determine whether it is an unusually high-mass neutron star or an unusually low-mass black hole. Either way, the authors argue that it formed by the earlier merger of two other neutron stars. —KTS

Science p. 275, 10.1126/science.adg3005; see also p. 259, 10.1126/science.adn1869

QUANTUM OPTICS

Generating optical grid states

Quantum computers under development are at the intermediate size scale and are already demonstrating quantum advantage over classical systems for specific tasks. Going to larger scale systems is challenging for some solid-state platforms. Optics provides a possible route, but optical systems require engineered photonic states to mitigate for loss and to run error correction codes. Konno *et al*. demonstrate the generation of Gottesman-Kitaev-Preskill states, or grid states, in which the wavefunction resembles a sharp-peaked two-dimensional array (see the Perspective by Pfister). These states have been predicted to be fault tolerant, allowing quantum error correction codes to be readily implemented. Realizing such engineered photonic states will be important in the development of large-scale optical quantum computers. —ISO

Science p. 289, 10.1126/science.adk7560; see also p. 264, 10.1126/science.adm9946

NANOMATERIALS

Boosting gold cluster emission with copper

Gold nanoclusters have potential applications as near-infrared emissive materials for biological applications but often exhibit low photoluminescence quantum yield (PLQY) in solution at room temperature. Shi *et al*. found that copper substitution to form an $Au_{16}Cu_6$ cluster improved the PLQY to more than 60% in oxygenated solutions at room temperature. The presence of copper led to ultrafast intersystem crossing to the long-lived triplet state and also suppressed nonradiative decay. —PDS

Science p. 326, 10.1126/science.adk6628

CATALYSIS

How silicon imparts stability

Copper–zinc oxide catalysts with elements such as silicon as structural promoters have long been used for methanol synthesis from carbon dioxide and hydrogen. Barrow *et al*. examined how silicon doping also results in almost doubling the lifetime of industrially used catalysts that also incorporate other metal oxides such as alumina. Their studies of model catalysts and of aged industrial catalysts show that silicon helps to slow particle growth (sintering), not only of copper, but more importantly of zinc oxide. Silicon's incorporation into the zinc oxide lattice improves its hydrothermal stability under reaction conditions. —PDS

Sci. Adv. (2024) 10.1126/sciadv.adk2081

THYMUS

How to promote thymus regeneration

The thymus undergoes transient but reversible involution in response to stress stimuli, but the molecular mechanisms mediating thymic regeneration after involution are not well defined. Nevo *et al*. performed single-cell RNA sequencing on the thymic non–T cell compartment after dexamethasone-induced acute thymic involution in mice. This treatment promoted a type 2 immune response driven by thymic-resident innate lymphoid cells type-2 (ILC2s), which were activated in response to tissue damage by cytokines produced by thymic tuft cells and fibroblasts. ILC2s were required for tissue regeneration by producing effector molecules that promoted medullary thymic epithelial cell differentiation. Together, these findings identify a thymic tuft cell–fibroblast–ILC2 axis required for thymic regeneration after acute thymic involution. —HMI

Sci. Immunol. (2024) 10.1126/sciimmunol.abq6930

HIV

Targeting the fusion peptide

The fusion peptide (FP) on the HIV-1 envelope protein is a conserved site that can be targeted by neutralizing antibodies. However, the extent of protection conferred by FP-directed antibodies in the context of mucosal infection is unclear. Pegu *et al*. tested three different anti-FP antibodies, including one human antibody and two rhesus

macaque antibodies, for their ability to neutralize HIV-1 in vitro and confer protection against simian-human immunodeficiency virus (SHIV) challenge of rhesus macaques in vivo. All three antibodies conferred protection against mucosal challenge with SHIV at clinically meaningful titers, supporting the further development of both anti-FP antibody therapies and vaccines designed to elicit anti-FP humoral responses. —CM

Sci. Transl. Med. (2024)
10.1126/scitranslmed.adh9039

that photoexcitation of a silver complex can release reactive trifluoromethyl and higher perfluoroalkyl radicals directly from more readily available carboxylates such as trifluoroacetate. The silver can then be reoxidized electrochemically to render the arene perfluoroalkylation reaction catalytic. The method also highlights silver's prospective role as a more general highly oxidizing photoredox catalyst. —JSY

Science p. 279, 10.1126/science.adk4919

DNA-COLLOID CRYSTALS

More tilings with short flexible linkers

Although nanocrystals with well-defined polyhedral shapes can be assembled into ordered lattices with long DNA linkers, the large linkers often do not preserve the packing shape. Zhou *et al*. have shown that placing a flexible spacer between the rigid DNA anchor strand and the single-stranded DNA sticky end decouples the binding of sticky ends from the nanocrystal shape. This approach allows imperfect polyhedral building blocks to assemble into highly ordered space-filling structures and for non–space-filling nanocrystals to co-pack as complementary shapes. —PDS

Science p. 312, 10.1126/science.adj1021

IMMUNOLOGY

HLA allotype pairs influence HIV

Highly variable human leukocyte antigen (HLA) molecules bind and present a diverse range of antigenic peptides to T lymphocytes, thereby eliciting an immune response. Maternally and paternally derived HLA molecules may bind very similarly to radically different sets of peptides. Viard *et al*. developed a method that estimates the difference in types of peptides bound between the pairs of HLA variants, which they termed "functional divergence." It is proposed that individuals with high functional divergence present a broader range of peptides and elicit stronger immune responses compared with people having low functional divergence. Accordingly, HIV-infected individuals with the highest functional divergence develop less severe disease relative to those with the lowest functional divergence. A similar effect might be observed for other diseases in which HLA plays a central role. —PNK

Science p. 319, 10.1126/science.adk0777

CHEMISTRY

Copper-coupling guide

Reactions that form carbon–nitrogen bonds are central to the synthesis of pharmaceuticals and agrochemicals. Although the general contours of the catalytic methods in current use have been known for decades, it remains surprisingly hard to predict ahead of time precisely which conditions will prove optimal to couple a pair of complex substrates. Samha *et al.* report the development and validation of a model to help optimize ligand selection in Ullman-type copper-catalyzed carbon–nitrogen coupling. The use of such models could save substantial time as chemists seek to access target compounds. —JSY

Sci. Adv. (2024) 10.1126/sciadv.adn3478

BIOCHEMISTRY

Detecting substrate and ubiquitination

Ubiquitination is a posttranslational modification that regulates protein stability and protein-protein interactions in many signaling pathways. Goncharov *et al*. generated bispecific antibodies that specifically detect the proinflammatory kinases RIP1 and RIP2 and their precise ubiquitin linkages. The antibodies revealed an increase in the ubiquitination of RIP2 in samples from individuals with inflammatory bowel disease, highlighting the potential utility of these antibodies in characterizing signaling pathway activity in disease contexts. —JFF

Sci. Signal. (2024) 10.1126/scisignal.abn1101

IN OTHER JOURNALS

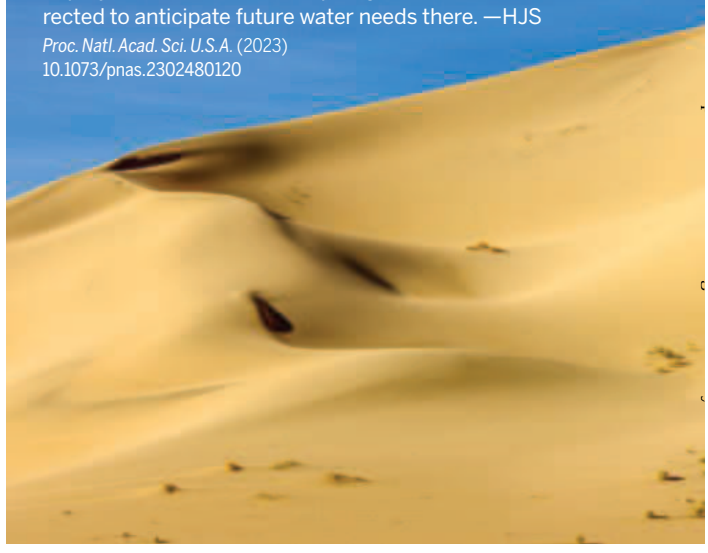
Edited by **Caroline Ash** *and* **Jesse Smith**

HYDROCLIMATE

Drier than expected

Both theoretical and model analyses predict that atmospheric water vapor concentrations will increase as the air warms. However, Simpson *et al.* report that observations in arid and semi-arid regions of the Southwestern US reveal that near-surface water vapor contents have not increased there over the last four decades, and that models fail to simulate that behavior. Given the importance of understanding hydroclimate in dry regions and the fact that models are the primary tools used to project climate, this discrepancy must be understood and corrected to anticipate future water needs there. —HJS

Proc. Natl. Acad. Sci. U.S.A. (2023) 10.1073/pnas.2302480120



SIGNAL TRANSDUCTION

JAKs suppress lung inflammation

Protein kinases of the Janus family (JAKs) are key components of inflammatory signaling by cytokines. Therefore, JAK inhibitors are being used and developed to treat diseases such as dermatitis and asthma. Such treatments are directed at cytokine signaling in hematopoietic cells, but Tamari *et al.* have discovered anti-inflammatory activities of JAKs in the lungs, where they suppress inflammation through signaling in sensory neurons. JAK signaling promotes synthesis of the neuropeptide calcitonin gene-related peptide β (CGRPβ) within neurons that innervate the lungs. Better understanding of such tissue-specific signaling may allow more effective targeting of JAKs to avoid conflicting effects that might limit the effectiveness of current experimental therapies. —LBR

Cell (2024) 10.1016/j.cell.2023.11.027

NEUROSCIENCE

Lost and now found

Named after Donald Hebb, a Canadian psychologist, Hebbian potentiation occurs when a stimulus is repeated, synapse and target cell–specific connections grow stronger, and an action becomes intuitive and "learned." Characterization of the plasticity of the underlying cellular and molecular mechanisms requires the identification of synaptic inputs and target cells. Grigoryan *et al.* investigated the receptors, ion channels, and molecular cascades contributing to long-term potentiation in the hippocampus of mice while learning

The atmosphere above the Southwestern US drylands is drier than predicted.

with their antibacterial properties. —MSL

Macromolecules (2023)
10.1021/acs.macromol.3c02026

SOCIAL CHEMOSIGNALING

Is there power in a woman's tears?

Many animals communicate through social chemosignaling, such that molecules in urine, milk, tears, or feces change behaviors among others within their species. For example, the production of tears among female rats has been linked to sexual or aggressive behaviors among male rats. The effect of chemical signals from human emotional tears has been poorly understood. Agron et al. harvested tears from human women donors and created odorless saline solutions that looked identical to tears. When subjects were exposed to human tears from women, rather than the saline controls, the authors observed reduced aggressive behavior among men and reduced activity levels in brain networks linked to aggression. —EEU

PLOS Biol. (2023)
10.1371/journal.pbio.3002442

the location of an object. The authors found evidence of retrograde signaling that required signaling cascades activated by metabotropic glutamate receptor 1α (mGluR1α)–coupled G proteins and a postsynaptic calcium burst. —PRS

Proc. Natl. Acad. Sci. U.S.A. (2023)
10.1073/pnas.2312752120

PSYCHIATRY

Social microbes

Social anxiety disorder, the intense feeling of being judged or rejected in social situations, is one of the most common mental disorders in teenagers. Sufferers experience severe anxiety or panic attacks in social situations. Because recent data have shown that an altered gut microbiome is associated with some psychiatric disorders, Ritz *et al.* investigated whether gut bacteria could play a role in anxiety. The authors transplanted gut microbiota from social anxiety disorder patients and healthy controls in germ-free mice and performed behavioral and gene expression analysis. Animals transplanted with patient microbiota showed mostly normal behavior, except for an increase in social fear associated with changes in peripheral immunity and the neuropeptide oxytocin in brain areas associated with fear and interpersonal relationships. —MMa

Proc. Natl. Acad. Sci. U.S.A. (2023)
10.1073/pnas.2308706120

HYDROGELS

A new route to hydrogels

Many routes to photopolymerizing monomers require either ultraviolet light to activate the added photoinitiators or the addition of a photosensitizer and co-initiator to use visible wavelengths. These processes may not be compatible with in vivo use because of the use of UV light or possible toxicity from the additives or their by-products. Cagnetta *et al.* evaluated a range of conjugated polymer nanoparticles that can operate across the full spectrum of visible light and can act as both photoinitiators and cross-linkers for the reactions and form a series of acrylic hydrogels showing high water uptake and swelling. Further, the authors found that the conjugated polymer nanoparticles were uniformly distributed within the gels and retained their photochemical activity, such as the formation of radical species under illumination, which endows the hydrogels

DNA SENSORS

A weak but reactive iron carbene

Single-walled carbon nanotubes (SWCNTs) can be wrapped with single-stranded DNA aptamers that bind neuromodulators such as dopamine. The subsequent increase in SWCNT near-infrared fluorescence can be used as an in vivo sensor. Rosenberg *et al.* studied how the wrapping of $(GT)_6$ and $(GT)_{15}$ aptamers changed upon dopamine binding in solution with x-ray scattering interference. They functionalized the oligomer ends with gold nanoparticles that acted as molecular rulers. Dopamine binding caused both axial elongation and radial constriction of single-stranded DNA on the SWCNT surface that in turn led to brighter fluorescence. —PDS

J. Am. Chem. Soc. (2024)
10.1021/jacs.3c09549

RESEARCH ARTICLE SUMMARY

LIQUID BIOPSY

Priming agents transiently reduce the clearance of cell-free DNA to improve liquid biopsies

Carmen Martin-Alonso†, Shervin Tabrizi†*, Kan Xiong†, Timothy Blewett, Sainetra Sridhar, Andjela Crnjac, Sahil Patel, Zhenyi An, Ahmet Bekdemir, Douglas Shea, Shih-Ting Wang, Sergio Rodriguez-Aponte, Christopher A. Naranjo, Justin Rhoades, Jesse D. Kirkpatrick, Heather E. Fleming, Ava P. Amini, Todd R. Golub, J. Christopher Love*, Sangeeta N. Bhatia*, Viktor A. Adalsteinsson*

INTRODUCTION: Liquid biopsies including the analysis of cell-free DNA (cfDNA) from blood can be used to diagnose, monitor, or molecularly profile disease. Despite the fast adoption of liquid biopsies in oncology, prenatal testing, infectious disease, and organ transplant monitoring, higher sensitivity is needed in many important clinical applications. In oncology, efforts to improve the sensitivity for detecting circulating tumor DNA (ctDNA) have mostly focused on ex vivo sequencing and analysis methods. However, an intrinsic challenge is the scarcity of ctDNA in vivo, which leaves little ctDNA to be collected and analyzed.

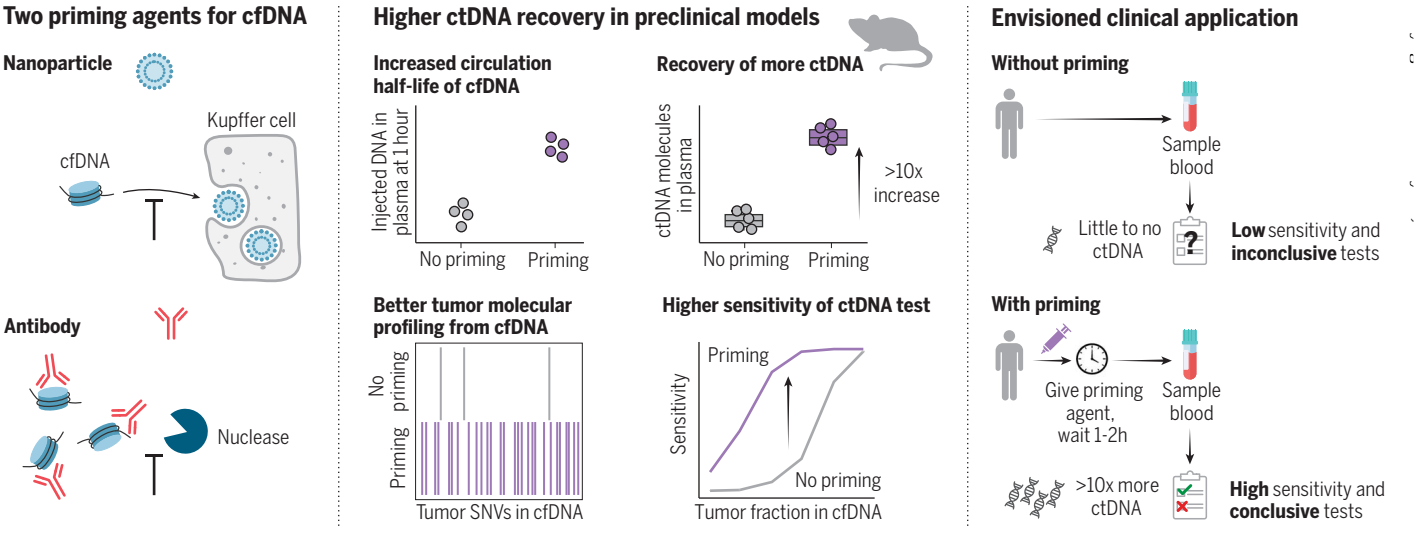
RATIONALE: We hypothesized that transiently attenuating cfDNA clearance in vivo would augment the levels of ctDNA in circulation and increase the amount recovered from a blood draw. The two natural mechanisms for clearing cfDNA are uptake by liver-resident macrophages and degradation by circulating nucleases. In this work, we sought to develop two intravenous priming agents given 1 to 2 hours before a blood draw that act on these mechanisms and enhance ctDNA recovery. Our priming agents comprise (i) nanoparticles that act on the cells responsible for cfDNA clearance and (ii) DNA-binding monoclonal antibodies (mAbs) that protect cfDNA.

RESULTS: We first investigated the nanoparticle priming strategy and identified a succinyl phosphoethanolamine–based liposomal agent that inhibited cfDNA uptake in vitro and transiently increased the recovery of cfDNA from blood in healthy mice. We confirmed that liposomes rapidly accumulated in the liver and that liver resident macrophages were necessary for cfDNA half-life extension. As an orthogonal strategy, we showed that DNA-binding mAbs interacted with elements of cfDNA and protected double-stranded DNA from nuclease digestion. Engineering the mAb to abrogate Fc-γ-receptor (FcγR) binding increased its persistence in circulation and the recovery of cfDNA from blood compared with that of the native mAb and an isotype control mAb in healthy mice. Using a bespoke ctDNA assay tracking 1822 tumor-specific single-nucleotide variants (SNVs) in plasma samples from mouse preclinical cancer models, we demonstrated that our two orthogonal priming strategies increase the recovery of ctDNA by >10-fold, enable more complete tumor molecular profiling from ctDNA, and increase the sensitivity for detection of small tumors from <10% to >75%.

CONCLUSION: By modulating cfDNA clearance in vivo, priming agents improved the sensitivity and robustness of ctDNA testing in tumor-bearing mice. Just as intravenous contrast agents have profoundly improved clinical imaging, we envision that priming agents will improve the sensitivity and utility of liquid biopsies across clinical applications. Additionally, the concept of delivering priming agents that transiently attenuate analyte clearance in vivo and boost diagnostic sensitivity may inform similar approaches to enhance the testing for other scarce biomarkers in oncology and beyond. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: shervin@broadinstitute.org (S.T.); clove@mit.edu (J.C.L.); sbhatia@mit.edu (S.N.B.); viktor@broadinstitute.org (V.A.A.)
†These authors contributed equally to this work.
Cite this article as C. Martin-Alonso *et al.*, *Science* **383**, eadf2341 (2024). DOI: 10.1126/science.adf2341

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adf2341



Priming agents reduce the clearance of cfDNA and enhance the sensitivity of liquid biopsies. Priming agents transiently attenuate natural clearance mechanisms for cfDNA and consist of nanoparticles that act on the cells responsible for cfDNA clearance (top left) or DNA-binding antibodies that protect cfDNA from cellular uptake and enzymatic digestion (bottom left). In preclinical models, priming agents increased the half-life of cfDNA, enhanced recovery of ctDNA, and improved tumor molecular profiling from ctDNA and sensitivity of ctDNA testing (middle). We envision that priming agents could be administered 1 to 2 hours prior to a blood draw in the clinic to improve the recovery of ctDNA and boost the sensitivity of many types of liquid biopsy tests (right).

RESEARCH ARTICLE

LIQUID BIOPSY

Priming agents transiently reduce the clearance of cell-free DNA to improve liquid biopsies

Carmen Martin-Alonso[1,2]†, Shervin Tabrizi[1,3,4,5]*†, Kan Xiong[3]†, Timothy Blewett[3], Sainetra Sridhar[3], Andjela Crnjac[3], Sahil Patel[1,3,6], Zhenyi An[3], Ahmet Bekdemir[1], Douglas Shea[3], Shih-Ting Wang[1], Sergio Rodriguez-Aponte[1,7], Christopher A. Naranjo[1], Justin Rhoades[3], Jesse D. Kirkpatrick[1,2], Heather E. Fleming[1], Ava P. Amini[8], Todd R. Golub[3,5,9], J. Christopher Love[1,3,10]*, Sangeeta N. Bhatia[1,2,3,11,12,13,14]*, Viktor A. Adalsteinsson[3]*

Liquid biopsies enable early detection and monitoring of diseases such as cancer, but their sensitivity remains limited by the scarcity of analytes such as cell-free DNA (cfDNA) in blood. Improvements to sensitivity have primarily relied on enhancing sequencing technology ex vivo. We sought to transiently augment the level of circulating tumor DNA (ctDNA) in a blood draw by attenuating its clearance in vivo. We report two intravenous priming agents given 1 to 2 hours before a blood draw to recover more ctDNA. Our priming agents consist of nanoparticles that act on the cells responsible for cfDNA clearance and DNA-binding antibodies that protect cfDNA. In tumor-bearing mice, they greatly increase the recovery of ctDNA and improve the sensitivity for detecting small tumors.

Liquid biopsies such as blood draws are a source of biological analytes such as cell-free DNA (cfDNA), and enable noninvasive diagnosis, monitoring, and molecular profiling of disease (*1*). The number of diagnostics based on liquid biopsies has grown rapidly over the last two decades in prenatal testing (*2*), infectious disease (*3*), oncology (*4*), and organ transplant monitoring (*5*), but the sensitivity of liquid biopsies remains inadequate for many applications. For example, in oncology the sensitivity of circulating tumor DNA (ctDNA)–based screening tests is low (~20 to 40% for Stage I cancer) (*6*, *7*), and liquid biopsies can be inconclusive in up to 40% of patients with advanced cancer (*8*). Additionally, up to 75% of patients who test negative for minimal residual disease after surgery experience recurrence (*9*–*11*).

To date, efforts to improve the sensitivity for detecting ctDNA have focused on sequencing and analysis (*12*, *13*), such as tracking multiple somatic variants (*9*, *10*, *14*–*16*) and integrating features such as DNA methylation or fragmentation patterns (*17*–*19*). An intrinsic challenge for all these methods is the scarcity of ctDNA in the collected samples, which limits sensitivity (*20*, *21*). One option to improve sensitivity is to draw larger volumes of blood (*4*) or to perform plasmapheresis (*16*). Large volumes, however, are impractical in frail or ill patients, and plasmapheresis carries major risks and requires expensive instrumentation. Alternatively, methods to sample more proximally to the tumor (*22*) or to increase tumor DNA shedding have been proposed (*23*, *24*). These methods require prior knowledge of tumor location, are limited to specific primary tumors, and often require specialized, expensive, and invasive procedures.

To realize a generalized approach for enhancing the amount of ctDNA recovered in any blood collection, we have developed two intravenous priming agents that transiently delay cfDNA clearance in vivo (Fig. 1A). The two natural mechanisms for clearing cfDNA are uptake by liver-resident cells of the mononuclear-phagocyte system (MPS) (*25*, *26*) and degradation by circulating nucleases (*27*) (Fig. 1B, left). Given that the majority of cfDNA circulates while bound to histone proteins as nanoparticulate mononucleosomes (~11 nm in diameter) (*1*), we hypothesized that a competing nanoparticle, such as a liposome, that is efficiently phagocytosed by the cells of the MPS would attenuate cfDNA cellular clearance (Fig. 1B, middle). Although the notion of saturating MPS uptake with a nanoparticle has been explored therapeutically to decrease the hepatic accumulation of nanomedicines (*25*, *28*–*31*), we have now applied this strategy to increase the abundance of an endogenous analyte for enhancing a diagnostic signal. As an orthogonal strategy, we also hypothesized that a DNA-binding priming agent could directly protect cfDNA itself from circulating DNases and extend its half-life in circulation (Fig. 1B, right). For this affinity-based approach, we selected monoclonal antibodies (mAbs) to develop given their persistence in circulation, ease of engineering, established manufacturing processes, and well-established safety and efficacy as biopharmaceuticals (*32*, *33*). In this work, we show that both approaches to priming agents improve recovery of ctDNA by more than 10-fold, enable better molecular profiling of tumors from blood samples, and increase the sensitivity for detection of small tumors from <10 to >75% in preclinical cancer models.

Nanoparticle priming agent attenuates cfDNA uptake by cells of the MPS

To test our hypothesis that administering liposomes inhibits cellular uptake of cfDNA, we first designed an in vitro two-dimensional assay using the murine macrophage cell line J774A.1 (Fig. 2A). Following pretreatment of J774A.1 cells with liposomes, we added Cy5-labeled mononucleosomes (fig. S1) and quantified their uptake. Empty liposomes were generated with cholesterol (50 mol%) and one of three different lipids [1,2-dipalmitoylsn-glycero-3-phosphoethanolamine-N-(succinyl) (SPE), 1,2-distearoyl-sn-glycero- 3-phospho-(1'-rac-glycerol) (DSPG), or 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC) (50 mol%)] sometimes used in FDA-approved liposomal formulations (*34*) (figs. S2 and S3). The average hydrodynamic diameter of the liposomes was between 230 and 260 nm and designed to match the size of the fenestrae of murine liver capillaries (*31*) such that they would preferentially target liver-resident macrophages over hepatocytes. The SPE- and DSPG-based formulations, but not the DSPC-based one, significantly ($P < 0.05$) inhibited the uptake of mononucleosomes by macrophages in a dose-dependent manner (Fig. 2, B and C). These two formulations are more negatively charged, consistent with prior reports that negatively charged particles display increased interactions with macrophages versus neutrally charged particles (*35*, *36*). Using SPE liposomes, we confirmed that inhibition of mononucleosome uptake was also dose dependent in the independent macrophage cell line RAW264 (fig. S4, A and B). Cell viability was not compromised with liposome treatment (fig. S4, C and D), and the liposomes did not impair phagocytosis of inactivated *Escherichia coli* in J774A.1 cells at the range of

[1]Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [2]Harvard-MIT Division of Health Sciences and Technology, Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [4]Department of Radiation Oncology, Massachusetts General Hospital, Boston, MA 02114, USA. [5]Harvard Medical School, Boston, MA 02115, USA. [6]Division of Pulmonary and Critical Care, Department of Medicine, Massachusetts General Hospital, Boston, MA 02124, USA. [7]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [8]Microsoft Research, Cambridge, MA 02142, USA. [9]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA. [10]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [11]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [12]Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA. [13]Wyss Institute at Harvard University, Boston, MA 02215, USA. [14]Howard Hughes Medical Institute, Cambridge, MA 02138, USA.
***Corresponding author. Email: shervin@broadinstitute.org (S.T.); clove@mit.edu (J.C.L.); sbhatia@mit.edu (S.N.B.); viktor@broadinstitute.org (V.A.A.)**
†These authors contributed equally to this work.

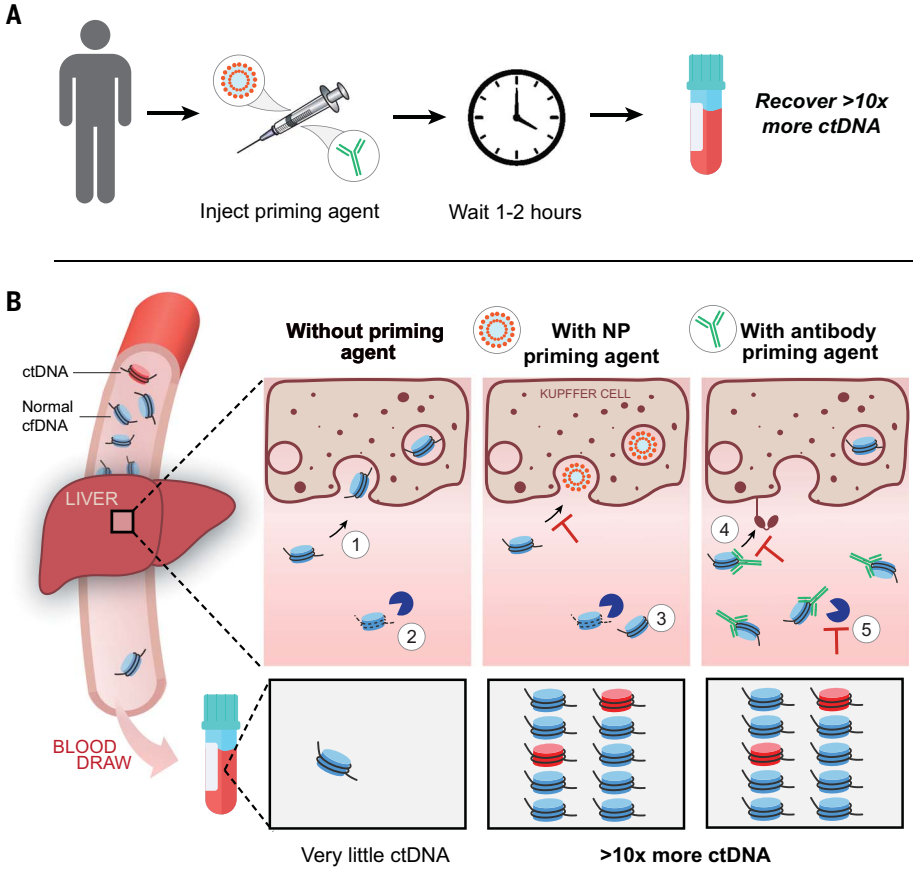


Fig. 1. Priming agents reduce clearance of cfDNA and improve the recovery of ctDNA. (**A**) Priming agents are injected 1 to 2 hours prior to a blood draw and improve the recovery of ctDNA by >10-fold. (**B**) (Left) In the absence of a priming agent, cfDNA (mostly in the form of mononucleosomes) is (i) rapidly taken up by macrophages of the MPS in the liver and (ii) degraded by circulating nucleases, yielding little ctDNA molecules in a blood draw. (Center) Following intravenous administration of a nanoparticle (NP) priming agent, (iii) cellular uptake is attenuated through MPS saturation. (Right) Intravenous administration of an antibody priming agent (iv) extends the half-life of cfDNA in circulation and (v) protects it from nuclease digestion. Both priming strategies enhance ctDNA recovery and improve mutation detection from a blood draw.

concentrations tested (fig. S5). These data suggest that our priming agent may not affect normal phagocytic pathways that play an important role in defending the host from infection (*37*).

We next assessed whether SPE-based liposome administration would decrease cfDNA clearance in vivo. We administered liposomes to healthy mice, followed by exogenous mononucleosomes carrying the Widom601 sequence (W601) (*38*), then quantified the levels of W601 in plasma over time. We observed that the percentage of injected W601 in plasma 60 min after administration increased as liposome doses increased [mean increase between 9- and 3198-fold at liposome doses of 50 and 300 mg/kg, respectively, relative to phosphate-buffered saline (PBS) administration; $P < 0.05$] (Fig. 2D).

To determine whether attenuated clearance of mononucleosomes would translate to higher concentrations of endogenous cfDNA in plasma, we next administered liposomes or PBS into healthy mice and measured cfDNA levels in blood collected longitudinally. We observed an increase in the concentration of endogenous cfDNA in plasma after liposome injection (Fig. 2E), with peak concentration achieved 30 min after administration of a 100 mg/kg dose (10.3-fold increase over PBS) or 3 hours after administration of a 300 mg/kg dose (78.0-fold increase over PBS). Notably, levels returned to baseline within 5 and 24 hours of liposome treatment at the lower and higher doses, respectively, suggesting a transient effect of liposomes on cfDNA levels. We confirmed that SPE liposomes rapidly accumulated in MPS organs in vivo (Fig. 2E, insert, and fig. S6), where they can interact with liver macrophages responsible for mononucleosome clearance (fig. S7), and that depletion of MPS macrophages by means of liposomal clodronate eliminated the effect of liposomes on cfDNA clearance (fig. S8). Together, these results suggest that uptake of SPE nanoparticles by MPS macrophages can attenuate the cellular uptake of mononucleosomes and increase the recovery of endogenous cfDNA from a blood draw.

Antibody priming agent protects cfDNA from clearance

We next investigated whether directly protecting cfDNA with a DNA-binding mAb could provide an alternative method to increase recovery of ctDNA in a blood draw. Of the eight known mouse anti-DNA immunoglobulin G (IgG) antibodies that we tested for double-stranded DNA (dsDNA) binding activity, four showed detectable binding at similar levels of affinity (fig. S9A). From these four, we selected a mouse IgG2a mAb (35I9) derived from a NZWxNZB F_1 lupus-prone mouse for further investigation, given its reported biochemical and binding characterization, with dissociation constant (K_d) of 90 and 700 nM to dsDNA and single-stranded DNA (ssDNA), respectively (*39*).

We first explored the interaction of 35I9 with elements of cfDNA (mononucleosomes and free dsDNA) using electrophoretic mobility shift assays (EMSAs). EMSA of 35I9 with a mixture of free and histone-bound 147-bp dsDNA revealed H3-negative bands corresponding to discrete ratios of mAb-to-dsDNA, as well as H3-positive bands corresponding to the binding of one or more than one mAb to histone-bound dsDNA (Fig. 3A and fig. S9B). 35I9 demonstrated rapid association and dissociation kinetics and similar binding affinity to various dsDNA oligonucleotides in vitro (fig. S10). To evaluate whether the observed interactions would interfere with nuclease activity, we next characterized the susceptibility of a fluorescence-quenched dsDNA probe to deoxyribonuclease (DNase) I degradation when incubated with different concentrations of 35I9. The fluorescent signal generated by cleavage of DNA in the presence of DNase I diminished with increasing concentrations of 35I9 (Fig. 3B). Together, these data demonstrate the ability of mAbs to interact with both free and histone-bound DNA and protect dsDNA from nuclease digestion.

To test mAb activity in vivo, we injected mononucleosomes carrying W601 with 35I9, without 35I9, or with IgG2a control into mice and measured the concentrations of W601 in plasma over time (Fig. 3C). Although the relative clearance of W601 was significantly ($P < 0.05$) delayed with mAb treatment (fig. S11), the absolute quantity of W601 recovered at 60 min was similar between 35I9 and the IgG2a control (Fig. 3C). We hypothesized that this lack of difference at 60 min was due to Fc-γ receptor (FcγR)–mediated clearance of dsDNA-35I9 complexes (*40*, *41*). This effect could relate to some of the larger complexes observed in vitro (Fig. 3A), which would be expected to be sequestered and cleared rapidly through FcγR in vivo (*42*). Coinjection of the W601-mAb

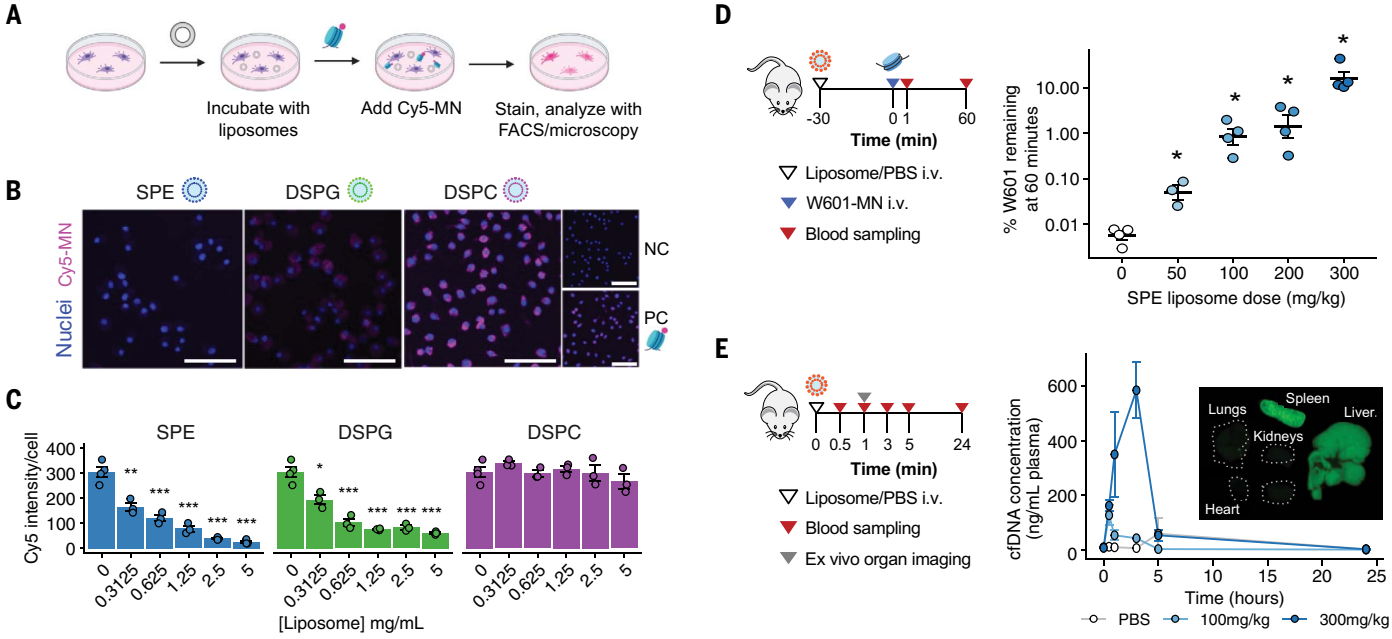


Fig. 2. SPE liposomes inhibit the uptake of mononucleosomes by macrophages in vitro and increase the recovery of cfDNA through decreased clearance in healthy mice. (**A**) Schematic of in vitro macrophage uptake inhibition assay. (**B**) Representative images of uptake of Cy5-labeled mononucleosomes (Cy5-MN) following incubation of J774A.1 with different liposomes at 5 mg/ml, without liposomes or Cy5-MN (negative control, NC), or with Cy5-MN only (positive control, PC). Scale bars, 100 nm. (**C**) Quantification of Cy5-MN uptake by J774A.1 cells from epifluorescence images after liposome pre-treatment (mean ± SEM, n = 3 to 4 wells per condition, N = 2). *P < 0.05; **P < 0.01; ***P < 0.001; one-way ANOVA. (**D**) (Left) Experimental approach to determine the plasma bioavailability of W601-mononucleosomes (W601-MN) following SPE liposome priming. (Right) Percentage of W601 remaining in plasma 60 min after administration of different SPE liposome doses (median, n = 3 to 4 mice per group, N = 2). *P < 0.05, two-tailed Mann-Whitney test. (**E**) (Left) Experimental approach to determine plasma cfDNA yields and liposome organ biodistribution. (Right) Plasma cfDNA concentration following Cy7-SPE administration (mean ± SEM, n = 3 mice per group). The largest elevation relative to the PBS group was at 30 min with the dose of 100 mg/kg liposome (10.3-fold, *P = 0.034) and at 3 hours with the dose of 300 mg/kg liposome (78.0-fold, **P = 0.005) (unpaired two-tailed t test; n = 3 mice per group, N = 1). (Insert) Organ biodistribution of Cy7-SPE liposomes 1 hour after administration. Images from a representative mouse are shown (n = 4 mice, N = 3).

preparation and antibodies blocking mouse Fc(γ)RI, Fc(γ)RII, and Fc(γ)RIII yielded higher W601 levels at 60 min (0.012 pg/μL versus 0.00043 pg/μL; P = 0.007) (Fig. 3C). Together, these results suggest that administration of DNA-binding mAbs can delay the clearance of dsDNA from blood, but that FcγR-mediated clearance of dsDNA bound to mAb reduces the benefits for prolonged stabilization of the dsDNA.

Engineered variants of the Fc domain of mAbs provide one way to modulate interactions with FcγR and have been used in biopharmaceutical candidates clinically (*43*). We selected three sets of sequence variants known to disrupt FcγR binding (*44*)– aglycosylated N297A (denoted aST2) (*45*–*47*), L234A/L235A/P329G (denoted aST3) (*48*), and D265A (denoted aST5) (*49*, *50*) (Fig. S12). All three variants still bound to dsDNA (Fig. S13). In vivo, aST3 yielded the highest recovery of W601 at 60 min (Fig. 3D, Fig. S14; 0.641 pg/μL vs. 0.004 pg/μL, P = 0.007), and was investigated further. We compared the pharmacokinetics of fluorophore-labeled aST3 and the Fc–wild type (WT) equivalent 3519 mAb and observed that aST3 levels were elevated in plasma (Fig. 3E). W601 levels were below the limit of detection by 24 hours with or without aST3 (fig. S15), consistent with a transient effect. We also compared the biodistribution of aST3 and 3519. The area-corrected accumulation of both mAbs was similar in the liver (fig. S7) but reduced in the spleen (Fig. 3F), suggesting differences in the clearance of aST3 by the two organs (*51*). Together, these data suggest that aST3, a DNA-binding mAb with abrogated FcγR binding, protects cfDNA from enzymatic digestion, has higher persistence in circulation, and increases cfDNA recovery from plasma compared with the native mAb and an IgG2a control.

Nanoparticle priming agent improves tumor detection

Because both liposomal and antibody priming agents showed increased recovery of cfDNA in healthy mice, we next explored whether they could enhance ctDNA-based tumor detection using a tumor-informed approach, tracking 1822 tumor-specific single-nucleotide variants (SNVs) (*9*, *52*) (fig. S16). After selecting one hour as the optimal time post liposome administration for blood sampling (fig. 2E) (*52*), we performed an experiment with escalating doses of liposomes in a flank tumor model (figs. S17 and S18 and data S1) and selected a liposome dose of 100 mg/kg for further testing in a more disease-relevant transplantation model of lung metastases (fig. S19 and data S2). In this model, the luciferized MC26 cell line (Luc-MC26) was injected intravenously to establish lung metastases. Plasma was collected once a week at different stages of tumor progression, each time at one hour after administration of liposomes or PBS (Fig. 4A). We observed that administration of liposomes significantly (P < 0.05) increased concentrations of plasma cfDNA (7-fold, 14-fold, and 28-fold at weeks one, two, and three, respectively) (Fig. 4B) and the number of mutant molecules recovered at each time point (4-fold, 19-fold, and 60-fold) (Fig. 4C) relative to PBS-treated mice (data S3; independent replicate at week two, fig S20 and data S4). The maximum improvement in mutant molecule recovery (~60-fold) was observed at week three. Moreover, additional SNVs were detected after the administration of liposomes (6-fold and 90-fold higher at two and three weeks, respectively) (figs. S21 an S22). Liposome administration did not significantly decrease the tumor fractions (the fraction of total cfDNA originating from the tumor) in this experiment (P > 0.05) (Fig. 4D) but did reach significance (P < 0.05) in an independent cohort (fig. S20). Incubating

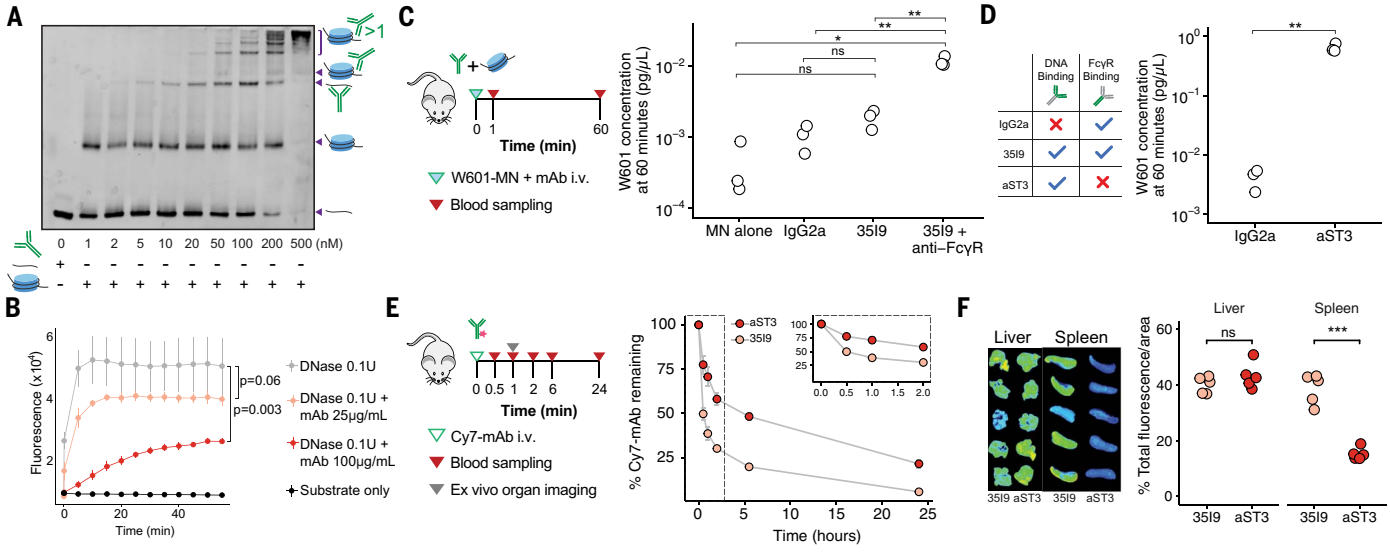


Fig. 3. Antibody priming agent binds cfDNA and attenuates its clearance in healthy mice. (**A**) EMSA of free and histone-bound dsDNA (4 ng/µL per lane) with varying concentrations of DNA-binding mAb 35I9 in PBS (N = 3). (**B**) Fluorescence signal from the digestion of a DNA substrate carrying a hexachlorofluorescein dye on one end and a dark quencher on the other, with or without 0.1 U of DNase I and mAb 35I9. Points indicate mean and lines indicate SEM of three technical replicates. Fluorescence signals across the whole experiment were compared by using mixed models with replicates as random effects (N = 2). (**C**) (Left) Experimental approach to evaluate the effect of priming mAb on dsDNA clearance. (Right) Concentration of W601 in plasma 60 min after injection of W601 only, coinjection with an unrelated IgG2a antibody, with 40 µg of DNA-binding antibody 35I9, or with 40 µg of 35I9 together with anti-FcγRI (20 µg), anti-Fc(γ)RII, and anti-Fc(γ)RIII (40 µg) (n = 3 mice per group, N = 2). (**D**) (Left) Overview of the DNA binding and FcγR binding properties of engineered mAb aST3 versus IgG2a control mAb and DNA-binding mAb 35I9. (Right) Concentration of W601 in plasma 60 min after coinjection of W601 with an unrelated IgG2a antibody or with the Fc-mutant aST3 DNA-binding antibody (n = 3 mice per group, N = 3). (**E**) (Left) Experimental approach to quantify pharmacokinetics of 3519 and aST3 labeled with AQuora750 in plasma. (Right) Plasma clearance of antibodies over time (mean ±SEM, n = 5 mice per group, N = 1). (**F**) Biodistribution and quantification of 3519 (Fc-WT antibody) or aST3 concentration in liver and spleen 1 hour after administration (n = 5 mice per group, N = 1). [(C), (D), (F)] ns, not significant; *P < 0.05; **P < 0.01; ***P < 0.001; one-way ANOVA.

primary murine white blood cells with liposomes in vitro led to a dose-dependent increase in the detection of DNA in conditioned medium, as measured using SYTOX green dye (fig. S23), suggesting that cfDNA release by white blood cells exposed to high concentrations of liposomes (*53*) may contribute to the modest decrease in tumor fraction observed.

We next assessed how the enhancement in recovered mutant molecules would impact the performance of ctDNA analyses, such as tumor genome profiling and sensitivity for tumor detection. In the absence of priming, most high-burden tumors (burden > total flux 1.5e8 p/s) were detectable, but priming with liposomes enabled detection of 67-fold (median) more SNVs than PBS (Fig. 4E), providing a more comprehensive molecular profile of these tumors. We next evaluated the sensitivity of ctDNA testing with and without priming by classifying each plasma sample as ctDNA positive only if the number of SNVs detected surpassed a given SNV threshold (between 2 and 10 SNVs, from lower to higher test stringencies). The liposomes improved the sensitivity of the ctDNA test (defined as the fraction of samples that were classified as ctDNA positive), regardless of SNV threshold, with the largest improvement in sensitivity observed in the group with the lowest tumor burden (burden < total flux 1.5e7 p/s) (Fig. 4F and fig. S24). By using a threshold of two SNVs, as has been previously applied to clinical samples (*9*), cancer was not detected in any of the untreated low–tumor burden mice, whereas 75% of liposome-primed mice were diagnosed as tumor-bearing with the same threshold. Improvements in sensitivity became smaller in the medium- and high-burden groups, as the untreated cohorts already had substantial levels of ctDNA prior to priming. Furthermore, the liposomes did not affect tumor progression (fig. S25) or evoke acute toxicity or weight loss after repeated dosing in healthy mice (fig. S26). Taken together, these results suggest that the nanoparticles enable profiling of more of the tumor genome and improve the sensitivity of a ctDNA-based test to enable detection of smaller tumors in preclinical models.

Antibody priming agent improves tumor detection

We next explored the effect of our antibody priming agent on ctDNA-based tumor detection in the same transplantation model of lung metastases. We tested our antibody priming agent at a range of doses (0.5 to 8 mg/kg aST3 versus IgG2a Control at 8 mg/kg) at a single time point (2 weeks) during tumor progression (Fig. 5A). We sampled blood two hours after administering the mAb, as this interval corresponded to the peak accumulation of endogenous cfDNA in plasma after injection of aST3 in healthy mice (fig. S27). Accordingly, we also observed significantly (P < 0.001) higher recovery of cfDNA from plasma at all concentrations of mAbs (compared with an IgG2a isotype control) in tumor-bearing mice (Fig. 5B).

Administration of mAb resulted in consistently higher concentrations of mutant molecules with aST3 compared with IgG2a control, with a dose-dependent improvement between 0.5 and 4.0 mg/kg (Fig. 5C, fig. S28, and data S5; independent replicate at 4.0 mg/kg aST3, fig. S29 and data S4). The maximum effect occurred at a dose of 4.0 mg/kg, with a median 19-fold improvement over IgG2a isotype control. With this agent, no difference in tumor fraction was observed between the groups post injection (Fig. 5D). We also detected more total SNVs when priming with the engineered mAb (median 77% of SNVs detected with 4.0 mg/kg versus 15% detected with IgG2a isotype control) (Fig. 5E), again suggesting that priming improves the genomic profiling of tumors from a liquid biopsy. Consistent with the nuclease protection afforded by the DNA-binding mAb (Fig. 3B), we also found that priming resulted in greater enrichment of parts of the genome close to or overlapping with DNase hypersensitivity

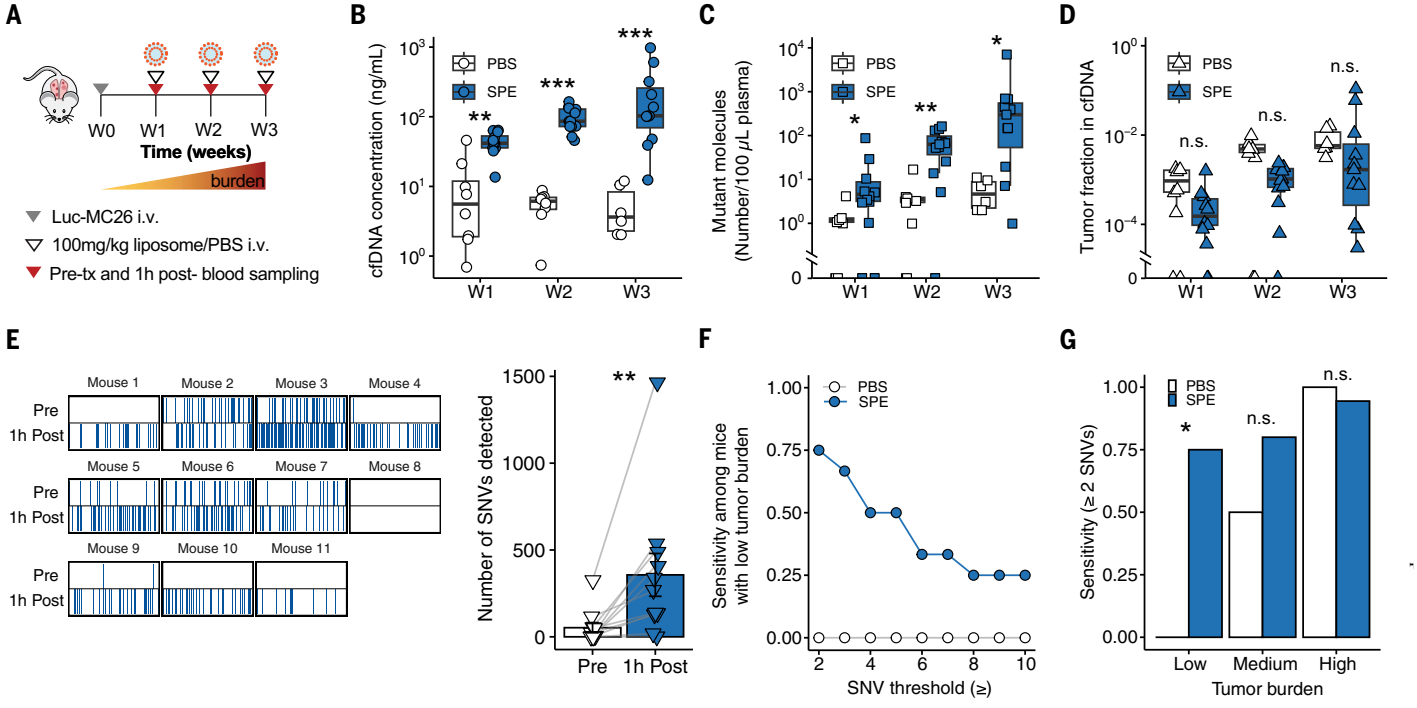


Fig. 4. Liposome priming improves ctDNA recovery and enables detection of smaller tumors in a murine lung metastasis model. (**A**) Experimental approach for the detection of mutations from the plasma of Luc-MC26 tumor-bearing mice using the liposome priming agent. For each mouse, blood was drawn prior to and 1 hour after i.v. administration of PBS or SPE liposomes (100 mg/kg) at 1 week (W1), 2 weeks (W2), and 3 weeks (W3) after tumor inoculation. (**B**) Plasma cfDNA concentrations, (**C**) concentration of mutant molecules detected, and (**D**) tumor fractions 1 hour after PBS (white) or SPE (blue) administration at W1, W2, or W3 (n = 6 to 12 mice per group). (**E**) (Left) Mutational fingerprints showing distinct SNVs detected pre– and post–SPE administration for mice with high tumor burden (burden > 1.5e8 p/s total flux, as measured by IVIS). Each vertical band corresponds to a SNV in our 1822-SNV panel and is colored blue if detected at least once in the plasma sample. (Right) Quantification of these distinct SNVs. (**F**) Sensitivity of ctDNA tests versus SNV threshold for tumor detection in mice with low tumor burden after administration of PBS or SPE liposomes (burden < 1.5e7 p/s total flux, as measured by IVIS). (**G**) Sensitivity of ctDNA tests for different tumor burdens after PBS or SPE administration (Low, burden < 1.5e7 p/s; Medium, 1.5e7 p/s ≤ burden ≤ 1.5e7 p/s; High, burden > 1.5e8 p/s). Sensitivity was calculated as the fraction of samples for which the number of SNVs detected in a blood sample was ≥ 2 (n = 6 to 12 mice per group; *P < 0.05, Chi-squared test) (fig. S20, independent replicate at week 2). Boxplots in (B), (C), (D), and (E) show median and interquartile range. ns, not significant; *P < 0.05; **P < 0.01; ***P < 0.001; one-way ANOVA.

peaks (figs. S30, A to C, and S31). We also observed enrichment of sites with higher GC content and those overlapping CpG islands (figs. S30, D and E, and S31) (*52*).

We next investigated the effect of our priming agent on the sensitivity of ctDNA assays. Recognizing that our conditions in this preclinical model may not be representative of current commercial assays that typically have smaller mutation panels (*9*, *54*, *55*), or of much lower tumor fractions typically observed in early detection and minimal residual disease settings, we estimated the benefit of priming in such settings through a computational downsampling approach (*52*). Across a wide range of panel sizes and detection thresholds, we consistently observed superior sensitivity with our priming agent compared with that of the IgG2a isotype control (Fig. 5F and fig. S32). We then evaluated the effect of our priming agent on ctDNA assay sensitivity at lower tumor DNA abundance (lower tumor fractions) (fig. S33) (*52*) and found that our priming agent resulted in similar sensitivity to that of the IgG2a isotype control at approximately 10-fold lower tumor fraction, irrespective of the SNV threshold used (fig. S34). By using a threshold of two SNVs, priming with aST3 improved the sensitivity across all different tumor fractions modeled, including at tumor fractions of 1 to 10 parts per million that are typical in the context of low tumor burden or minimal residual disease (Fig. 5G) (*11*, *56*). These detection levels were reached in samples of mouse plasma with mean volumes of only 0.33 mL (SD, 0.09 mL), >10-fold less than plasma from a typical blood draw in humans (4 mL).

Discussion

We have developed intravenous priming agents for liquid biopsies: agents that are given 1 to 2 hours prior to a blood draw to enable recovery of more cfDNA in a blood sample. The liposomal nanoparticles attenuate the uptake capacity of cfDNA by liver macrophages, whereas the DNA-binding antibody aST3 protects the cfDNA itself from nuclease degradation and plasma clearance. Both agents increase the recovery of ctDNA molecules from blood >10-fold, enable more of the tumor genome to be recovered in a blood draw, and enhance the sensitivity of ctDNA diagnostic tests.

Our priming agents intervene in vivo on the natural clearance pathways of cfDNA to boost ctDNA recovery, addressing the well-recognized barrier of low quantities of input cfDNA that limits the sensitivity of liquid biopsy tests (*16*, *57*, *58*). Sampling larger blood volumes has traditionally been used to increase the total quantity of cfDNA available for assays, but with only modest linear increases in recovery given the notable practical limitations on sampling large volumes of blood. The priming agents we describe increase the concentration of cfDNA in blood prior to sampling. These approaches are also distinct from those that rely on local sampling, such as lymph fluid or bronchoalveolar lavage (*22*, *59*), because they preserve the advantages of a blood draw: sampling from all potential disease sites and avoiding the need for specialized, invasive, and disease-specific sampling procedures. Our antibody

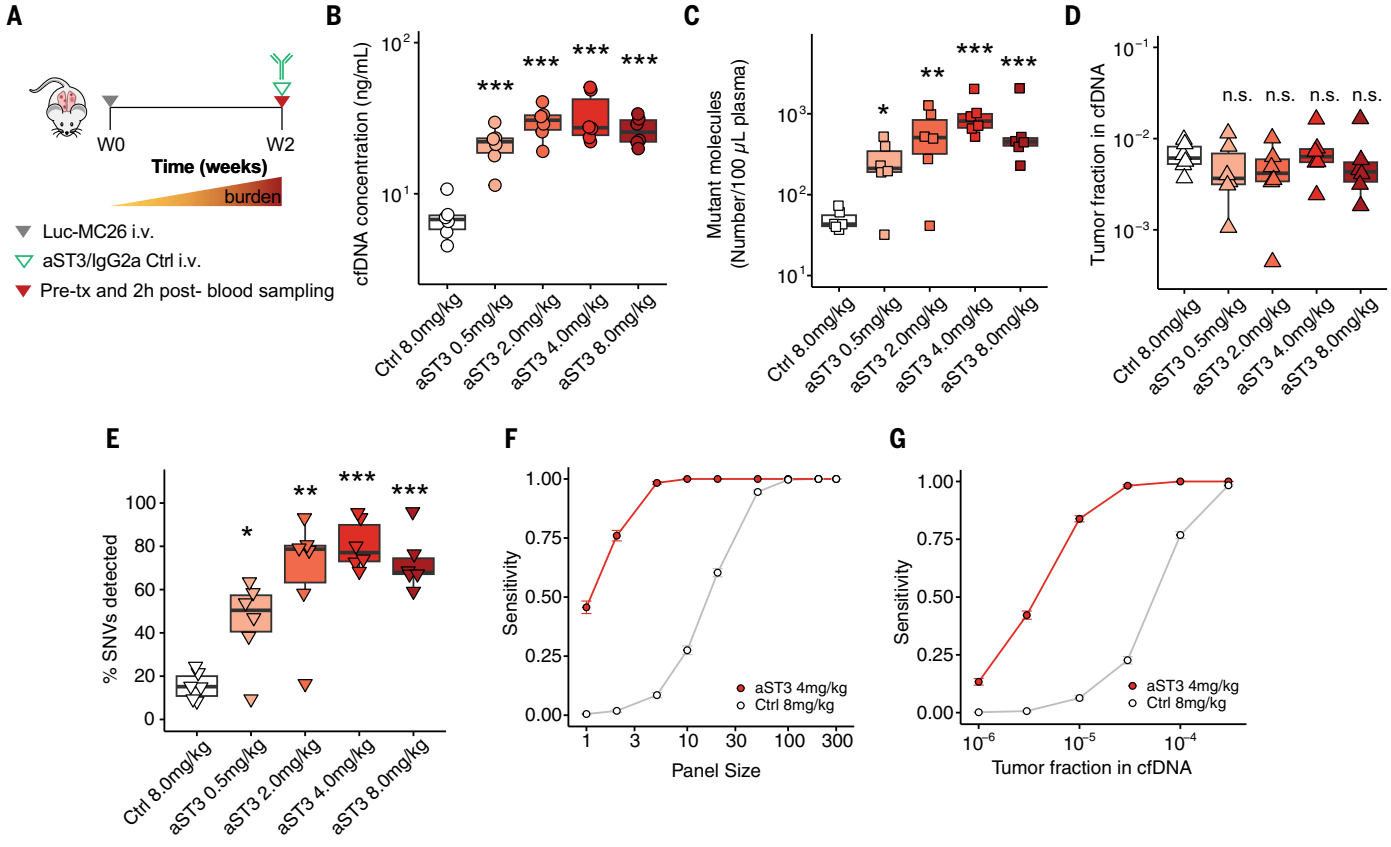


Fig. 5. Antibody priming agent improves ctDNA recovery in murine lung metastasis model. (**A**) Experimental approach for the detection of mutations from the plasma of Luc-MC26 tumor-bearing mice with the antibody priming agent aST3. (**B**) Plasma cfDNA concentrations, (**C**) concentration of mutant molecules detected, and (**D**) tumor fractions detected 2 hours after administration of IgG2a control mAb or various doses of aST3 (n = 6 mice per group) (fig. S29, independent replicate at aST3 4.0 mg/kg). (**E**) Percentage of distinct SNVs from an 1822-SNV panel detected in plasma with control mAb or various doses of aST3. (**F** and **G**) Estimation of sensitivity for detection of ctDNA upon administration of 8 mg/kg of IgG2a control or 4 mg/kg of aST3 versus (F) panel size (G) or tumor fraction based on binomial down-sampling of mutant molecules, with a detection threshold of ≥ 2 SNVs (mean ± SEM, n = 100 replicates). Boxplots in (B) to (E) show median and interquartile range. ns, not significant; *P < 0.05; **P < 0.01; ***P < 0.001; one-way ANOVA.

priming agent showed 86% sensitivity at $^{1}/_{100,000}$ tumor fraction in 0.33-mL mouse plasma samples, a sensitivity on par with the best-performing genome-wide cfDNA tests reported to date, which use >10-fold higher plasma volumes from patient plasma samples (*57*, *60*). When scaling the sample volumes from mouse plasma to typical clinical blood draws, the sensitivity afforded by our priming agents could far exceed those reported in the literature. Furthermore, because these priming agents are given prior to collecting and processing liquid biopsies, they could also enhance existing genome-wide workflows (*16*, *18*, *19*) to maximize sensitivity.

Although results from our proof-of-concept studies in preclinical models are encouraging, it remains to be determined how these strategies would translate clinically. Further development prior to clinical testing of either agent would involve preclinical optimization, formulation, testing, and tolerability in other animal models. For the nanoparticles, optimizing formulations by using emerging technologies in nanoparticle engineering (*61*, *62*) could improve potency and mitigate dose-dependent reductions in tumor fractions. Additionally, investigating the cellular mechanisms driving the inhibition of cfDNA uptake, which may involve changes in membrane availability or composition (i.e., competition for or internalization of receptors) or feedback mechanisms in phagocytic signaling networks (*30*, *63*), could reveal additional avenues for development. For the antibody, higher affinity or alternative cfDNA binders could be explored to further improve recovery of cfDNA. One clinically relevant observation to support the translational potential of the antibody is from studies of the autoimmune disease systemic lupus erythematosus. A feature of this disease is elevated levels of anti-DNA antibodies. Higher concentrations of cfDNA have been associated with increased titers of anti-DNA antibodies along with reduced degradation of extracellular DNA (*64*, *65*). These observations support the potential efficacy of an antibody priming agent in humans. Furthermore, engineering of the Fc-effector function, as we demonstrated, could reduce or eliminate potential safety risks related to Fc-mediated immune activation (*66*–*68*) for transient administration of low doses, as tested here (*49*, *69*). In our testing, no sign of acute toxicity was observed with either agent. Future development work will be needed to evaluate safety in other animal models prior to first-in-human testing.

Because the two approaches have different targets (liver macrophages for nanoparticles and cfDNA in blood for antibodies), each has distinct advantages as a priming agent. For nanoparticles, interfering with the uptake capacity of macrophages could potentially enhance the recovery of other circulating analytes cleared through similar pathways. For antibodies, their target specificity could be further engineered to enhance the recovery of other analytes or of subpopulations of cfDNA molecules, such as those carrying specific epigenetic marks. The optimal approach would depend on the intended application. With our two approaches targeting different processes, a broad range of potential diagnostic applications as well as possible combinations of the two could be considered.

We envision that the initial clinical use of our priming agents could be in patients with a previous cancer diagnosis in which tumor detection or monitoring sensitivity is currently lacking. Priming could boost the sensitivity of minimal residual disease tests to guide clinical decisions, such as the use of adjuvant therapy or evaluating the efficacy of nonsurgical organ-preserving treatments. In patients with advanced cancer, priming could enable the detection of rare targetable mutations missed by conventional liquid biopsies. Looking ahead, priming could also boost the sensitivity of liquid biopsy cancer screening tests and would be especially useful for individuals at elevated risk of cancer, with nonspecific symptoms that may be associated with cancer, or with indeterminate findings from other diagnostics such as imaging scans. A notable example would be indeterminate nodules on lung computed tomography scans. Furthermore, given that our priming agents modulate cfDNA clearance, their use could be considered in applications beyond oncology. Priming could improve detection of microbial cfDNA during early or deep-seated infections (*70*), where diagnosis is critical for therapy selection but remains challenging. Liquid biopsy–based applications in cardiovascular disease and Alzheimer's disease are other areas where the low abundance of cfDNA is a limitation, and where priming agents may be beneficial (*71*, *72*). Deeper characterization of the effect of priming on other aspects of cfDNA, such as epigenetics and fragmentomics, could reveal further insights into cfDNA biology and motivate other applications. We believe that the concept of a priming agent capable of perturbing endogenous biomarker clearance in vivo can change how we think about the limit of diagnostic detection. These approaches should spark interest in the field, not only for further development of related priming agents for cfDNA detection, but also for improved detection of other circulating biomarkers.

In this work, we present liquid biopsy priming agents that improve the sensitivity and the robustness of ctDNA testing in tumor-bearing mice by modulating cfDNA clearance. Just as iodinated and gadolinium contrast agents greatly improve the sensitivity of clinical imaging, we envision that priming agents can boost the sensitivity of liquid biopsies in cancer care and for indications beyond oncology.

Materials and methods summary

Liposome synthesis and characterization

Liposomes were prepared using the lipid film rehydration method with slight modifications from the protocol described by Saunders *et al.* (*31*). Briefly, ovine cholesterol (50 mol %, cat. 700000P, Avanti Polar Lipids) was solubilized in chloroform and added to 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine-N-(succinyl) (sodium salt) (SPE) (50 mol %, cat. 870225P, Avanti Polar Lipids), 1,2-distearoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (sodium salt) (DSPG) (50 mol %, cat. 840465P, Avanti Polar Lipids), or 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC) (50 mol %, cat. 850365P, Avanti Polar Lipids) with 1:1 (v/v) methanol. The solution was evaporated under nitrogen flow to form a thin dry film and vacuumed overnight to remove any traces of organic solvent. The lipid film was hydrated at 60°C with sterile Dulbecco's phosphate-buffered saline (DPBS) to a total lipid concentration of 50 mg/ml. Extrusion was performed at 60°C with 1-μm (cat. WHA110410, MilliporeSigma) and 0.4-μm polycarbonate membranes (cat. WHA10417101, MilliporeSigma), 21 and 20 times respectively, using the 1000-μL Mini-Extruder from Avanti Polar Lipids (cat: 610023). For the fluorescent liposome used for biodistribution studies, 0.2 mol % of SPE was replaced for Cy7-SPE (cat: 810347C, Avanti Polar Lipids) prior to solubilization with organic solvents. The hydrodynamic diameter and polydispersity index of liposomes was characterized using a Zetasizer NanoZS (Malvern Instruments). The morphology of liposomes was confirmed by cryo–transmission electron microscopy imaging.

Mononucleosome preparation and labeling

To prepare mononucleosomes, chromatin was extracted from CT26 cells following manufacturer's recommendations using the Nucleosome Preparation Kit (cat. 53504, Active Motif). The enzymatic digestion time was optimized as 30 min, and the resulting mononucleosomes were confirmed via electrophoresis through a 1.5% agarose gel. Subsequently, aliquots of 10 μg mononucleosomes were washed and buffer-exchanged into PBS. Four washes were performed using 30-kDa Amicon filters (cat. UFC503024, EMD Millipore) by centrifugation at 12,000 rpm for 10 min at 4°C. The protein yield was calculated using a commercial HeLa mononucleosome standard by measuring absorbance at 230 nm using a Nanodrop 8000 Spectrophotometer (cat. ND-8000-GL, Thermo Fisher). To label mononucleosomes, sulfonated-Cy5 (cat. 13320, Lumiprobe) was added at a 25:1 molar ratio of dye to protein, and the reaction incubated at 4°C in an Eppendorf Thermomixer C Model 5382 (Eppendorf) at 550 rpm overnight. Excess dye was removed using Micro Bio-Spin Columns with Bio-Gel P-6 (cat.7326221, BioRad) by centrifugation at 1000*g* for 2 min at room temperature. Labeling efficiency was quantified by measuring Cy5 intensity at 650/680 nm against a Cy5 standard using an Infinite F200 Pro reader (Tecan) fluorometer and protein yield was estimated by measuring absorbance at 230 nm using a Nanodrop 8000 Spectrophotometer.

In vitro macrophage mononucleosome uptake inhibition assay with liposomes

J774A.1 (TIB-67, ATCC) cells were plated at a density of 30,000 and 45,000 cells per chamber, respectively, in 8-well chamber slides (cat.80806, Ibidi). Following overnight acclimatization, cells were incubated with 300 μL of liposomes (SPE, DSPG, or DSPC) diluted in Dulbecco's Modified Eagle Medium (DMEM) at the desired concentrations (0.1 to 5 mg/ml) for 4 hours at 37°C. Next, 30 μL of mononucleosomes were spiked into each well to achieve a final mononucleosome concentration of 10 nM and further incubated for 2 hours at 37°C. Cells incubated with DMEM followed by mononucleosome addition were used as a positive control for uptake, and cells incubated only with DMEM were used as a negative control. At the end of the incubation, cells were washed once with DMEM, stained with Hoechst 33342 (cat. H3570, ThermoFisher) at 1:2000 dilution in DMEM for 10 min at room temperature, and further washed (twice with DMEM and once with PBS) to remove any extracellular mononucleosomes. Subsequently, cells were fixed with 4% PFA for 20 min at room temperature and washed with PBS prior to imaging on an Eclipse Ti microscope (Nikon).

To quantify cellular uptake, four fields of view per well were obtained at 10X magnification, and mean Cy5 fluorescence intensity per cell was quantified using custom scripts in QuPath (*73*). Results are displayed after background subtraction using the mean Cy5 fluorescence intensity per cell from the negative control.

Electrophoretic mobility shift assays (EMSA)

Widom601 dsDNA complexed with recombinant human histones was purchased from Epicypher (cat. 16-0009). dsDNA (free and/or histone bound) was combined at a final concentration of 4 ng/μL total DNA with varying concentrations of 35I9 (Abcam ab27156) in PBS (21-040-CM, Corning) in 10 μL total volume. 1 μL of Novex high density TBE sample buffer (cat. LC6678, Thermo Fisher Scientific) was added, and 10 μL of mixture was loaded into 6% DNA Retardation Gels (cat. EC6365BOX, Thermo Fisher Scientific). Gels were run at 4°C, 100 V for 120 min in 0.5x TBE, stained with SYBR Safe (cat. S3312, Thermo Fisher Scientific) at 1:10000 dilution in 0.5x TBE for 30 min, and imaged on an ImageQuant LAS4000.

DNase protection assays

To measure sensitivity to DNase digestion, the DNaseAlert kit (cat. 11-02-01-04, IDT) was used in combination with various concentrations of recombinant DNase I and antibody 35I9 in 100-μL reactions incubated at 37°C in a Tecan microplate-reader with initial measurement before addition of DNase I and subsequent measurements every 5 min after addition of DNase I (excitation 365 nm, emission 556 nm).

Animal models

All animal studies were approved by the Massachusetts Institute of Technology Committee

on Animal Care (MIT Protocols 042002323, 2301000462). Female BALB/c mice (6 to 10 weeks, Taconic Biosciences) were used for all healthy mice experiments. To generate the CT26 flank tumor model, female BALB/c mice (6 weeks, Taconic Biosciences) were injected subcutaneously with 2×10^6 CT26 cells resuspended in Opti-Mem (cat. 11058021, Thermo Fisher) into bilateral rear flanks. Tumors were measured every other day for 2 weeks, and tumor volumes were calculated by the modified ellipsoidal formula $V = 0.5 \times (l \times w^2)$, where l and w are the tumor length and width, respectively. To generate the transplantation model of lung metastasis, 1×10^5 Luc-MC26 cells in 100 μL of DPBS were injected intravenously (i.v.) into female BALB/c mice (6 weeks, Taconic Biosciences). Tumor growth was monitored by luminescence using the In Vivo Imaging System (IVIS, PerkinElmer) on days 6, 13, and 20 after tumor inoculation.

Blood collection

Retroorbital blood draws (70 μL in general, 35 μL for antibody pharmacokinetic study) were collected by means of nonheparinized capillary tubes from mice under isoflurane anesthesia, alternating between eyes for serial draws. Blood was immediately displaced from the capillary tube into 70 μL of 10-mM EDTA (cat. AM9260G, Thermo Fisher Scientific) in PBS. For terminal bleed samples, blood was collected through cardiac puncture into a syringe filled with 200 μL of 10-mM EDTA in PBS. Total volume was measured and an additional 10 mM of EDTA in PBS was added to reach a 1:1 ratio of blood to EDTA. Blood with EDTA was kept on ice and centrifuged within 90 min at 8000*g* for 5 min at 4°C. The plasma fraction was collected and stored at −80°C until further processing.

cfDNA extraction and quantification

Frozen plasma was thawed and centrifuged at 15,000*g* for 10 min to remove residual cells and debris. PBS was then added into plasma to make the total volume 2.1 ml for cfDNA extraction using the QIAsymphony Circulating DNA kit (cat:937556, Qiagen). The extracted cfDNA was quantified using a Taqman quantitative polymerase chain reaction (qPCR) assay targeting a locus in the mouse genome and then kept at 4°C until ready for further processing.

In vivo mononucleosome pharmacokinetic studies

SPE liposomes or sterile DPBS were administered i.v. into awake mice (50 to 300 mg/kg, 200 μl). 30 min after liposome injection, 1 μg of recombinant mononucleosomes carrying the Widom601 (W601) sequence (cat. 81070, ActiveMotif) suspended in 10 μL of DPBS were injected i.v. into anesthetized mice. In the study evaluating the percentage of exoNCP remaining at 60 min (n = 4 per group), 70 μL of blood was drawn retro-orbitally 1 and 60 min after mononucleosome injection. For the mAb pharmacokinetic assay, 10 to 20 ng of W601 sequence (cat. 81070, ActiveMotif) was combined with antibody in 200 μL of PBS. Engineered variants were produced in house; 35I9 was purchased from Abcam (cat. ab27156); mouse IgG2a control (clone 20102, cat. MAB003), anti-FcγRII/III (rat anti-mouse, clone 190909, cat. MAB1460), and anti-FcγRI (rat anti-mouse, clone 29035, cat. MAB2074) were purchased from R&D systems. 40 μg of anti-FcγRII/III and 20 μg of anti-FcγRI were used in FcγR-blocking conditions. Each mouse was anesthetized with inhaled isoflurane and injected i.v. with 200 μL of mixture. At 1 min after injection, 70 μL of blood was collected through a retro-orbital blood draw. Mice were allowed to recover after this and between subsequent blood draws (all 70 μL). Percentage of W601 remaining was calculated as the percentage of W601 remaining at 60 min relative to 1 min, as quantified using Taqman qPCR.

Plasma cfDNA concentration measurements following liposome administration

100 or 300 mg/kg SPE liposomes (200 μL in sterile DPBS) or DPBS were administered i.v. in awake mice (n = 3 mice per group). At 1 and 30 min and 1, 3, 5, and 24 hours after liposome administration, 70 μL of blood was collected retro-orbitally. Only two blood samples were collected from each mouse to prevent repeated sampling from the same capillary bed. Plasma cfDNA concentration was quantified as described above. Given that cfDNA recovery was highest 30 min and 3 hours after liposome administration for the 100 mg/kg and 300 mg/kg doses, respectively, we decided to sample blood 1 hour after liposome administration, which allowed us to compare results from animals treated with different liposome doses in our tumor models.

Antibody expression and purification

Desired Fc changes were introduced into the heavy-chain sequence (as determined by liquid chromatography–mass spectrometry de novo sequencing) and codon-optimized for expression in HEK293 cells. Gene blocks for the heavy and light chains were cloned into the same gWiz plasmid, separated by the T2A ribosome skipping sequence (*74*, *75*). Expi293F cells at a density of 3×10^6 cells/mL were transfected with 1 mg/L of culture of plasmid complexed with PEI Max 40K (cat. 24765-100, Polysciences) in a 1:2 plasmid:PEI w/w ratio in 40 mL of Opti-MEM (cat. 31985062, Thermo Fisher Scientific) per 1L culture. Flasks were kept in a shaking incubator (125 rpm) at 37°C and 8% CO_2. 24 hours after transfection, flasks were supplemented with glucose and valproic acid (cat. P4543, Millipore Sigma) to final concentrations of 0.4% v/v and 3 mM, respectively. Culture supernatant was harvested after 5 to 6 days and purified using Protein A affinity chromatography (AKTA, Cytiva), buffer exchanged into PBS, and sterile filtered and stored at −80°C.

Cell-line and buffy coat sequencing and fingerprint design

Genomic DNA (gDNA) was extracted from CT26 cells, Luc-MC26 cells, and Balb/c buffy coat, then sheared to 150 bp. gDNA libraries were prepared using the Kapa HyperPrep Library Construction kit (cat. KK8504, Roche Diagnostics). Whole-genome sequencing was performed to 30× coverage for CT26 and Luc-MC26, and 15× coverage for Balb/c buffy coat. Tumor fingerprints consisting of 98 and 1822 single-nucleotide variants (SNVs) were designed for CT26 and MC26 (data S1 and S2, respectively), as previously described (*9*).

Library construction, hybrid capture, and sequencing

cfDNA libraries were constructed using the Kapa Hyper Prep kit (cat: 07962363001, Roche) with custom dual index duplex UMI adapters (IDT), as previously described (*9*). Hybrid capture (HC) using tumor specific panels was performed using the xGen hybridization and wash kit (cat: 1080584, IDT) with xGen Universal blockers (cat: 1075476; IDT). For the ctDNA diagnostic test, libraries were pooled up to maximum 12-plex, with a library mass equivalent to 25× DNA mass into library construction for each sample, and a panel consisting of 120-bp long probes (IDT) targeting tumor-specific SNVs was applied. After the first round of HC, libraries were amplified by 16 cycles of PCR and then carried through a second HC. After the second round of HC, libraries were amplified through 8 to 16 cycles of PCR, quantified, and then pooled for sequencing (151 bp paired-end runs) with a targeted raw depth of 40,000×per site per 20 ng of DNA input. Sequencing data were processed by our duplex consensus calling pipeline as previously described, yielding measurements of the total number of mutant duplexes detected, the unique number of loci detected, and the tumor fractions (*9*). Relative duplex depth at each site was computed by subtracting mean overall depth for the library and then dividing by the standard deviation to obtain a relative duplex depth.

Assessing the performance of liposomal priming agent for tumor detection

Six days after tumor inoculation, mice bearing Luc-MC26 metastatic tumors were randomized into different treatment groups [100 mg/kg SPE liposomes (n = 12 mice) or PBS (n = 8 mice)] such that total burden was equivalent across different treatment groups (1.08 ± 0.5e7 photons/s for 100 mg/kg SPE liposomes versus 9.95 ± 5.2e6 photons/s for PBS). To determine how

our liposomal priming affected ctDNA performance at different tumor burdens, priming was performed 1, 2, and 3 weeks after tumor inoculation. At each timepoint, 70 μL of blood was sampled retro-orbitally from each mouse prior to treatment as an internal control. Subsequently, 100 mg/kg SPE liposomes (in 200 μL sterile DPBS) or sterile DPBS were administered i.v. into awake mice. 1 hour after treatment, 70 μL of blood was collected retro-orbitally from the contralateral eye, and a terminal bleed was then performed. cfDNA concentration measurement and ctDNA detection was performed on all samples as described above.

To calculate the sensitivity of the ctDNA test for tumor detection, mice were grouped as a function of tumor burden into those with small (total burden < 1.5e7 photons/s), medium (1.5e7 photons/s < total burden < 1.5e8 photons/s), and large (total burden > 1.5e8 photons/s) tumors. Retro-orbital plasma samples were classified as ctDNA positive if the number of distinct SNVs detected surpassed a given SNV threshold (between 2 and 10 SNVs, from lower to higher stringency of the test), and sensitivity was calculated as the % of samples that were ctDNA positive per group.

Assessing the performance of antibody priming agent for tumor detection

Between days 10 and 12 post–tumor inoculation, the performance of aST3 on ctDNA testing was assessed in Luc-MC26 tumor-bearing mice. As an internal control, 70 μL of blood was sampled retro-orbitally from each mouse prior to treatment. Subsequently, 4.0 mg/kg of aST3 (in 200 μL sterile DPBS) or 4.0 mg/kg of IgG2a isotype were administered into awake mice i.v. 2 hours after treatment, 70 μL of blood was collected retro-orbitally from the contralateral eye, and the remainder of the blood was collected by means of cardiac puncture. The 2 hour time point was chosen as it resulted in the highest endogenous cfDNA concentration in healthy mice after injection of aST3 (fig. S27). cfDNA concentration measurement and ctDNA detection was performed on all samples as described above.

ctDNA sensitivity estimation

To estimate sensitivity at smaller panel sizes, we used a bootstrap procedure down-sampling with replacement from our 1822-site panel to smaller panel sizes. Sensitivity at different detection thresholds was estimated as the fraction of mice that had mutant molecules detected at the given threshold. For each panel size and dose, 100 replicates were generated, and the mean sensitivity and standard error was computed. To estimate sensitivity at lower tumor fractions, we first confirmed that the distribution of mutant molecules (n_{ij}) and the distribution of the ratio of mutant molecules to total molecules (n_{ij}/t_{ij}) could be accurately recapitulated through a binomial sample $n_{ij} \sim \text{Binom}(t_{ij}, f_i)$, where n_{ij} is the number of mutant molecules at site j in sample i, t_{ij} is the number of total molecules at site j in sample i, and f_i is the global tumor fraction in sample i. To estimate sensitivity at lower tumor fractions, we then generated distributions of mutant molecules under lower f_i for each sample, also incorporating various panel sizes as above, and computed sensitivity for detection of mutant molecules under various detection thresholds. Sensitivity at each f_i, dose, and panel size was estimated by taking the mean and standard error from 100 replicates.

Statistical analysis

One-way analysis of variance (ANOVA) was used for statistical testing unless noted otherwise. A suite of scripts (Miredas) was used for calling mutations and creating metrics files (*9*, *15*). All other analysis was performed using GraphPad Prism v9, custom Python scripts, and R (v4.0.3) [code available on Zenodo (*76*)]. Detailed statistical information is provided in figure captions. For each animal experiment, mice were randomized such that groups would have comparable tumor burden. Investigators were not blinded to the groups or the treatments during the experiments.

Full materials and methods are available in the supplementary materials (*52*).

REFERENCES AND NOTES

1. Y. M. D. Lo, D. S. C. Han, P. Jiang, R. W. K. Chiu, Epigenetics, fragmentomics, and topology of cell-free DNA in liquid biopsies. *Science* **372**, eaaw3616 (2021). doi: 10.1126/science.aaw3616; pmid: 33833097
2. N. C. Rose *et al*., Screening for Fetal Chromosomal Abnormalities: ACOG Practice Bulletin, Number 226. *Obstet. Gynecol.* **136**, e48–e69 (2020). doi: 10.1097/AOG.0000000000004084; pmid: 32804883
3. T. A. Blauwkamp *et al*., Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. *Nat. Microbiol.* **4**, 663–674 (2019). doi: 10.1038/s41564-018-0349-6; pmid: 30742071
4. E. Heitzer, I. S. Haque, C. E. S. Roberts, M. R. Speicher, Current and future perspectives of liquid biopsies in genomics-driven oncology. *Nat. Rev. Genet.* **20**, 71–88 (2019). doi: 10.1038/s41576-018-0071-5; pmid: 30410101
5. I. De Vlaminck *et al*., Circulating Cell-Free DNA Enables Noninvasive Diagnosis of Heart Transplant Rejection. *Sci. Transl. Med.* **6**, 241ra77 (2014).
6. M. C. Liu *et al*., Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA.*Ann. Oncol.* **31**, 745–759 (2020). doi: 10.1016/j.annonc.2020.02.011; pmid: 33506766
7. A. Jamshidi *et al*., Evaluation of cell-free DNA approaches for multi-cancer early detection. *Cancer Cell* **40**, 1537–1549.e12 (2022). doi: 10.1016/j.ccell.2022.10.022; pmid: 36400018
8. M. G. O. Fernandes, N. Cruz-Martins, J. C. Machado, J. L. Costa, V. Hespanhol, The value of cell-free circulating tumour DNA profiling in advanced non-small cell lung cancer (NSCLC) management. *Cancer Cell Int.* **21**, 675 (2021). doi: 10.1186/s12935-021-02382-0; pmid: 34915883
9. H. A. Parsons *et al*., Sensitive Detection of Minimal Residual Disease in Patients Treated for Early-Stage Breast Cancer. *Clin. Cancer Res.* **26**, 2556–2564 (2020). doi: 10.1158/1078-0432.CCR-19-3005; pmid: 32170028
10. A. R. Parikh *et al*., Minimal Residual Disease Detection using a Plasma-only Circulating Tumor DNA Assay in Patients with Colorectal Cancer. *Clin. Cancer Res.* **27**, 5586–5594 (2021). doi: 10.1158/1078-0432.CCR-21-0410; pmid: 33926918
11. E. J. Moding, B. Y. Nabet, A. A. Alizadeh, M. Diehn, Detecting Liquid Remnants of Solid Tumors: Circulating Tumor DNA Minimal Residual Disease. *Cancer Discov.* **11**, 2968–2986 (2021). doi: 10.1158/2159-8290.CD-21-0634; pmid: 34785539
12. A. M. Newman *et al*., Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat. Biotechnol.* **34**, 547–555 (2016). doi: 10.1038/nbt.3520; pmid: 27018799
13. M. W. Schmitt *et al*., Detection of ultra-rare mutations by next-generation sequencing. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 14508–14513 (2012). doi: 10.1073/pnas.1208715109; pmid: 22853953
14. D. M. Kurtz *et al*., Enhanced detection of minimal residual disease by targeted sequencing of phased variants in circulating tumor DNA. *Nat. Biotechnol.* **39**, 1537–1547 (2021). doi: 10.1038/s41587-021-00981-w; pmid: 34294911
15. G. Gydush *et al*., Massively parallel enrichment of low-frequency alleles enables duplex sequencing at low depth. *Nat. Biomed. Eng.* **6**, 257–266 (2022). doi: 10.1038/s41551-022-00855-9; pmid: 35301450
16. A. Zviran *et al*., Genome-wide cell-free DNA mutational integration enables ultra-sensitive cancer monitoring. *Nat. Med.* **26**, 1114–1124 (2020). doi: 10.1038/s41591-020-0915-3; pmid: 32483360
17. L. Keller, Y. Belloum, H. Wikman, K. Pantel, Clinical relevance of blood-based ctDNA analysis: Mutation detection and beyond. *Br. J. Cancer* **124**, 345–358 (2021). doi: 10.1038/s41416-020-01047-5; pmid: 32968207
18. F. Chemi *et al*., cfDNA methylome profiling for detection and subtyping of small cell lung cancers. *Nat. Cancer* **3**, 1260–1270 (2022). doi: 10.1038/s43018-022-00415-9; pmid: 35941262
19. S. Cristiano *et al*., Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature* **570**, 385–389 (2019). doi: 10.1038/s41586-019-1272-6; pmid: 31142840
20. C. Fiala, E. P. Diamandis, Can a Broad Molecular Screen Based on Circulating Tumor DNA Aid in Early Cancer Detection? *J. Appl. Lab. Med.* **5**, 1372–1377 (2020). doi: 10.1093/jalm/jfaa138; pmid: 33057613
21. M. Ignatiadis, G. W. Sledge, S. S. Jeffrey, Liquid biopsy enters the clinic - implementation issues and future challenges. *Nat. Rev. Clin. Oncol.* **18**, 297–312 (2021). doi: 10.1038/s41571-020-00457-x; pmid: 33473219
22. A. Tivey, M. Church, D. Rothwell, C. Dive, N. Cook, Circulating tumour DNA - looking beyond the blood. *Nat. Rev. Clin. Oncol.* **19**, 600–612 (2022). doi: 10.1038/s41571-022-00660-y; pmid: 35915225
23. L. Zhu, A. Nazeri, C. P. Pacia, Y. Yue, H. Chen, Focused ultrasound for safe and effective release of brain tumor biomarkers into the peripheral circulation. *PLOS ONE* **15**, e0234182 (2020). doi: 10.1371/journal.pone.0234182; pmid: 32492056
24. J. M. Noh *et al*., Targeted Liquid Biopsy Using Irradiation to Facilitate the Release of Cell-Free DNA from a Spatially Aimed Tumor Tissue. *Cancer Res. Treat.* **54**, 40–53 (2022). doi: 10.4143/crt.2021.151; pmid: 34044476
25. A. Kustanovich, R. Schwartz, T. Peretz, A. Grinshpun, Life and death of circulating cell-free DNA. *Cancer Biol. Ther.* **20**, 1057–1067 (2019). doi: 10.1080/15384047.2019.1598759; pmid: 30990132
26. S. Khier, P. B. Gahan, Hepatic Clearance of Cell-Free DNA: Possible Impact on Early Metastasis Diagnosis. *Mol. Diagn. Ther.* **25**, 677–682 (2021). doi: 10.1007/s40291-021-00554-2; pmid: 34427906
27. D. S. C. Han, Y. M. D. Lo, The Nexus of cfDNA and Nuclease Biology. *Trends Genet.* **37**, 758–770 (2021). doi: 10.1016/j.tig.2021.04.005; pmid: 34006390
28. M. Germain *et al*., Priming the body to receive the therapeutic agent to redefine treatment benefit/risk profile. *Sci. Rep.* **8**, 4797 (2018). doi: 10.1038/s41598-018-23140-9; pmid: 29556068
29. T. Liu, H. Choi, R. Zhou, I.-W. Chen, RES blockade: A strategy for boosting efficiency of nanoparticle drug. *Nano Today* **10**, 11–21 (2015). doi: 10.1016/j.nantod.2014.12.003
30. B. Ouyang *et al*., The dose threshold for nanoparticle tumour delivery. *Nat. Mater.* **19**, 1362–1371 (2020). doi: 10.1038/s41563-020-0755-z; pmid: 32778816
31. N. R. M. Saunders *et al*., A Nanoprimer To Improve the Systemic Delivery of siRNA and mRNA. *Nano Lett.* **20**, 4264–4269 (2020). doi: 10.1021/acs.nanolett.0c00752; pmid: 32357299
32. M. Jain, N. Kamal, S. K. Batra, Engineering antibodies for clinical applications. *Trends Biotechnol.* **25**, 307–316 (2007). doi: 10.1016/j.tibtech.2007.05.001; pmid: 17512622
33. R.-M. Lu *et al*., Development of therapeutic antibodies for the treatment of diseases. *J. Biomed. Sci.* **27**, 1 (2020). doi: 10.1186/s12929-019-0592-z; pmid: 31894001

34. P. Liu, G. Chen, J. Zhang, A Review of Liposomes as a Drug Delivery System: Current Status of Approved Products, Regulatory Environments, and Future Perspectives. *Molecules* **27**, 1372 (2022). doi: 10.3390/molecules27041372; pmid: 35209162
35. C. D. Walkey, J. B. Olsen, H. Guo, A. Emili, W. C. W. Chan, Nanoparticle size and surface chemistry determine serum protein adsorption and macrophage uptake. *J. Am. Chem. Soc.* **134**, 2139–2147 (2012). doi: 10.1021/ja2084338; pmid: 22191645
36. Y.-N. Zhang, W. Poon, A. J. Tavares, I. D. McGilvray, W. C. W. Chan, Nanoparticle-liver interactions: Cellular uptake and hepatobiliary elimination. *J. Control. Release* **240**, 332–348 (2016). doi: 10.1016/j.jconrel.2016.01.020; pmid: 26774224
37. C. Rosales, E. Uribe-Querol, Phagocytosis: A Fundamental Process in Immunity. *BioMed Res. Int.* **2017**, 9042851 (2017). doi: 10.1155/2017/9042851; pmid: 28691037
38. P. T. Lowary, J. Widom, New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**, 19–42 (1998). doi: 10.1006/jmbi.1997.1494; pmid: 9514715
39. N. H. H. Heegaard, D. T. Olsen, K.-L. P. Larsen, Immuno-capillary electrophoresis for the characterization of a monoclonal antibody against DNA. *J. Chromatogr. A* **744**, 285–294 (1996). doi: 10.1016/0021-9673(96)00425-6; pmid: 8843677
40. E. Ben Chetrit, E. H. Dunsky, S. Wollner, D. Eilat, In vivo clearance and tissue uptake of an anti-DNA monoclonal antibody and its complexes with DNA. *Clin. Exp. Immunol.* **60**, 159–168 (1985). pmid: 3874013
41. I. Mellman, H. Plutner, Internalization and degradation of macrophage Fc receptors bound to polyvalent immune complexes. *J. Cell Biol.* **98**, 1170–1177 (1984). doi: 10.1083/jcb.98.4.1170; pmid: 6715404
42. F. Junker, J. Gordon, O. Qureshi, Fc Gamma Receptors and Their Role in Antigen Uptake, Presentation, and T Cell Activation. *Front. Immunol.* **11**, 1393 (2020). doi: 10.3389/fimmu.2020.01393; pmid: 32719679
43. R. Liu, R. J. Oldham, E. Teal, S. A. Beers, M. S. Cragg, Fc-Engineering for Modulated Effector Functions-Improving Antibodies for Cancer Treatment. *Antibodies (Basel)* **9**, 64 (2020). doi: 10.3390/antib9040064; pmid: 33212886
44. X. Wang, M. Mathieu, R. J. Brezski, IgG Fc engineering to modulate antibody effector functions. *Protein Cell* **9**, 63–73 (2018). doi: 10.1007/s13238-017-0473-8; pmid: 28986820
45. M. R. Walker, J. Lund, K. M. Thompson, R. Jefferis, Aglycosylation of human IgG1 and IgG3 monoclonal antibodies can eliminate recognition by human cells expressing Fc γ RI and/or Fc γ RII receptors. *Biochem. J.* **259**, 347–353 (1989). doi: 10.1042/bj2590347; pmid: 2524188
46. M. H. Tao, S. L. Morrison, Studies of aglycosylated chimeric mouse-human IgG. Role of carbohydrate in the structure and effector functions mediated by the human IgG constant region. *J. Immunol.* **143**, 2595–2601 (1989). doi: 10.4049/jimmunol.143.8.2595; pmid: 2507634
47. M. K. Leabman *et al*., Effects of altered FcγR binding on antibody pharmacokinetics in cynomolgus monkeys. *MAbs* **5**, 896–903 (2013). doi: 10.4161/mabs.26436; pmid: 24492343
48. T. Schlothauer *et al*., Novel human IgG1 and IgG4 Fc-engineered antibodies with completely abolished immune effector functions. *Protein Eng. Des. Sel.* **29**, 457–466 (2016). doi: 10.1093/protein/gzw040; pmid: 27578889
49. R. A. Clynes, T. L. Towers, L. G. Presta, J. V. Ravetch, Inhibitory Fc receptors modulate in vivo cytotoxicity against tumor targets. *Nat. Med.* **6**, 443–446 (2000). doi: 10.1038/74704; pmid: 10742152
50. L. Baudino *et al*., Crucial role of aspartic acid at position 265 in the CH2 domain for murine IgG2a and IgG2b Fc-associated effector functions. *J. Immunol.* **181**, 6664–6669 (2008). doi: 10.4049/jimmunol.181.9.6664; pmid: 18941257
51. S. W. Hosea, E. J. Brown, M. I. Hamburger, M. M. Frank, Opsonic requirements for intravascular clearance after splenectomy. *N. Engl. J. Med.* **304**, 245–250 (1981). doi: 10.1056/NEJM198101293040501; pmid: 7442756
52. Materials and methods are available as supplementary materials.
53. A. N. Ilinskaya, M. A. Dobrovolskaia, Nanoparticles and the blood coagulation system. Part I: Benefits of nanotechnology. *Nanomedicine (Lond.)* **8**, 773–784 (2013). doi: 10.2217/nnm.13.48; pmid: 23656264
54. J. Tie *et al*., Circulating Tumor DNA Analysis Guiding Adjuvant Therapy in Stage II Colon Cancer. *N. Engl. J. Med.* **386**, 2261–2272 (2022). doi: 10.1056/NEJMoa2200075; pmid: 35657320
55. P. M. Kasi *et al*., Tumor-informed assessment of molecular residual disease and its incorporation into practice for patients with early and advanced-stage colorectal cancer (CRC-MRD Consortia). *J. Clin. Oncol.* **38**, 4108 (2020). doi: 10.1200/JCO.2020.38.15_suppl.4108
56. S. Avanzini *et al*., A mathematical model of ctDNA shedding predicts tumor detection size. *Sci. Adv.* **6**, eabc4308 (2020). doi: 10.1126/sciadv.abc4308; pmid: 33310847
57. J. C. M. Wan *et al*., ctDNA monitoring using patient-specific sequencing and integration of variant reads. *Sci. Transl. Med.* **12**, eaaz8084 (2020). doi: 10.1126/scitranslmed.aaz8084; pmid: 32554709
58. B. R. McDonald *et al*., Personalized circulating tumor DNA analysis to detect residual disease after neoadjuvant therapy in breast cancer. *Sci. Transl. Med.* **11**, eaax7392 (2019). doi: 10.1126/scitranslmed.aax7392; pmid: 31391323
59. V. S. Nair *et al*., Genomic Profiling of Bronchoalveolar Lavage Fluid in Lung Cancer. *Cancer Res.* **82**, 2838–2847 (2022). doi: 10.1158/0008-5472.CAN-22-0554; pmid: 35748739
60. H. A. Parsons *et al*., Circulating tumor DNA association with residual cancer burden after neoadjuvant chemotherapy in triple-negative breast cancer in TBCRC 030. *Ann. Oncol.* **34**, 899–906 (2023). doi: 10.1016/j.annonc.2023.08.004; pmid: 37597579
61. X. Hou, T. Zaks, R. Langer, Y. Dong, Lipid nanoparticles for mRNA delivery. *Nat. Rev. Mater.* **6**, 1078–1094 (2021). doi: 10.1038/s41578-021-00358-0; pmid: 34394960
62. N. Boehnke *et al*., Massively parallel pooled screening reveals genomic determinants of nanoparticle-cell interactions. *Science* **377**, 384 (2022). doi: 10.1126/science.abm5551
63. C. S. Zent, M. R. Elliott, Maxed out macs: Physiologic cell clearance as a function of macrophage phagocytic capacity. *FEBS J.* **284**, 1021–1039 (2017). doi: 10.1111/febs.13961; pmid: 27863012
64. A. Hakkim *et al*., Impairment of neutrophil extracellular trap degradation is associated with lupus nephritis. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 9813–9818 (2010). doi: 10.1073/pnas.0909927107; pmid: 20439745
65. J. A. Chen *et al*., Sensitive detection of plasma/serum DNA in patients with systemic lupus erythematosus. *Autoimmunity* **40**, 307–310 (2007). doi: 10.1080/08916930701356317; pmid: 17516216
66. D. A. Isenberg, J. J. Manson, M. R. Ehrenstein, A. Rahman, Fifty years of anti-ds DNA antibodies: Are we approaching journey's end? *Rheumatology (Oxford)* **46**, 1052–1056 (2007). doi: 10.1093/rheumatology/kem112; pmid: 17500073
67. S. Yung, T. M. Chan, Anti-DNA antibodies in the pathogenesis of lupus nephritis—The emerging mechanisms. *Autoimmun. Rev.* **7**, 317–321 (2008). doi: 10.1016/j.autrev.2007.12.001; pmid: 18295737
68. K. Ohnishi *et al*., Comparison of pathogenic and non-pathogenic murine antibodies to DNA: Antigen binding and structural characteristics. *Int. Immunol.* **6**, 817–830 (1994). doi: 10.1093/intimm/6.6.817; pmid: 8086372
69. K. A. Fenton, B. Tømmerås, T. N. Marion, O. P. Rekvig, Pure anti-dsDNA mAbs need chromatin structures to promote glomerular mesangial deposits in BALB/c mice. *Autoimmunity* **43**, 179–188 (2010). doi: 10.3109/08916930903305633; pmid: 19835488
70. S. G. Thakku *et al*., Genome-wide tiled detection of circulating Mycobacterium tuberculosis cell-free DNA using Cas13. *Nat. Commun.* **14**, 1803 (2023). doi: 10.1038/s41467-023-37183-8; pmid: 37002219
71. T. M. Soelter, J. H. Whitlock, A. S. Williams, A. A. Hardigan, B. N. Lasseigne, Nucleic acid liquid biopsies in Alzheimer's disease: Current state, challenges, and opportunities. *Heliyon* **8**, e09239 (2022). doi: 10.1016/j.heliyon.2022.e09239; pmid: 35469332
72. I. A. Polina, D. V. Ilatovskaya, K. Y. DeLeon-Pennell, Cell free DNA as a diagnostic and prognostic marker for cardiovascular diseases. *Clin. Chim. Acta* **503**, 145–150 (2020). doi: 10.1016/j.cca.2020.01.013; pmid: 31978408
73. P. Bankhead *et al*., QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017). doi: 10.1038/s41598-017-17204-5; pmid: 29203879
74. P. Sharma *et al*., 2A peptides provide distinct solutions to driving stop-carry on translational recoding. *Nucleic Acids Res.* **40**, 3143–3151 (2012). doi: 10.1093/nar/gkr1176; pmid: 22140113
75. J. H. Kim *et al*., High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLOS ONE* **6**, e18556 (2011). doi: 10.1371/journal.pone.0018556; pmid: 21602908
76. S. Tabrizi, priming agents, Zenodo (2023); https://doi.org/10.5281/zenodo.10237042.

ACKNOWLEDGMENTS

We thank the Koch Institute's Robert A. Swanson (1969) Biotechnology Center for the technical support. Specifically, we thank K. Cormier and R. Bronson from the Koch Institute Histology Core for tissue processing and pathological assessment, respectively, and D. S. Yun from the Nanotechnology Materials Core (RRID:SCR_018674) for assistance on the TEM and cryo–transmission electron microscopy imaging. We thank G. Ha for advice on running Griffin software. We thank S. Cowles, E. Lutz, and K. D. Wittrup for advice and assistance with mammalian expression. We thank G. Gydush for his help setting up the data analysis pipeline leveraged for this work. We thank L. Gaffney for her assistance in designing the graphics for this work. **Funding:** This work was supported in part by a Cancer Center Support (core) grant P30-CA14051 from the National Cancer Institute, a Core Center grant P30-ES002109 from the National Institute of Environmental Health Sciences, the Koch Institute's Marble Center for Cancer Nanomedicine (S.N.B.), the Gerstner Family Foundation (V.A.A. and T.R.G.), the Koch Institute Frontier Research Program through the Casey and Family Foundation Cancer Research (J.C.L. and S.N.B.) the Virginia and D. K. Ludwig Fund for Cancer Research (S.N.B.), and the Bridge Project, a partnership between the Koch Institute for Integrative Cancer Research at MIT and the Dana-Farber/Harvard Cancer Center (J.C.L. and V.A.A.). C.M.-A. acknowledges support from a fellowship from La Caixa Foundation (ID 100010434). The fellowship code is LCF/BQ/AA19/11720039. C.M.-A. also acknowledges support from The Ludwig Center Fellowship at MIT's Koch Institute. S.T. acknowledges support from an ASCO Conquer Cancer Foundation Young Investigator Award (2021YIA-5688173400) and a Prostate Cancer Foundation Young Investigator Award (21YOUN20). S.P. acknowledges support from a T32 (T32HL116275). S.N.B. is a Howard Hughes Medical Institute investigator. **Author contributions:** Conceptualization: J.C.L., V.A.A., S.N.B., C.M.-A., S.T., and K.X.; Methodology: C.M.-A., S.T., K.X., T.B., S.S., A.C., S.P., Z.A., A.B., D.S., S.-T.W., S.R.-A., C.A.N., J.R., J.D.K., and A.P.A.; Investigation: C.M.-A., S.T., K.X., T.B., S.S., A.C., S.P., Z.A., A.B., D.S., S.-T.W., S.R.-A., C.A.N., J.R., J.D.K., and A.P.A.; Visualization: C.M.-A., S.T., K.X., and T.B.; Funding acquisition: J.C.L., V.A.A., S.N.B., and T.R.G.; Project administration: C.M.-A., S.T., and K.X.; Supervision: J.C.L., S.N.B., and V.A.A.; Writing – original draft: C.M.-A., S.T., and K.X.; Writing – review and editing: C.M.-A., S.T., K.X., A.P.A., H.F., J.C.L., S.N.B., and V.A.A. **Competing interests:** A patent application has been filed on this work: "METHODS AND COMPOSITIONS FOR INCREASING THE CONCENTRATION OF CELL FREE DNA" WO/2022/159910 (J.C.L., C.M.-A., S.N.B., S.T., V.A.A., and K.X.). T.R.G. holds equity in Sherlock Biosciences and is a paid adviser and equity holder in Anji Pharmaceuticals, Dewpoint Therapeutics, and Braidwell Inc., none of which are related to the present study. T.R.G. also receives research funding from Calico Life Sciences and Deerfield Management, both unrelated to the present study. J.C.L. has interests in Sunflower Therapeutics PBC, Honeycomb Biotechnologies, OneCyte Biotechnologies, QuantumCyte, and Repligen, which were not involved in this study. V.A.A. is a coinventor on a patent application (US 2023/0203568, pending) licensed to Exact Sciences and receives research funding from Exact Sciences. S.N.B. reports consulting roles and/or equity in Sunbird Bio, Satellite Bio, Catalio Capital, Port Therapeutics, Matrisome Bio, Xilio Therapeutics, Ochre Bio, Vertex Pharmaceuticals, Moderna, Johnson & Johnson, Ropirio Therapeutics, Danaher, and Owlstone Medical, which were not involved in this study. J.C.L.'s and S.N.B.'s interests were reviewed and managed under MIT's policies for potential conflicts of interest. T.R.G., J.C.L., V.A.A., and S.N.B. cofounded Amplifyer Bio based on this work. The remaining authors report no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. All materials used in this study are available from the authors upon request. A suite of scripts (Miredas) was used for analysis of ctDNA data (*9*, *15*). Other custom code is available on Zenodo (https://doi.org/10.5281/zenodo.10237042) (*76*). All sequencing data generated in this study have been deposited into SRA (access ID: PRJNA1037081).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adf2341
Materials and Methods
Figs. S1 to S34
References (*77*, *78*)
MDAR Reproducibility Checklist
Data S1 to S5

Submitted 17 October 2022; resubmitted 7 August 2023
Accepted 1 December 2023
10.1126/science.adf2341

RESEARCH ARTICLE SUMMARY

CORONAVIRUS

Persistent complement dysregulation with signs of thromboinflammation in active Long Covid

Carlo Cervia-Hasler, Sarah C. Brüningk, Tobias Hoch, Bowen Fan, Giulia Muzio, Ryan C. Thompson, Laura Ceglarek, Roman Meledin, Patrick Westermann, Marc Emmenegger, Patrick Taeschler, Yves Zurbuchen, Michele Pons, Dominik Menges, Tala Ballouz, Sara Cervia-Hasler, Sarah Adamo, Miriam Merad, Alexander W. Charney, Milo Puhan, Petter Brodin, Jakob Nilsson, Adriano Aguzzi, Miro E. Raeber, Christoph B. Messner, Noam D. Beckmann, Karsten Borgwardt, Onur Boyman*

INTRODUCTION: Acute infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) causes various clinical phenotypes, ranging from asymptomatic to life-threatening COVID-19. About 5% of all infected individuals do not recover from acute disease but develop long-term complications, called Long Covid. Current hypotheses on factors contributing to Long Covid include tissue damage, viral reservoirs, autoimmunity, and persistent inflammation. There are currently no diagnostic tests or therapeutic solutions for affected patients.

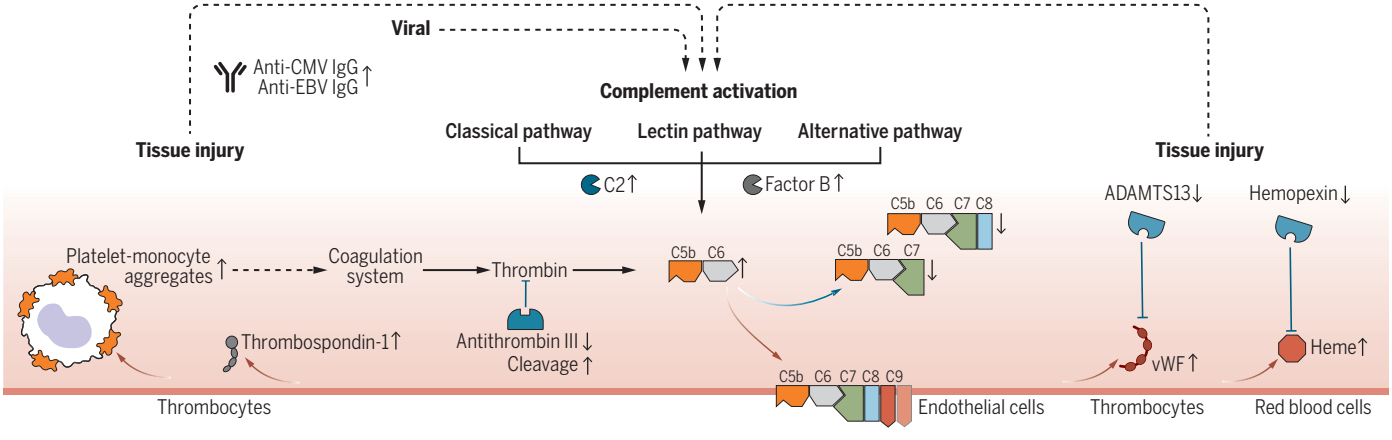
RATIONALE: We followed 39 healthy controls and 113 COVID-19 patients for up to 1 year after initial confirmation of acute SARS-CoV-2 infection to identify biomarkers associated with Long Covid. At 6-month follow-up, 40 patients had Long Covid symptoms. Repeated clinical assessments were paired with blood draws, resulting in a total of 268 longitudinal blood samples. We measured >6500 proteins in serum by proteomics. Top candidate biomarkers were identified using computational tools and further evaluated experimentally.

RESULTS: Long Covid patients exhibited increased complement activation during acute disease, which also persisted at 6-month follow-up. The complement system is part of the innate immune system and contributes to immunity and homeostasis by targeting pathogens and damaged cells, among other functions. Interestingly, blood complement levels normalized in Long Covid patients recovering before their 6-month follow-up. The complement system can be activated by various triggers, resulting in formation of the terminal complement complex (TCC), made of the complement components C5b-9. These complexes can integrate into cell membranes and induce cell activation or lysis. Long Covid patients showed imbalanced TCC formation, marked by increased soluble C5bC6 complexes and decreased levels of C7-containing TCC formations that can incorporate into cell membranes. This suggested increased membrane insertion of TCCs in Long Covid patients, contributing to tissue damage. Accordingly, Long Covid patients showed elevated tissue injury markers in blood and a thromboinflammatory signature, characterized by markers of endothelial activat such as von Willebrand factor (vWF), and blood cell lysis. Low antithrombin III levels in Long Covid patients were accompanied by signs of increased cleavage by thrombin, a driver of TCC formation. Furthermore, Long Covid patients had elevated platelet activation markers and monocyte–platelet aggregates at 6-month follow-up, particularly in cases where Long Covid persisted for 12 months or more. These patients also showed signs of antibody-mediated activation of the classical complement pathway, which was associated with increased anti-CMV (cytomegalovirus, also known as human herpesvirus 5) and anti-EBV (Epstein-Barr virus) immunoglobulin G (IgG) antibody levels.

CONCLUSION: Our data suggest that active Long Covid is accompanied by a blood protein signature marked by increased complement activation and thromboinflammation, including activated platelets and markers of red blood cell lysis. Tissue injury may also be complement-mediated and, in turn, activate the complement system. Moreover, complement activation may be driven by antigen–antibody complexes, involving autoantibodies and antibodies against herpesviruses, as well as cross-talk with a dysregulated coagulation system. In addition to offering a basis for new diagnostic solutions, our work provides support for clinical research on complement modulators for patients suffering from Long Covid. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: onur.boyman@uzh.ch
Cite this article as C. Cervia-Hasler *et al.*, *Science* **383**, eadg7942 (2024). DOI: 10.1126/science.adg7942

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adg7942



Pathomechanistic model of Long Covid. Model of complement-mediated thromboinflammation, showing increased and decreased biomarkers (up arrows and down arrows, respectively) measured at 6-month follow-up in patients with persistent Long Covid symptoms compared with recovered COVID-19 patients and healthy controls. Measurements were done using proteomics, spectral flow cytometry, single-cell transcriptomics, high-throughput antibody measurements, and targeted assays. Red arrows mark activating protein interactions, and blue arrows mark inhibiting protein interactions. Dashed arrows connect changes in different biological pathways.

RESEARCH ARTICLE

CORONAVIRUS

Persistent complement dysregulation with signs of thromboinflammation in active Long Covid

Carlo Cervia-Hasler[1], Sarah C. Brüningk[2,3], Tobias Hoch[1], Bowen Fan[2,3], Giulia Muzio[2,3], Ryan C. Thompson[4,5,6], Laura Ceglarek[1], Roman Meledin[1], Patrick Westermann[7], Marc Emmenegger[8], Patrick Taeschler[1], Yves Zurbuchen[1], Michele Pons[1], Dominik Menges[9], Tala Ballouz[9], Sara Cervia-Hasler[1], Sarah Adamo[1], Miriam Merad[4], Alexander W. Charney[4,5,6], Milo Puhan[9], Petter Brodin[10,11], Jakob Nilsson[1], Adriano Aguzzi[8], Miro E. Raeber[1], Christoph B. Messner[7], Noam D. Beckmann[4,5,6,12], Karsten Borgwardt[2,3], Onur Boyman[1,13]*

Long Covid is a debilitating condition of unknown etiology. We performed multimodal proteomics analyses of blood serum from COVID-19 patients followed up to 12 months after confirmed severe acute respiratory syndrome coronavirus 2 infection. Analysis of >6500 proteins in 268 longitudinal samples revealed dysregulated activation of the complement system, an innate immune protection and homeostasis mechanism, in individuals experiencing Long Covid. Thus, active Long Covid was characterized by terminal complement system dysregulation and ongoing activation of the alternative and classical complement pathways, the latter associated with increased antibody titers against several herpesviruses possibly stimulating this pathway. Moreover, markers of hemolysis, tissue injury, platelet activation, and monocyte–platelet aggregates were increased in Long Covid. Machine learning confirmed complement and thromboinflammatory proteins as top biomarkers, warranting diagnostic and therapeutic interrogation of these systems.

Acute infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) causes a variety of clinical phenotypes ranging from asymptomatic to life-threatening COVID-19 (*1*). Continuous transmission of SARS-CoV-2 in a previously naïve population has been accompanied by mounting evidence that about 20% of patients diagnosed with COVID-19 and about 5% of all SARS-CoV-2–infected persons do not recover from acute disease but develop long-term complications, called Long Covid (*2*–*4*). Multifaceted symptoms, including fatigue, post-exertional malaise, and cognitive impairment (*5*), can resemble other postviral conditions and myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) (*6*). Current hypotheses on Long Covid include tissue damage, viral reservoirs, autoimmunity, and persistent inflammation (*7*).

The immune response to SARS-CoV-2 has been extensively studied and is initiated at first contact with the virus, leading to strong activation of the innate and adaptive immune systems (*8*–*12*). These include interferons and immunoglobulin G3 (IgG3), the latter an IgG subset involved in antiviral immune responses and considered a protective factor against Long Covid and ME/CFS (*13*–*15*). Acute SARS-CoV-2–related immunopathology has been associated with excess inflammation, complement activation, hypercoagulation, and vascular injury (*16*, *17*). We applied two high-throughput proteomics approaches to a prospective cohort that included healthy controls, Long Covid patients, and patients who fully recovered from mild or severe COVID-19. Based on longitudinal measurements of >6500 serum proteins, we found evidence of persistent complement-mediated immunopathology associated with thromboinflammation in individuals experiencing Long Covid.

Results

Cohort characteristics

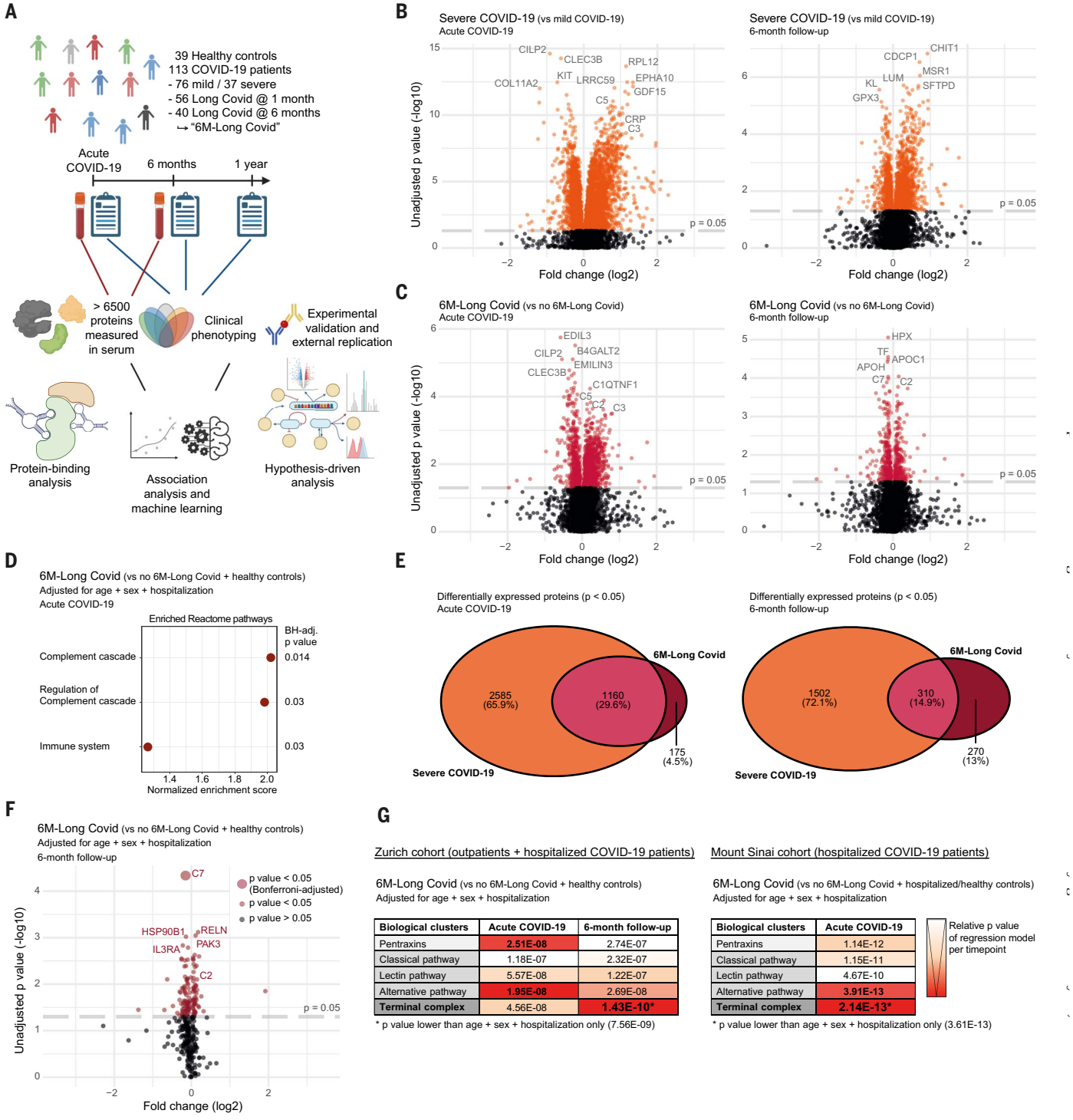
In this multicenter study, we followed 113 COVID-19 patients for up to 1 year. Study participants were included after confirmed acute SARS-CoV-2 infection, determined by a reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) test, and were seen for follow-up visits at 6 and 12 months after acute COVID-19. Moreover, 39 healthy adults were included as controls (fig. S1A and table S1). Of the 113 COVID-19 patients, 37 (33%) had a disease course classified as severe according to the World Health Organization (WHO) criteria (*1*, *18*). Sixty-five patients (58%) recovered fully, whereas 48 patients reported one or more COVID-19–related symptoms that persisted at 6-month follow-up (hereafter referred to as 6-month Long Covid). Eight patients indicated only changes in smell or taste; as isolated chemosensory disorders may reflect only a local pathology of the olfactory system (*19*), we excluded these individuals from our 6-month Long Covid definition. Therefore, we used samples from 40 patients to investigate systemic mechanisms underlying Long Covid. The 6-month Long Covid group showed a higher prevalence of severe acute COVID-19 compared with the 73 other COVID-19 patients (table S1). Sixteen patients without 6-month Long Covid experienced symptoms for longer than 1 month but recovered before the 6-month follow-up (Fig. 1A). In the 12-month follow-up, 22 of the 40 6-month Long Covid patients reported persisting symptoms, whereas 10 had recovered and 8 were lost to follow-up (fig. S1A).

Serum proteome of patients with persistent Long Covid at 6-month follow-up

Serum was collected from healthy controls and COVID-19 patients during acute COVID-19 and at 6-month follow-up. To analyze serum proteins, we used the SomaScan platform (*20*). This platform is based on synthetic aptamers that are short single-stranded nucleic acids with unique binding specificities to proteins, similar to antibodies. The platform included 7289 different aptamers targeting distinct human protein epitopes. As some proteins were targeted by several aptamers, a total of 6596 different human proteins were measured (Fig. 1A). We detected differences in serum protein levels between patients with severe COVID-19 and those with mild acute COVID-19 (Fig. 1B), as well as differences between 6-month Long Covid patients and patients without 6-month Long Covid (Fig. 1C), both during acute COVID-19 and at 6-month follow-up.

Severe COVID-19 is associated with autoantibody formation and persistent inflammation even after recovery (*21*–*24*) and thus can confound Long Covid–associated changes. We therefore searched for differences in the serum proteome that were distinct from those related to severe COVID-19. To this end, we tested all aptamer measurements for association with 6-month Long Covid, while considering potential confounders, such as patient age, sex, and hospitalization status (*13*, *25*). Statistical

[1]Department of Immunology, University Hospital Zurich, University of Zurich, 8091 Zurich, Switzerland. [2]Department of Biosystems Science and Engineering, ETH Zurich, 4058 Basel, Switzerland. [3]Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland. [4]Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [5]Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [6]Mount Sinai Clinical Intelligence Center, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [7]Precision Proteomics Center, Swiss Institute of Allergy and Asthma Research, University of Zurich, 7265 Davos, Switzerland. [8]Institute of Neuropathology, University Hospital Zurich, University of Zurich, 8091 Zurich, Switzerland. [9]Epidemiology, Biostatistics and Prevention Institute, University of Zurich, 8001 Zurich, Switzerland. [10]Unit for Clinical Pediatrics, Department of Women's and Children's Health, Karolinska Institute, 17165 Solna, Sweden. [11]Department of Immunology and Inflammation, Imperial College London, London W12 0NN, UK. [12]Division of Data Driven and Digital Medicine (D3M), Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [13]Faculty of Medicine and Faculty of Science, University of Zurich, 8006 Zurich, Switzerland.
*Corresponding author. Email: onur.boyman@uzh.ch



Biological clusters	Acute COVID-19	6-month follow-up
Pentraxins	**2.51E-08**	2.74E-07
Classical pathway	1.18E-07	2.32E-07
Lectin pathway	5.57E-08	1.22E-07
Alternative pathway	**1.95E-08**	2.69E-08
Terminal complex	4.56E-08	**1.43E-10***

* p value lower than age + sex + hospitalization only (7.56E-09)

Biological clusters	Acute COVID-19
Pentraxins	1.14E-12
Classical pathway	1.15E-11
Lectin pathway	4.67E-10
Alternative pathway	**3.91E-13**
Terminal complex	**2.14E-13***

* p value lower than age + sex + hospitalization only (3.61E-13)

Fig. 1. Differentially expressed proteins and pathways in Long Covid. (**A**) Study overview indicating the number of individuals used to generate data. [Created with BioRender.com] (**B** and **C**) Volcano plots showing differential serum levels of 7289 human protein epitopes in (B) severe versus mild COVID-19 patients and (C) patients with Long Covid (6-month Long Covid) versus no Long Covid at 6-month follow-up (no 6-month Long Covid), measured during acute COVID-19 (left) and at 6-month follow-up (right), using two-tailed *t* tests. (**D**) Significantly enriched Reactome pathways using 6408 different proteins measured during acute COVID-19, ranked by means of logistic regression. Benjamini-Hochberg (BH) adjusted. (**E**) Venn diagram showing overlap of proteins ($P < 0.05$) in (B) and (C). (**F**) Volcano plot showing differential expression of 331 epitopes [from 270 proteins from (E), right] selected at 6-month follow-up. Logistic regression of 6-month Long Covid (versus no 6-month Long Covid and healthy controls) adjusted for patient age, sex, and hospitalization. Bonferroni adjusted. (**G**) Association of protein clusters (table S2) with 6-month Long Covid (versus no 6-month Long Covid and healthy controls), including covariates patient age, sex, and hospitalization, in the Zurich cohort during acute COVID-19 and at 6-month follow-up (left) and the Mount Sinai cohort during acute COVID-19 (right; $n = 280$, including 145 6-month Long Covid patients). *P* values were calculated using logistic regression with likelihood-ratio test, colored according to their relative value at each time point (columns) and highlighted with asterisks when lower than a regression model based on covariates only.

modeling was applied to data of patients with 6-month Long Covid, patients without 6-month Long Covid, and healthy controls. The association between individual aptamer measurements, representing protein serum levels, and 6-month Long Covid was quantified for further analysis of enriched biological pathways (*26*). The pathways "complement cascade," "regulation of complement cascade," and "immune system" were significantly enriched in 6-month Long Covid patients during acute COVID-19 (Fig. 1D and fig. S1B) but not at 6-month follow-up (fig. S1C). Next, we compared proteins enriched in severe acute COVID-19 patients with proteins enriched in 6-month Long Covid patients and found a large overlap, particularly during acute COVID-19 (Fig. 1E and data S1 to S4). To identify biomarkers of Long Covid that are also applicable to patients with mild acute disease, we examined proteins that were specific to 6-month Long Covid patients at 6-month follow-up (Fig. 1E) and tested for their association with 6-month Long Covid, while considering patient age, sex, and hospitalization status during acute illness. We found complement component 7 (C7) measurements by one aptamer (seq.2888.49) to be significantly decreased in 6-month Long Covid patients (Fig. 1F).

C7 belongs to the complement system, which is part of the innate immune response and comprises a protein cascade targeting pathogens and cell debris, among other functions. The complement system can be activated by antigen–antibody complexes (classical pathway), bacterial sugars (lectin pathway), or spontaneously on cell surfaces (alternative pathway), requiring regulatory control on host cells. All three pathways lead to cleavage of C5 into C5b, which sequentially binds to C6, C7, C8, and C9, thus forming the terminal complement complex (or TCC; C5b-9). This complex can integrate into cell membranes and mediate cell activation or lysis. Pentraxins, a protein group involved in acute immune reactions, can also activate complement pathways (*27*, *28*).

Based on our finding of decreased C7, we analyzed association of 6-month Long Covid with protein clusters of selected biological pathways related to C7, including other TCC components and upstream complement activation pathways (table S2). Association of the selected pathways with 6-month Long Covid was analyzed, considering patient age, sex, and hospitalization, using logistic regression with likelihood-ratio test. During acute COVID-19, pentraxins and the alternative complement activation pathway were most strongly associated with 6-month Long Covid (Fig. 1G). At 6-month follow-up, TCC components were most differentially expressed in 6-month Long Covid patients, with C7 (measured by two different aptamers) representing the top associated protein cluster within the TCC (Fig. 1G and fig. S1D).

To confirm complement dysregulation in Long Covid, we analyzed SomaScan measurements in acute COVID-19 samples of an independent external cohort (the Mount Sinai cohort), including 198 patients hospitalized due to acute COVID-19, of which 145 (73.2%) developed 6-month Long Covid (fig. S1E and table S3). In the Mount Sinai cohort, TCC components were most differentially expressed in 6-month Long Covid patients, compared with COVID-19 patients without 6-month Long Covid, patients hospitalized for reasons unrelated to COVID-19, and healthy controls, with C5bC6 complexes representing the top associated protein cluster within the TCC during acute COVID-19 (Fig. 1G and fig. S1D). Altogether, we found evidence of a dysregulated complement system in 6-month Long Covid patients, particularly affecting the terminal pathway including C7.

Reduced levels of complement component 7 complexes in active Long Covid

C7 can be found as monomeric protein and as C7 complexes within the TCC formations C5b-7, C5b-8, and C5b-9 (Fig. 2A) (*27*). Thus, we determined C7 aptamer (seq.2888.49) specificity by an in-house enzyme-linked immunosorbent assay (ELISA) using C7 aptamers to capture different complement components. ELISA confirmed C7 aptamer binding to monomeric and complexed C7 but no other TCC components (fig. S2A). C7 aptamer–based pull-down and ELISA of monomeric and complexed TCC components in the presence of a polyanionic competitor, representing similar conditions as in the SomaScan assay, revealed strong aptamer binding to C7 complexes but not monomeric C7 (fig. S2, B and C), showing increased specificity of the C7 aptamer for complexed C7. Further analysis of C7 aptamer measurements by SomaScan revealed significantly reduced serum C7 complexes in patients with 6-month Long Covid at 6-month follow-up (Fig. 2B, left panel). Notably, C7 complexes were consistently reduced at 6-month follow-up, both in mild and severe cases with 6-month Long Covid (Fig. 2B, right panel). No significant differences could be detected when comparing healthy controls with mild and severe COVID-19 cases (Fig. 2C).

To assess whether C7 complexes were low only in Long Covid patients with active disease, we grouped patients without 6-month Long Covid into patients who recovered within 1 month (no Long Covid) and patients experiencing prolonged symptoms for more than 1 month but who fully recovered from Long Covid before 6-month follow-up (hereafter referred to as recovered Long Covid). Only 6-month Long Covid patients had decreased C7 complexes at 6-month follow-up, whereas recovered Long Covid patients had normal serum levels (Fig. 2D). Paired comparison of acute COVID-19 and 6-month follow-up samples revealed no significant changes of C7 levels with time (Fig. 2, E and F). Increasing patient age was not correlated with decreased C7 complexes (fig. S2D). Additionally, COVID-19 vaccination status did not influence the association of low C7 complexes with 6-month Long Covid (fig. S2E and table S4).

As monomeric C7 is 1000 times more abundant than C7 in TCCs in serum (*29*), we applied an ELISA using polyclonal C7-specific antibodies to quantify total C7. In contrast to the C7 aptamer, polyclonal anti-C7 antibodies target multiple epitopes on C7 and detect both monomeric and complexed C7. Total C7 levels were comparable in individuals with and without 6-month Long Covid, suggesting an isolated decrease in C7 complexes in 6-month Long Covid (Fig. 2G). A second aptamer targeting a different C7 epitope (seq.13731.14) was also unchanged in 6-month Long Covid patients compared with patients without 6-month Long Covid (fig. S2F) and correlated with total C7 levels measured by antibody-based ELISA (fig. S2G), thus most likely recognizing an epitope accessible on monomeric C7. The ratio of C7 complexes (measured by seq.2888.49) over total C7 (measured by seq.13731.14) was consistently decreased in 6-month Long Covid patients at 6-month follow-up (fig. S2H). We did not detect increased C7-specific autoantibody reactivity in 6-month Long Covid patients, suggesting no interference with C7 complex formation by autoantibodies (Fig. 2H and fig. S2, I and J). Thus, we observed low serum levels of C7 complexes in active Long Covid at 6-month follow-up, which normalized upon recovery.

Increased C5bC6 levels and complement activity in Long Covid

Formation of the TCC is initiated by association of C5b and C6 in stable bimolecular C5bC6 complexes. The recruitment of C7 to the C5bC6 complex enables subsequent binding to C8 and multiple C9 molecules, ultimately forming the potentially lytic membrane attack complex (MAC) or soluble C5b-9 (sC5b-9) (*27*). C7 is a central regulator of this process, as it can either scavenge C5bC6, forming soluble C7 complexes, or incorporate C5bC6 into cell membranes, leading to MAC formation and subsequent target-cell activation or lysis (*30*, *31*). In agreement with previous reports on increased complement activation in severe COVID-19 (*32*), we observed the highest C5bC6 levels in severe cases of acute COVID-19, which remained elevated at 6-month follow-up (Fig. 3A). In patients with 6-month Long Covid, we found higher C5bC6 levels both during acute COVID-19 and at 6-month follow-up compared with patients without 6-month Long Covid (Fig. 3B) and those who recovered before 6-month follow-up (fig. S3A). Paired comparison of acute COVID-19 and 6-month follow-up samples showed that in 6-month Long Covid patients, C5bC6 levels

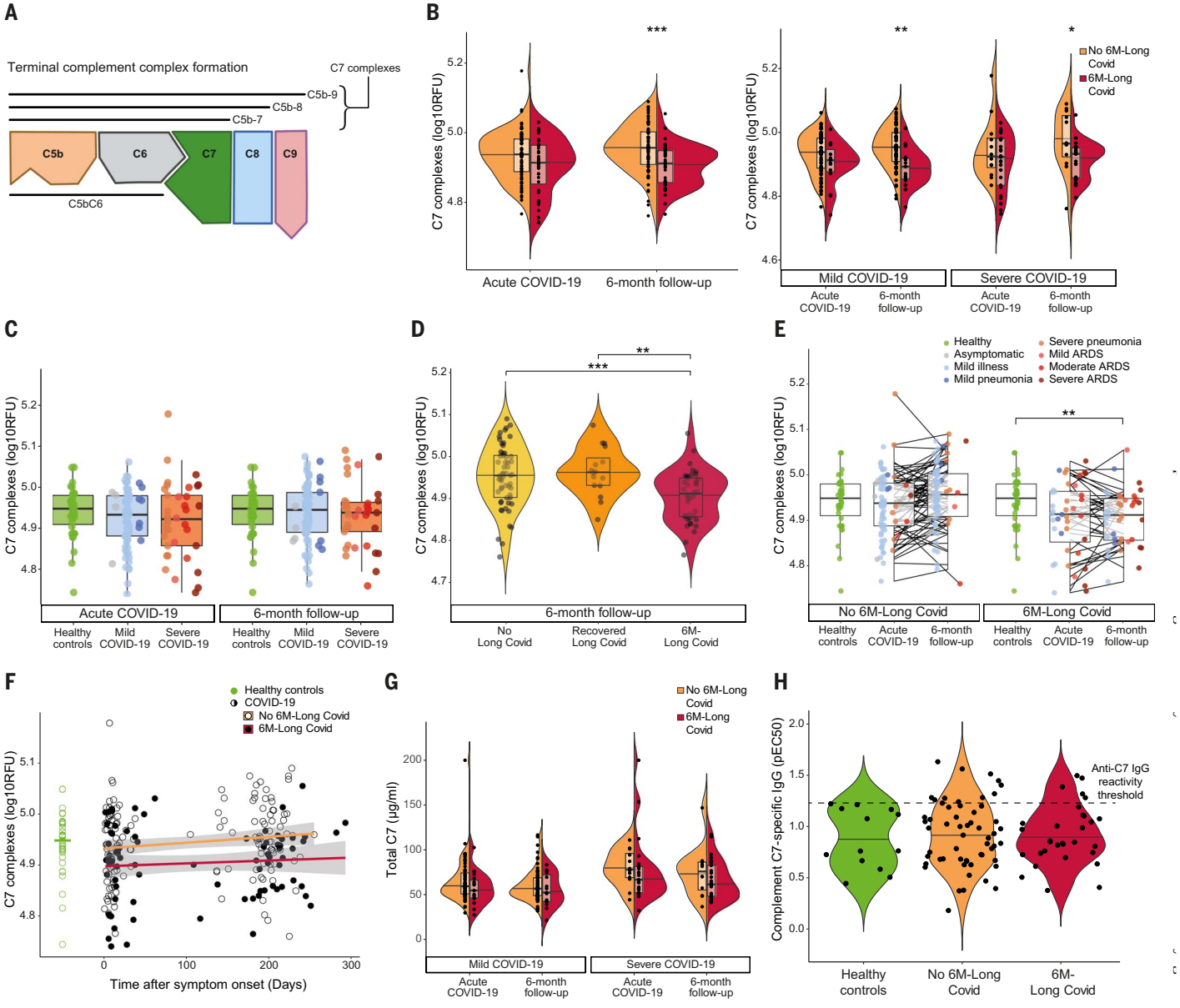


Fig. 2. Reduced levels of complement component 7 complexes in active Long Covid. (**A**) Schematic of terminal complement complexes. (**B**) Complement component 7 (C7) complexes in $\log_{10}$-transformed relative fluorescence units (log10RFU), measured by C7 aptamer seq.2888.49 in patients without (orange) or with 6-month Long Covid (red), during acute COVID-19 (left) and at 6-month follow-up (right). Comparison of all patients (left panel), and mild versus severe cases (right panel). (**C**) C7 complexes in healthy controls, mild or severe COVID-19 patients, at acute COVID-19 (left) and 6-month follow-up (right). (**D**) C7 complexes in patients with no Long Covid (symptom duration < 1 month), Long Covid recovered before 6-month follow-up, or 6-month Long Covid, at 6-month follow-up. (**E**) C7 complexes in patients without (left) or with 6-month Long Covid (right), at acute COVID-19 and 6-month follow-up (paired, Wilcoxon signed-rank test), compared with healthy controls. Lines connect corresponding patients. (**F**) C7 complexes in patients without (open circles) or with 6-month Long Covid (black dots), relative to time after symptom onset, including linear models. Healthy controls with median. (**G**) Total C7 (ELISA) in mild and severe COVID-19 patients without or with 6-month Long Covid, at acute COVID-19 and 6-month follow-up. (**H**) C7-specific IgG in healthy controls, patients without or with 6-month Long Covid at 12-month follow-up (n = 101), with median. Dots represent individual patients. Two-sided Wilcoxon test, unless otherwise specified. *P < 0.05, **P < 0.01, ***P < 0.001.

were highest during acute COVID-19 and remained elevated over time (Fig. 3C). ELISA measurements of sC5b-9 complexes, detecting a neoepitope of the fully assembled TCC, showed increasing concentrations with acute COVID-19 severity, but no differences at 6-month follow-up (Fig. 3D). Long Covid patients who recovered before 6-month follow-up and 6-month Long Covid patients had elevated sC5b-9 levels during the initial acute COVID-19 phase, whereas sC5b-9 levels were unchanged at 6-month follow-up (Fig. 3E and fig. S3B). In contrast, complement activity, determined by a CH50 equivalent assay measuring total sC5b-9 formation upon in vitro complement activation, revealed increased lytic activity, both during acute COVID-19 and at 6-month follow-up, in 6-month Long Covid patients but not in recovered Long Covid patients (Fig. 3, F and G, and fig. S3C).

To determine which TCC components were associated with increased sC5b-9 formation upon complement activation, we applied statistical modeling (linear mixed-effects model),

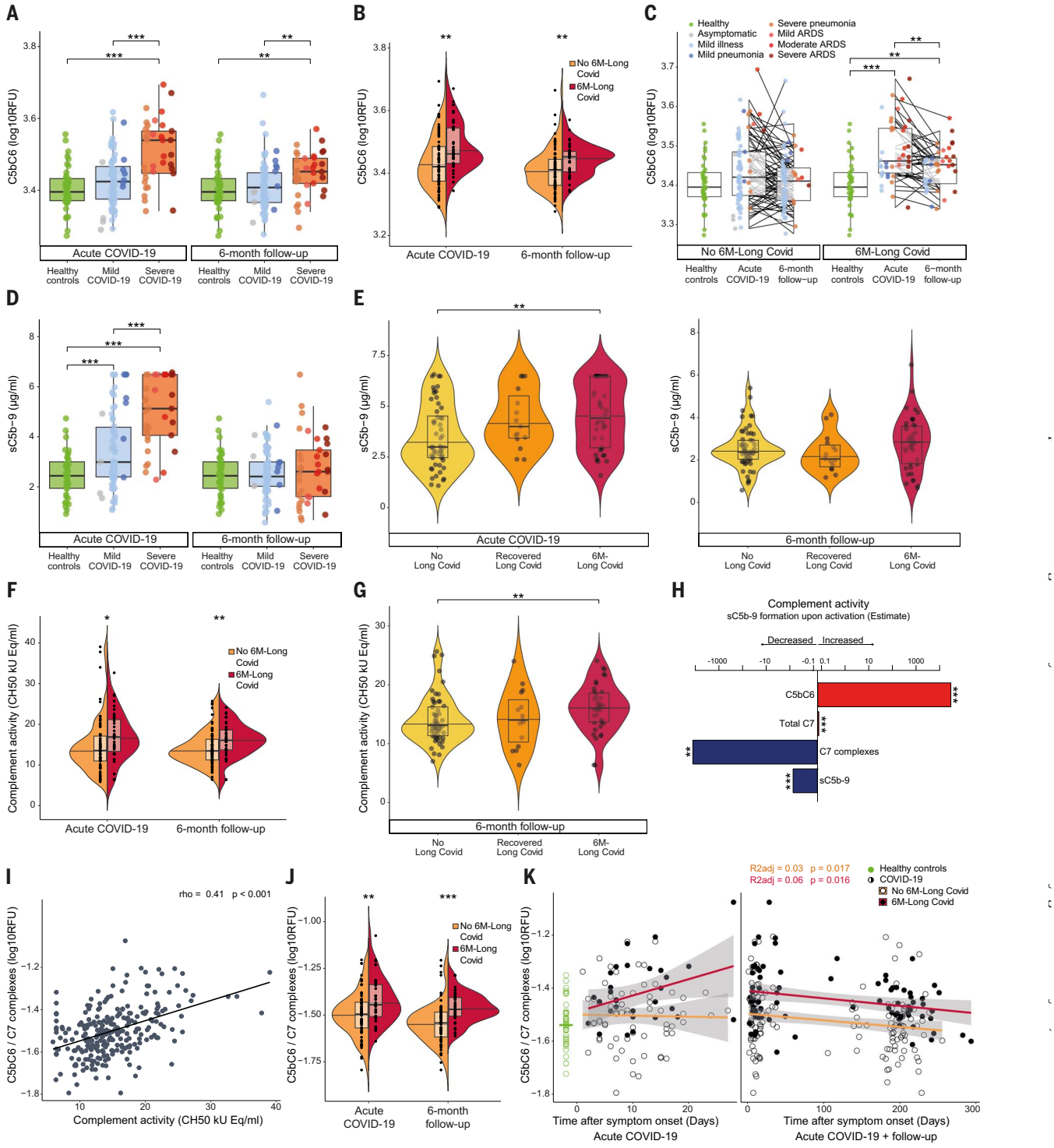


Fig. 3. Increased complement activity in Long Covid. (**A**) Complement C5bC6 complex (C5bC6) in healthy controls and mild and severe COVID-19 patients, at acute COVID-19 and 6-month follow-up. (**B** and **C**) C5bC6 in patients without or with 6-month Long Covid, at acute COVID-19 and 6-month follow-up (B), (paired test) compared to healthy controls (C). Lines connect corresponding patients. (**D** and **E**) Soluble C5b-9 (sC5b-9; ELISA) in healthy controls and mild or severe COVID-19 patients (D), and patients with no Long Covid, recovered Long Covid, or 6-month Long Covid (E), at acute COVID-19 and 6-month follow-up. (**F** and **G**) Complement activity in patients without or with 6-month Long Covid (F) and patients with no Long Covid, recovered Long Covid, or 6-month Long Covid (G), at acute COVID-19 and 6-month follow-up. (**H**) Linear mixed-effects model of complement activity as a function of complement levels, adjusted for age, sex, severe COVID-19, 6-month Long Covid, and sampling time point, using acute COVID-19 and 6-month follow-up data. (**I**) Spearman correlation of C5bC6/C7 complex ratio and complement activity. (**J**) C5bC6/C7 complex ratio in patients without or with 6-month Long Covid, at acute COVID-19 and 6-month follow-up. (**K**) C5bC6/C7 complex ratio in patients without or with 6-month Long Covid, relative to time, with linear models. Healthy controls with median (left). Dots represent individual patients. Two-sided Wilcoxon test, unless specified otherwise. *P < 0.05, **P < 0.01, ***P < 0.001.

considering patient age, sex, COVID-19 severity, 6-month Long Covid status, and sampling time point (Fig. 3H and fig. S3D). Combining acute COVID-19 and 6-month follow-up data, we found increased complement activity in COVID-19 patients to be associated with high levels of early TCC components, including C5bC6 and total C7, as well as low levels of late TCC formations, including C7 complexes and sC5b-9 (Fig. 3H). As C5bC6 and C7 complexes showed a negative correlation (fig. S3E), we calculated the C5bC6/C7 complex ratio, which strongly correlated with complement activity in study participants (Fig. 3I). The C5bC6/C7 complex ratio was increased in 6-month Long Covid patients, both during acute COVID-19 and 6 months later (Fig. 3J), with a tendency toward higher levels during acute COVID-19, followed by a gradual decrease with time (Fig. 3K), which was not influenced by patient age (fig. S3F). C7 complexes and the C5bC6/C7 complex ratio were similarly reduced at 6-month follow-up in all patients with 6-month Long Covid, regardless of whether they progressed to 12-month Long Covid (fig. S3G) and independent of frequency or quality of Long Covid symptoms (fig. S3, H to K). Collectively, increased complement activity in Long Covid patients at 6-month follow-up was marked by elevated early TCC formations, whereas reduced late TCC formations suggested membrane insertion of TCCs.

Association of classical and alternative complement activation with Long Covid

The classical and lectin complement activation pathways converge in activation of complement C2 and C4, followed by conversion of C3 to active C3b, thus initiating the assembly of C5 convertase, which results in TCC formation (*27*). Factor B is a central component of the alternative pathway. Moreover, C5 can be activated directly by thrombin by cross-talk with the blood coagulation system and initiate TCC formation (*33*, *34*) (Fig. 4A). We therefore investigated whether changes to the upstream complement activation pathways were associated with Long Covid. We detected increased C2 levels in 6-month Long Covid patients during acute COVID-19 and at 6-month follow-up (Fig. 4B), indicating classical or lectin pathway activation. ELISA measurement of the C4 cleavage product, C4d, revealed transiently increased C4d during acute COVID-19 but no differences based on Long Covid history (Fig. 4C and fig. S4A). Conversely, mannose-binding lectin (MBL), a central component of the lectin pathway, was unaltered between patients with and without 6-month Long Covid at 6-month follow-up (fig. S4B).

During acute COVID-19, factor B levels were increased in 6-month Long Covid patients compared with patients without Long Covid and healthy controls (Fig. 4D and fig. S4C). ELISA measurement of the cleavage product, factor Ba, confirmed increased levels in Long Covid patients during acute COVID-19 (fig. S4, C and D). At 6-month follow-up, factor Ba levels were persistently elevated in 6-month Long Covid patients compared with patients without Long Covid and healthy controls (Fig. 4E and fig. S4E). Complement C3d, the final degradation product of C3, was also elevated in 6-month Long Covid patients at 6-month follow-up compared with healthy controls (Fig. 4F). We analyzed serum levels of the thrombin inhibitor antithrombin III and found persistently low levels in 6-month Long Covid patients during acute COVID-19 and at 6-month follow-up (Fig. 4G and fig. S4E). In a subset of the Mount Sinai cohort (n = 21), we consistently found increased complement C3d and decreased antithrombin III in 6-month Long Covid patients at 3-month follow-up compared with patients without 6-month Long Covid, patients hospitalized for reasons unrelated to COVID-19, and healthy controls (Fig. 4H; fig. S4, F and G; and table S3).

To validate complement system dysregulation further, we performed mass spectrometry on 6-month follow-up serum samples of COVID-19 patients in our cohort. Complement components were among the top differentially abundant proteins in 6-month Long Covid patients compared with patients without 6-month Long Covid (Fig. 4I). Consistent with our previous analyses, complement C5 and factor B were increased and antithrombin III was decreased in 6-month Long Covid patients (Fig. 4J). In 6-month Long Covid patients, C5 was the only TCC component with significantly altered total levels (fig. S4H). Analysis of additional complement components showed increased complement regulatory proteins factor I, factor H, and C4-binding protein beta (C4BPB) and confirmed increased C2 in 6-month Long Covid patients (Fig. 4K). Interestingly, C2 was most elevated in 6-month Long Covid patients experiencing fatigue (fig. S4I) and 6-month Long Covid patients progressing to 12-month Long Covid (Fig. 4L) but showed no significant association with the number of 6-month Long Covid symptoms (fig. S4J). Instead, C2 highly correlated with C4BPB serum levels (fig. S4J), thus possibly explaining the high C2 levels in the presence of normal C4d levels. C2 levels measured by mass spectrometry at 6-month follow-up distinguished between patients developing 12-month Long Covid and those recovering before 12-month follow-up, as quantified by an area under the curve (AUC) of 0.81 (Fig. 4M). Other proteins of the complement system, such as C3 desArg, C5a, clusterin, vitronectin, and the pentraxins CRP, pentraxin 3 (PTX3), and serum amyloid A4 measured by SomaScan, mass spectrometry, or ELISA were unchanged (fig. S4K).

Analysis of antithrombin III peptides by mass spectrometry showed a decrease of peptide intensities after the reactive site with thrombin in 6-month Long Covid patients (fig. S4L). The ratio of peptide intensities on either side of this cleavage site was increased in 6-month Long Covid patients (Fig. 4N), suggesting increased cleavage of antithrombin III (*35*). Peptides at a heparin-binding region of antithrombin III were also decreased, accompanied by a decrease of the potential binding partner heparan sulfate proteoglycan 2 (HSPG2; fig. S4, L and M) (*35*–*37*). Analysis of other components of the coagulation system showed increased coagulation factor XI, fibrinogen beta, protein C, and heparin cofactor II in 6-month Long Covid patients (Fig. 4O). Altogether we demonstrated, by two approaches, classical and alternative complement system activation as well as low antithrombin III levels and increased cleavage of antithrombin III at the thrombin-reactive site.

Signs of complement-mediated tissue injury in Long Covid

Downstream effects of pathological complement activation entail hemolysis; induction of thromboinflammatory responses, including platelet and endothelial activation; and changes in innate immune cells, such as neutrophils (*38*–*40*). Combining SomaScan and mass spectrometry proteomics, we assessed a total of 84 measurements of 49 different serum biomarkers in the above-mentioned cell types (data S5) (*41*–*46*). Statistical modeling identified eight measurements associated with 6-month Long Covid, including decreased hemopexin (three times), ICAM-1, and S100-A8/A9, as well as increased thrombospondin-1 (TSP-1; twice) and von Willebrand factor (vWF) (Fig. 5A). ICAM-1 levels were persistently low during acute COVID-19 and at 6-month follow-up in 6-month Long Covid, suggesting low baseline levels (fig. S5A). Low levels of hemopexin, a marker of increased heme levels, were accompanied by normal hemoglobin and myoglobin levels (Fig. 5B and fig. S5B). An additional colorimetric assay was performed on 6-month follow-up samples of COVID-19 patients to quantify heme in serum, which revealed an association of increased heme with 6-month Long Covid and confirmed correlation of decreased hemopexin with increased heme (Fig. 5C and fig. S5C). Erythrocyte counts were unchanged, whereas haptoglobin was increased in 6-month Long Covid patients, along with slightly elevated interleukin-6 (IL-6) levels (Fig. 5B and fig. S5B). Measurements of transferrin, a marker of iron homeostasis, revealed a tendency toward reduced levels in 6-month Long Covid patients at 6-month follow-up (fig. S5B). Other markers of hemolysis and muscle cell injury as well as neutrophil and thrombocyte counts were unchanged overall (Fig. 5D and fig. S5, D and E).

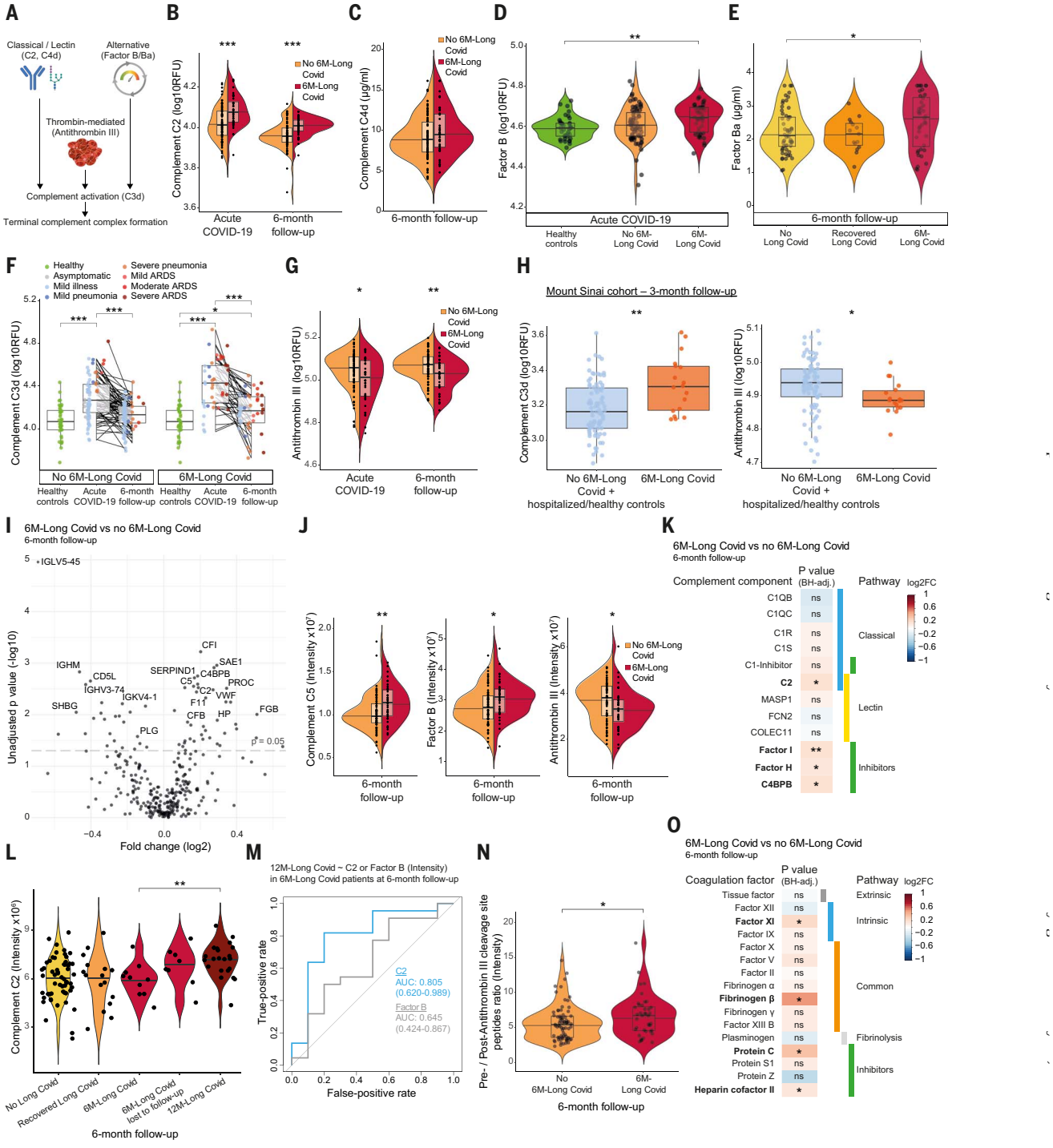


Fig. 4. Complement activation pathways at 6-month follow-up. (**A**) Schematic of pathways. [Created with BioRender.com] (**B** and **C**) C2 (B) and C4d (C) in patients without or with 6-month Long Covid, at acute COVID-19 or 6-month follow-up. (**D**) Factor B in healthy controls, patients without or with 6-month Long Covid, at acute COVID-19. (**E**) Factor Ba (ELISA) in patients with no Long Covid, recovered Long Covid, or 6-month Long Covid, at 6-month follow-up. (**F**) C3d in patients without or with 6-month Long Covid, at acute COVID-19 and 6-month follow-up (paired test), compared with healthy controls. Lines connect corresponding patients. (**G**) Antithrombin III in patients without or with 6-month Long Covid, at acute COVID-19 and 6-month follow-up. (**H**) C3d and antithrombin III in healthy and patient controls (blue; $n = 85$), and 6-month Long Covid patients (red; $n = 18$) of Mount Sinai cohort at 3-month follow-up. (**I**) Differential proteins (mass spectrometry) in patients with versus without 6-month Long Covid, at 6-month follow-up. (**J**) C5, factor B, and antithrombin III in patients without or with 6-month Long Covid, at 6-month follow-up. (**K**) Differential complement proteins (BH-adjusted) in patients with versus without 6-month Long Covid, at 6-month follow-up. (**L**) C2 in patients with no Long Covid, recovered Long Covid, 6-month Long Covid, 6-month Long Covid lost to follow-up, or 12-month Long Covid, at 6-month follow-up. (**M**) Area under the curve (AUC) for prediction of 12-month Long Covid as a function of C2 or factor B, in 6-month Long Covid patients at 6-month follow-up, using logistic regression. (**N**) Ratio of antithrombin III peptides intensity before and after thrombin-reactive site, in patients without or with 6-month Long Covid, at 6-month follow-up. (**O**) Differential coagulation proteins in patients with versus without 6-month Long Covid, at 6-month follow-up. Dots represent individual patients. Two-sided Wilcoxon test, unless specified otherwise. ns, nonsignificant; $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

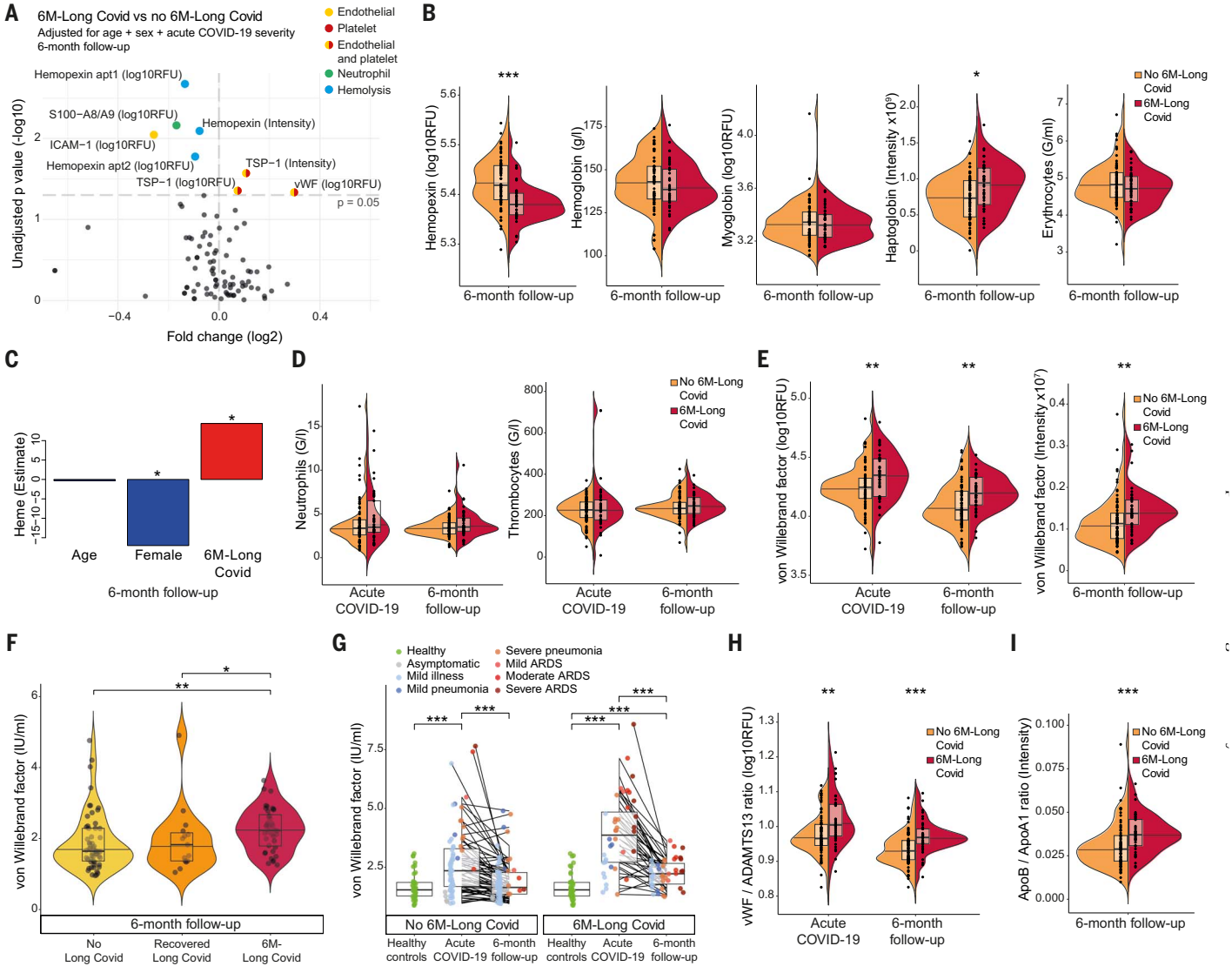


Fig. 5. Tissue injury markers in Long Covid. (**A**) Differential proteins [SomaScan aptamers (apt) and mass spectrometry; logistic regression] in patients with versus without 6-month Long Covid, at 6-month follow-up, including 84 measurements of 49 tissue injury markers (data S5). (**B**) Hemopexin, hemoglobin, myoglobin, haptoglobin, and erythrocyte counts in patients without or with 6-month Long Covid, at 6-month follow-up. (**C**) Linear model of heme (colorimetric assay) as a function of patient age, sex, and 6-month Long Covid, in COVID-19 patients at 6-month follow-up. (**D**) Neutrophil and thrombocyte counts in patients without or with 6-month Long Covid, at acute COVID-19 and 6-month follow-up. (**E**) Von Willebrand factor [vWF; SomaScan (left) and mass spectrometry (right)] in patients without or with 6-month Long Covid, at acute COVID-19 and 6-month follow-up. (**F**) vWF (ELISA) in patients with no Long Covid, recovered Long Covid, and 6-month Long Covid, at 6-month follow-up. (**G**) vWF in patients without or with 6-month Long Covid, at acute COVID-19 and 6-month follow-up (paired test), compared with healthy controls. Lines connect corresponding patients. (**H** and **I**) vWF/ADAMTS13 ratio (H) or ApoB/ApoA1 ratio [mass spectrometry; (I)] in patients without or with 6-month Long Covid, at acute COVID-19 or 6-month follow-up. Dots represent individual patients. Two-sided Wilcoxon test, unless specified otherwise. *P < 0.05, **P < 0.01, ***P < 0.001.

Serum levels of vWF, measured by SomaScan, mass spectrometry, and an ELISA calibrated to WHO standards, were increased in 6-month Long Covid patients both during acute COVID-19 and at 6-month follow-up but not in recovered Long Covid patients (Fig. 5, E to G, and fig. S5, F and G). The regulatory counterpart of vWF, ADAMTS13 (*47*), was decreased in 6-month Long Covid patients at 6-month follow-up (fig. S5H). The vWF/ADAMTS13 ratio, a marker of thromboinflammation and endothelial dysregulation (*45*, *47*, *48*), was increased in 6-month Long Covid patients both during acute COVID-19 and at 6-month follow-up (Fig. 5H). Both, vWF and the vWF/ADAMTS13 ratio showed positive correlation with complement activity (fig. S5I). Notably, at 6-month follow-up, vWF levels were lower in hospitalized patients who had received antiviral treatment with remdesivir during acute COVID-19 when compared with hospitalized patients who had not been treated with remdesivir (fig. S5J). Moreover, we detected increased coagulation factor VIII (FVIII) levels in 6-month Long Covid patients at 6-month follow-up (fig. S5K), which could be explained by vWF acting as a carrier of FVIII. In line with this thromboinflammatory signature, we also observed decreased platelet-activating factor acetylhydrolase (PAF-AH) and an increased ratio of low-density lipoprotein component apolipoprotein B (ApoB) to high-density lipoprotein component ApoA1 in 6-month Long Covid patients (Fig. 5I and fig. S5L), which is

an established cardiovascular risk factor (*49*). Antithrombin III and vWF levels were not affected by the presence of fatigue (fig. S5M). Collectively, Long Covid was characterized by elevated tissue injury markers, such as hemolysis and endothelial and platelet activation.

Increased monocyte–platelet aggregates in Long Covid patients

Next, we explored thromboinflammatory signatures on a cellular level. We performed spectral flow cytometry on peripheral blood mononuclear cells (PBMCs) from seven healthy controls, five COVID-19 patients without 6-month Long Covid, and ten 6-month Long Covid patients, of whom six progressed to 12-month Long Covid, all sampled at 6-month follow-up (Fig. 6A). These analyses were paired with single-cell transcriptomics of flow cytometry–sorted monocytes. Comparing cellular markers in these four groups, we found different phenotypes in cell clusters corresponding to classical monocytes (Fig. 6, B to D, and fig. S6A). We also measured CD41, a platelet marker usually absent on monocytes; thus, CD41-positive monocytes correspond to aggregates of monocytes with platelets (*50*). Based on CD41 surface abundance, classical monocytes were further manually subclustered into $CD41^{high}$, $CD41^{dim}$, and $CD41^{neg}$ monocytes (Fig. 6E). Frequencies of $CD41^{high}$ monocytes were highest in 12-month Long Covid patients (Fig. 6, B and F).

Flow cytometry–sorted monocytes of healthy controls were negative for CD41; however, adding platelets to these monocytes increased their CD41 abundance, regardless of activation by lipopolysaccharide (Fig. 6G and fig. S6, B to D), demonstrating that $CD41^{high}$ monocytes represented monocyte–platelet aggregates. On direct ex vivo analysis of cells, CD41 surface expression on monocytes was lowest in healthy controls and highest in Long Covid patients (Fig. 6H). Notably, $CD41^{high}$ monocytes showed different marker expression profiles than $CD41^{dim}$ and $CD41^{neg}$ monocytes, including increased CD55, also known as complement decay-accelerating factor (Fig. 6I). Single-cell transcriptomics of sorted monocytes (Fig. 6J) did not reveal a prothrombotic transcriptional signature of monocytes (Fig. 6K) but did confirm their absent *Cd41* mRNA expression (fig. S6E). We found low abundance of *IL1B* and *NR4A1*, particularly in classical monocytes, and increased expression of the interferon-induced transmembrane protein *IFITM3* (Fig. 6K and fig. S6, F to H). Altogether, we found monocyte–platelet aggregates in Long Covid patients at 6-month follow-up, which were highest in patients progressing to 12-month Long Covid.

Association of classical complement activation with IgG against herpesviruses

To identify antibodies potentially driving classical complement activation in 6-month Long Covid patients, we investigated autoantibodies and antiviral antibodies. Using a highly sensitive indirect immunofluorescence assay, we assessed antinuclear antibody (ANA) titers at 6-month follow-up and detected an increased prevalence of ANA positivity in 6-month Long Covid patients compared with patients without 6-month Long Covid (fig. S7A). ANA fluorescence patterns were mainly of nonspecific or fine speckled fluorescence across the nucleoplasm in 6-month Long Covid patients, whereas in ANA-positive patients without 6-month Long Covid, nucleolar patterns were more frequent (fig. S7B), which suggests different autoantibody specificities. Anti-chemokine antibodies, previously described to be protective in Long Covid, were measured in a subset of 88 COVID-19 patients. Statistical modeling revealed low anti-CXCL13 to be significantly associated with 6-month Long Covid (fig. S7, C and D), confirming published data (*51*).

High-throughput measurement of antiviral antibodies was performed by phage immunoprecipitation sequencing technology (VirScan) for a total of 87,890 epitopes in a subset of 57 COVID-19 patients during acute disease, including 22 that developed 6-month Long Covid, and 18 healthy controls (Fig. 7A and table S5). Although we found no overall increase in reactivity to viral epitopes in 6-month Long Covid patients compared with patients without 6-month Long Covid and healthy controls, IgG titers against herpesviruses 1 to 8 were increased proportionally (Fig. 7, A and B). On the epitope level, we observed low IgG against enterovirus B and high IgG against human herpesvirus 5 [also known as cytomegalovirus (CMV)] in 6-month Long Covid patients (Fig. 7, C and D). Statistical models showed associations of low anti-enterovirus IgG with patient age and of high anti-CMV IgG with 6-month Long Covid (fig. S7E).

On the basis of these data and previous reports on Epstein-Barr virus (EBV) reactivation (*52*), we measured serum titers of IgM and IgG against CMV and EBV at 6-month follow-up. Whereas antiviral IgM is increased during active viral replication, antiviral IgG indicates an immune response to past infection (*53*). Overall serum positivity for CMV- and EBV-specific IgG and, thus, prevalence of CMV or EBV infection did not differ between patients with and patients without 6-month Long Covid (Fig. 7E). We also found no difference in anti-CMV and anti-EBV IgM titers (Fig. 7F and fig. S7F). However, both anti-CMV and anti-EBV IgG were increased in 6-month Long Covid patients (Fig. 7F and fig. S7F). Statistical modeling confirmed the association of high anti-CMV and anti-EBV IgG titers with 6-month Long Covid (Fig. 7G). Notably, C2 levels were associated with anti-CMV IgG titers and 12-month Long Covid (Fig. 7H). We found no differences in serum IgG antibodies targeting four different SARS-CoV-2 spike epitopes in patients with and patients without 6-month Long Covid (Fig. 7I and fig. S7G) and no evidence of SARS-CoV-2 or herpesviral transcripts in monocytes (Fig. 7J). Analysis of previously reported inflammatory markers in Long Covid showed that interferon alpha 2 inversely correlated with C7 complexes, and IL-6 correlated with C5bC6 complexes in COVID-19 patients at 6-month follow-up (fig. S7H). Altogether, increased titers of antibodies against herpesviruses in Long Covid patients were associated with C2 of the classical complement pathway, suggesting that viral antigen–antibody immune complexes could be involved in activation of this pathway.

Serum proteomics for Long Covid diagnosis and prediction

To translate the observed differences into a predictive model for 6-month Long Covid, we included up to 61 uncorrelated protein epitope measurements of involved biological pathways (with absolute Spearman correlation coefficients below 0.3), two protein ratios, and information on medical history (14 variables) (data S8). For diagnostic purposes, we included only protein measurements and medical history obtained at 6-month follow-up to identify diagnostic biomarkers associated with 6-month Long Covid and, thus, normal in recovered patients and healthy controls. We applied a random forest classifier, a machine learning algorithm suitable for analyzing large datasets. Through iterative assessment of acute COVID-19 and 6-month follow-up data, the algorithm identified the most influential biomarkers for predicting Long Covid. The models were trained through fivefold nested cross-validation, with model hyperparameters being tuned on the training part of the data (Fig. 8A). Relevant data splits were stratified for acute COVID-19 severity. The resulting predictions performed well in the unseen test sets (corresponding to patient data not used for training the prediction model), as quantified by the area under the receiver operator characteristic (AUROC; ≥0.74 ± 0.10 for 6-month Long Covid) and relative average precision scores accounting for Long Covid prevalence and class imbalance (Fig. 8B and fig. S8A).

Studying the intricate interplay of features driving the prediction, we identified patient age, C5bC6/C7 complex ratio, vWF/ADAMTS13 ratio, and patient body mass index (BMI) as the driving variables (Fig. 8C). These four variables were sufficient to predict 12-month Long Covid (mean AUC of 0.81 ± 0.08) in a cohort of both healthy and COVID-19 subjects. Addition of the two protein ratios to age and BMI improved model performance in all scenarios (Fig. 8B and fig. S8, A and B). Inclusion of healthy controls in the group not developing Long Covid reduced the association of Long Covid with potentially preexisting risk factors

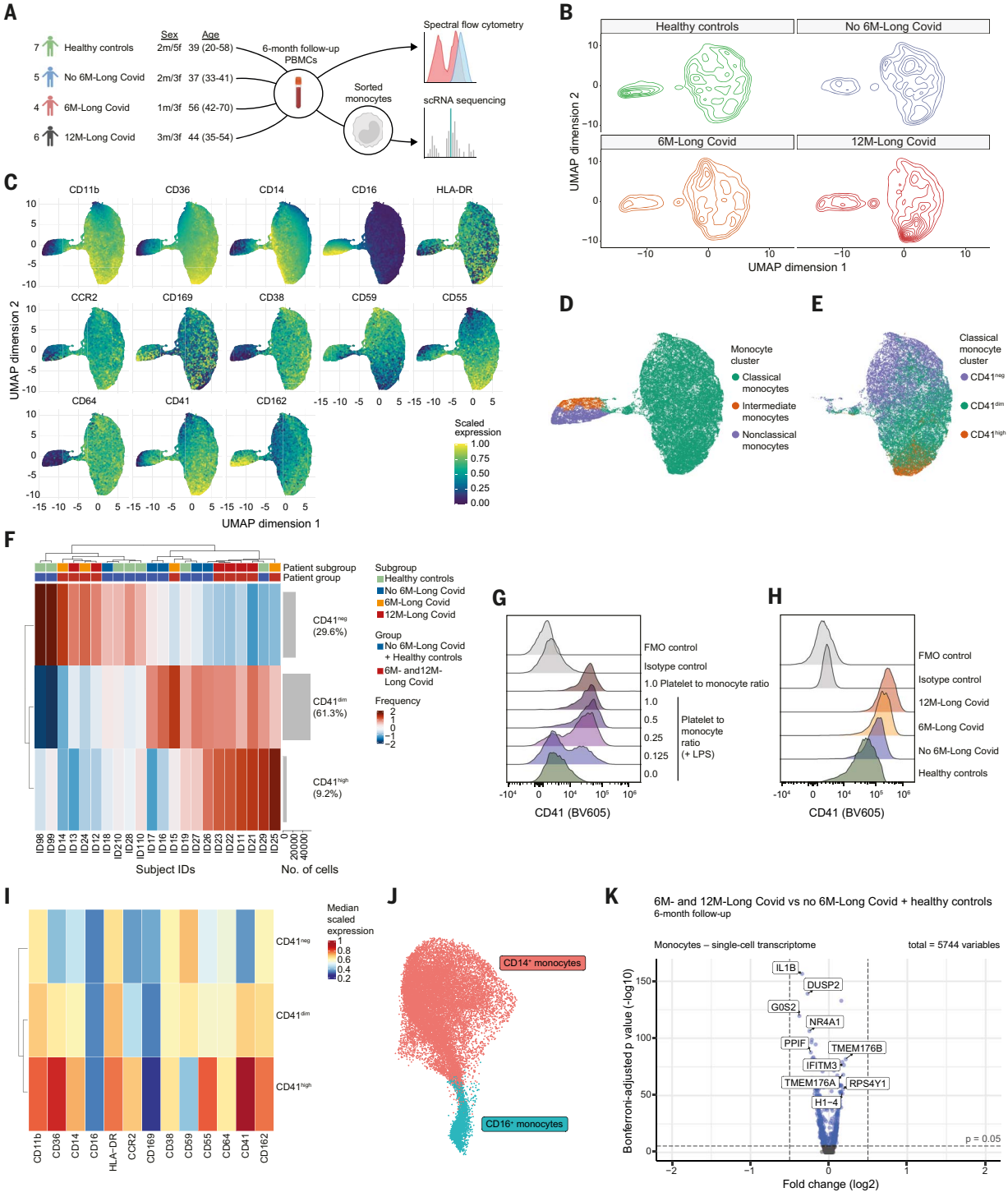


Fig. 6. Increased monocyte-platelet aggregates in Long Covid. (**A**) Overview of groups for spectral flow cytometry and sorting of PBMCs isolated at 6-month follow-up, for single-cell RNA (scRNA) sequencing of monocytes. Six-month Long Covid patients were subgrouped into patients with or without recovery before 12-month follow-up. For scRNA sequencing, five out of seven healthy controls were analyzed [sex: 1m/4f; age: 44 (23–65)]. A total of 19,247 monocytes were sequenced. [Created with BioRender.com] (**B**) Uniform manifold approximation and projection (UMAP) density plots of gated monocytes (spectral flow cytometry) of indicated groups. (**C**) UMAPs of scaled marker expression of monocytes. (**D** and **E**) UMAPs of all (D) and classical monocytes manually subclustered on the basis of CD41 expression profiles (E). (**F**) Relative (inside heatmap) and cumulative (gray bars) abundance of classical monocytes subclustered on the basis of CD41 expression, in indicated individual donors. Frequencies of CD41-subclustered classical monocytes indicate Z-score-normalized proportion of subclusters within total classical monocytes. (**G**) CD41 surface abundance on sorted monocytes incubated with indicated ratios of platelets, without or with lipopolysaccharide (LPS) prestimulation. (**H**) CD41 surface abundance on classical monocytes from indicated groups. (**I**) Scaled marker expression in CD41-subclustered classical monocytes. (**J**) UMAP of all monocytes (scRNA sequencing). (**K**) Volcano plot of differentially expressed RNA transcripts in monocytes of Long Covid patients (6- and 12-month Long Covid) versus controls (no 6-month Long Covid and healthy controls), at 6-month follow-up. Vertical dashed lines indicate $\log_2$ fold change of −0.5 and 0.5.

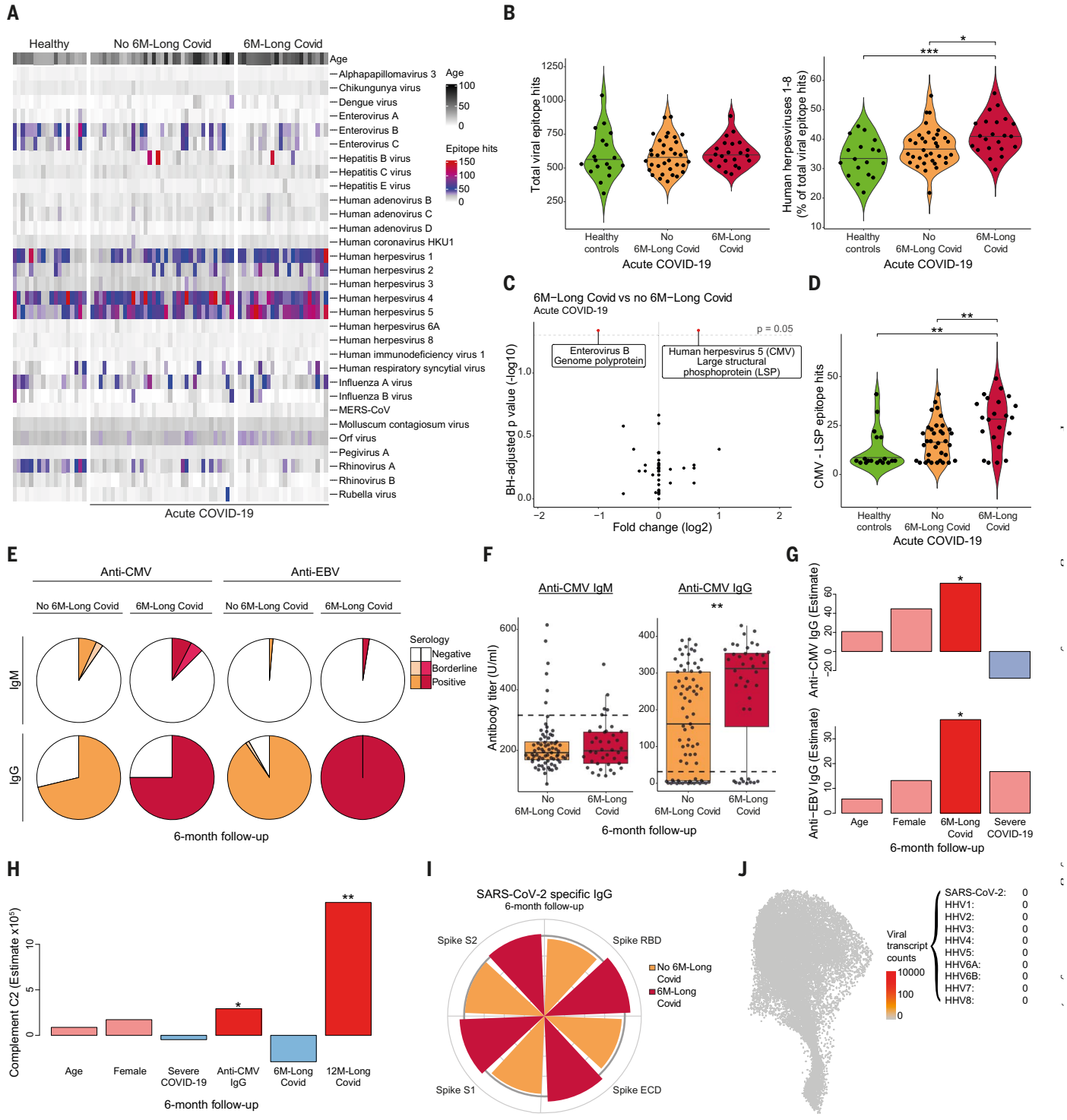


Fig. 7. Increased antibody reactivity against herpesviruses in Long Covid. (**A**) Antiviral IgGs (VirScan) in healthy controls (n = 18) and patients without (n = 35) or with 6-month Long Covid (n = 22), during acute COVID-19. Viruses with >300 epitope hits. (**B**) Total viral (left) and percentage of herpesviral (right) epitope hits in healthy controls and patients without or with 6-month Long Covid, with median. (**C**) Differential abundance of viral epitope-specific IgG in patients with versus without 6-month Long Covid, BH-adjusted. (**D**) CMV LSP epitope hits in healthy controls and patients without or with 6-month Long Covid. (**E**) Serum reactivity for anti-CMV (left) and anti-EBV (right), and IgM (top) and IgG (bottom) in patients without (orange) or with (red) 6-month Long Covid, at 6-month follow-up (ELISA; n = 113). (**F**) Anti-CMV IgM (left) and IgG (right) in patients without or with 6-month Long Covid. (**G** and **H**) Linear model coefficients for anti-CMV or anti-EBV IgG titers (G) and C2 (H). (**I**) Radar plot. Wedge sizes represent median IgG titers of patients without or with 6-month Long Covid, normalized to median titers of all patients, at 6-month follow-up. (**J**) UMAP of monocytes showing absence of intracellular herpesviral and SARS-CoV-2 transcripts (scRNA sequencing). ECD, extracellular domain; RBD, receptor binding domain. Dots represent individual patients. Two-sided Wilcoxon test, unless specified otherwise. *P < 0.05, **P < 0.01, ***P < 0.001.

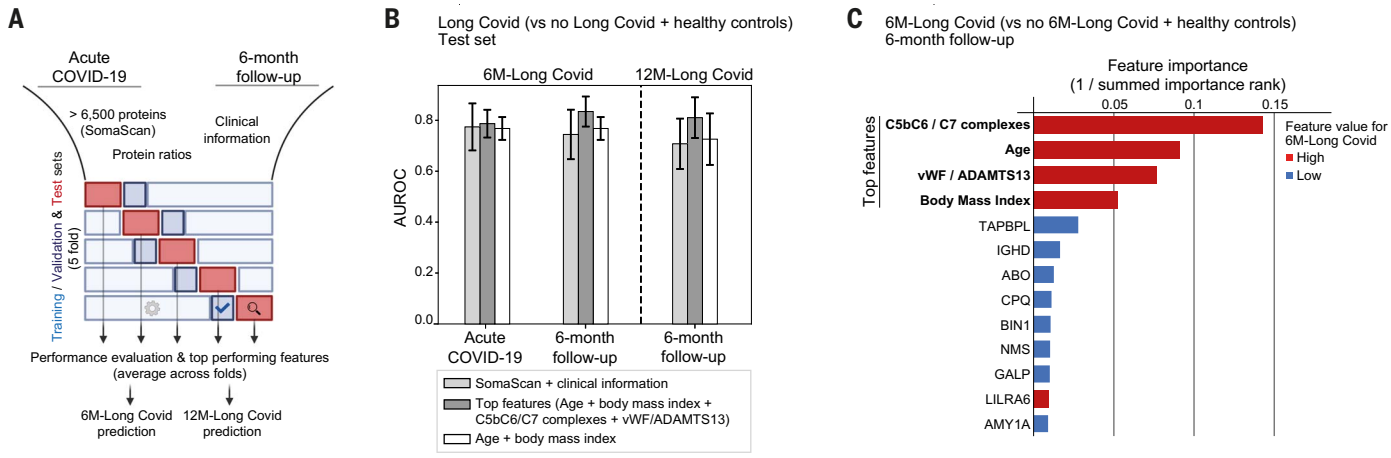


Fig. 8. Predictive and diagnostic models of Long Covid. (**A**) Overview of model training strategy by nested cross-validation. For each fold, models were trained on training portion of data with hyperparameters adjusted on the basis of validation set. Each trained model was evaluated on an unseen test set. [Created with BioRender.com] (**B**) Area under the receiver operating characteristics curve (AUROC) of a random forest predictor using up to 61 uncorrelated protein measurements, 2 protein ratios, 14 general clinical variables, top selected features of 6-month Long Covid prediction selected on training data [age, body mass index (BMI), C5bC6/C7 complexes, and vWF/ADAMTS13 ratio], or age and BMI alone. Input measurements obtained at acute COVID-19 or 6-month follow-up for prediction of 6-month Long Covid versus no 6-month Long Covid and healthy controls (n = 40, 73, and 39, respectively), and at 6-month follow-up for prediction of 12-month Long Covid versus no 12-month Long Covid and healthy controls (n = 22, 73, and 39, respectively; mean ± SD). (**C**) Top features of prediction of 6-month Long Covid versus no 6-month Long Covid and healthy controls, using 6-month follow-up measurements. Summed importance rank was calculated over all five cross-validation folds. Top selected features were identified by elbow analysis.

in healthy controls, thus increasing specificity of identified biomarkers for active Long Covid. We also tested the models upon exclusion of healthy controls and consistently confirmed the diagnostic and predictive potential of the protein biomarkers (fig. S8C). Collectively, machine learning algorithms independently identified complement and thromboinflammatory markers as top protein biomarkers of active Long Covid.

Discussion

New SARS-CoV-2 variants have confronted the world with repeated SARS-CoV-2 infection (*54*), contributing to a high prevalence of Long Covid (*55*). We identified common patterns in the serum proteome of Long Covid patients not recovered at 6 months after acute infection by studying a prospective cohort of mild and severe COVID-19 patients and healthy controls.

Analysis of >7000 protein measurements revealed the complement system as the top dysregulated biological pathway in Long Covid. At 6-month follow-up, late TCC formations, comprising the soluble C7 complexes C5b-7, C5b-8, and C5b-9, were reduced in individuals experiencing active Long Covid. These findings were paralleled by an increase of the early TCC formation C5bC6 and increased complement activity. The imbalance in TCC formations was best quantified by a C5bC6/C7 complex ratio, also identified by machine learning as the top predictive biomarker of Long Covid. Binding of C7 to the stable bimolecular complex C5bC6 enables the trimolecular C5b-7 complex to integrate into cell membranes (*30*). Our observation of increased complement activity in the presence of elevated C5bC6 levels and decreased soluble C7 complexes suggested increased membrane insertion of TCCs in active Long Covid.

Increased factor Ba levels suggested alternative complement pathway activation as a possible driver of TCC formation (*27*). Moreover, increased C2, reflecting classical complement pathway activation, was associated with Long Covid persistence. Our finding of increased anti-CMV and anti-EBV IgG titers in Long Covid patients at 6-month follow-up suggested increased viral antigen exposure and, thus, antiviral antibody formation in Long Covid patients, perhaps through reactivation of herpesviruses. The association of C2 levels with anti-CMV IgG titers and Long Covid persistence may connect our finding of increased complement activation to recent reports of herpesvirus reactivation in Long Covid (*52*). Furthermore, C5 could be activated directly through thrombin, as supported by our findings of a dysregulated coagulation cascade, low antithrombin III levels in Long Covid patients in two independent cohorts, and increased antithrombin III cleavage at the thrombin-reactive site (*33*–*36*).

Elevated vWF and TSP-1 along with decreased protective factors, such as ADAMTS13, PAF-AH, and ApoA1, indicated thromboinflammatory responses, which can also drive complement activation (*38*, *39*, *47*). Both vWF and TSP-1 can be secreted by endothelial cells and thrombocytes. Persistently low ICAM-1 levels in Long Covid patients argue against strong endothelial activation and have been associated with genetic ICAM-1 variants (*41*, *56*). The finding of increased monocyte–platelet aggregates in Long Covid patients suggested platelet activation at 6-month follow-up (*57*). However, normal platelet counts despite the presence of increased monocyte–platelet aggregates indicated compensatory mechanisms. As single-cell monocyte transcriptomes lacked signs of a prothrombotic signature, aggregates may be driven by platelet activation rather than by monocytes (*50*). However, monocytes showed distinct transcriptomic changes in Long Covid patients, including decreased *NR4A1* and increased transcripts for interferon-induced transmembrane proteins. Interestingly, NR4A1-dependent monocyte subsets have been associated with endothelial homeostasis in mice (*58*). Moreover, increased interferon-induced transcripts have been associated with severe cases of SARS-CoV-2 infection (*59*). As with monocytes, we did not observe any increased markers of neutrophil activation and NETosis, and neutrophil counts were normal at 6-month follow-up.

Pathological complement activation has been previously associated with thromboinflammatory and microangiopathic manifestations (*38*). Chronic diseases involving complement activation include neurodegenerative diseases (*60*) as well as atypical hemolytic uremic syndrome and paroxysmal nocturnal hemoglobinuria, which are marked by complement-mediated hemolysis (*38*). We found signs of a hemolytic process in Long Covid, including low hemopexin

and increased heme levels in Long Covid patients. Hemopexin–heme complexes are formed upon hemolysis to prevent heme-mediated oxidative damage. These findings were accompanied by increased haptoglobin and IL-6, as well as normal hemoglobin and erythrocyte counts, suggestive of a chronic inflammatory process (*42*, *61*, *62*). As tissue injury can be mediated by complement, and in turn also activates the complement system, the observed markers may be both consequence and cause of complement activation. Clinical history and the overall pattern of our findings suggested chronic processes, possibly maintained by external drivers, including chronic herpesvirus infection, or by a self-perpetuating cycle of thromboinflammation.

Excessive complement activation has been reported in severe acute COVID-19 (*32*). MBL, a central member of the lectin pathway, can directly bind SARS-CoV-2 (*63*) and is increased in severe pediatric cases (*64*) but was unchanged at 6-month follow-up in our cohort. Increased PTX3 has been associated with short-term COVID-19 mortality (*65*) and Long Covid at 8-month follow-up (*66*) but was not altered in our study by two different methods. Other pathways of complement activation include intracellular mechanisms in SARS-CoV-2–infected host cells (*67*, *68*). COVID-19–associated thromboinflammation has been described in hospitalized patients (*32*) and has also been proposed to underlie Long Covid (*69*–*71*). Brain autopsies of COVID-19 patients showed signs of neurovascular injury accompanied by classical complement component deposition on endothelial cells and platelets (*72*). Moreover, the vWF/ADAMTS13 ratio was associated with increased thrombogenicity in Long Covid patients (*73*), and vWF and C7 were among the top enriched proteins in amyloid-like microclots found in Long Covid patients, supporting a central role of C7 (*74*).

Our multicenter, longitudinal study provides evidence of an inflammatory signature restricted to patients with active Long Covid, with diagnostic accuracy 6 months after symptom onset and independent of any information on COVID-19 history, thus facilitating clinical applicability. Moreover, our findings address the gap between observed complement-activating properties of SARS-CoV-2 and reports of microclots, vascular inflammation, and cardiovascular complications (*75*). We were able to validate our hypothesis of dysregulated complement activation by different experimental approaches as well as an independent longitudinal cohort of hospitalized COVID-19 patients. However, external validation of our findings with larger cohorts, including Long Covid patients with mild acute COVID-19 and later blood sampling time points, is needed. Larger cohorts might allow more granular assessment of Long Covid subgroups. Other limitations include caveats of high-throughput biomarker discovery, which were addressed by multiple testing correction, hypothesis-driven analysis, and validation of results by additional methods. The specificity of C7 aptamer for C7 complexes, but not monomeric C7, highlights the need of experimental validations when applying high-throughput approaches.

Given our data, early cardiovascular assessment of Long Covid patients should be considered. Moreover, antivirals targeting SARS-CoV-2 or herpesviruses could reduce thromboinflammatory responses. Available therapeutics targeting the terminal complement pathway could offer new treatment strategies for Long Covid and possibly other postinfection syndromes.

Materials and methods

Experimental study design and Zurich cohort

This study was approved by the Cantonal Ethics Committee of Zurich (BASEC #2016-01440). Adult individuals were included in the study between April 2020 and April 2021 on the basis of RT-qPCR–confirmed SARS-CoV-2 infection. The latest inclusion time point for acute SARS-CoV-2–infected patients was January 2021. Follow-up was continued until 1 year after acute COVID-19 disease, with the last follow-up visit in February 2022.

Following written informed consent, 152 study participants attended one to three study visits (fig. S1A), consisting of a structured questioning by trained physicians and a blood draw. Thirty-nine healthy controls donated blood during the first study visit, and 113 COVID-19 patients donated blood during acute infection as well as 6 months later at the second study visit (6-month follow-up). Eighty-five COVID-19 patients attended a third study visit 12 months after acute COVID-19 (12-month follow-up). Patients unable to attend the 12-month follow-up visit in person were questioned by phone by a trained physician. Eighteen patients declined a 12-month follow-up or could not be contacted. The multicenter study was composed of patients from four different hospitals in the region of Zurich, Switzerland, including the University Hospital Zurich (n = 77), City Hospital Triemli (n = 18), Limmattal Hospital (n = 13), and Uster Hospital (n = 5). No further inclusion or exclusion criteria were applied on analysis of the 152 study participants. All healthy controls were included at the University Hospital Zurich and reported no history of SARS-CoV-2 infection and no acute or active illness prior to blood sampling, and they had a negative serology for SARS-CoV-2 spike S1–specific IgA and IgG. Participants were mostly Caucasian, and ethnicity was not systematically assessed.

Mount Sinai COVID-19 cohort

This cohort was approved by the Human Research Protection Program at the Icahn School of Medicine at Mount Sinai (STUDY-20-00341). Following written informed consent, 280 individuals were enrolled in the Mount Sinai COVID-19 cohort between April and June 2020 and had venous blood sampled for proteomics analysis at SomaLogic, Inc. (Boulder, CO). Included in this cohort were 198 patients hospitalized with COVID-19 and 82 controls, including 35 healthy controls and 47 patients hospitalized for reasons unrelated to COVID-19. Patients with acute COVID-19 were followed-up with 6 months or later [median: 414 days; interquartile range (IQR): 326 to 503] after COVID-19 hospital admission (fig. S1E). No further inclusion or exclusion criteria were applied on analysis of the 280 study participants with complete proteomics analysis and follow-up. Acute COVID-19 sampling was performed at a median time point of 1 day (IQR: 0 to 2) after hospital admission and before discharge. A subset of 21 patients underwent venous blood collection for further proteomics analysis after hospital discharge at a median time point of 89 days (IQR: 53 to 125) after hospital admission, i.e., 3-month follow-up (fig. S4F). Demographic characteristics of the cohort are given in table S3.

Definitions

Acute COVID-19 severity was graded according to recommendations of the World Health Organization (WHO) classification criteria into mild cases, including asymptomatic cases and patients with mild illness or pneumonia, and severe cases, consisting of patients with severe pneumonia, or mild, moderate, or severe acute respiratory distress syndrome (ARDS) (*1*, *18*, *76*). Symptomatic patients with uncomplicated SARS-CoV-2 infection, defined by an absence of abnormal vital signs and pathological lung examination findings, were classified as patients with mild illness, whereas patients with complicated SARS-CoV-2 infection, defined by the presence of one or more aforementioned criteria, were classified as patients with pneumonia. Patients with pneumonia requiring supplemental oxygen therapy, presenting with symptoms of severe respiratory distress, an increased respiratory rate (>30 breaths/min), and/or blood oxygen saturation levels below 93% on ambient air, were classified as patients with severe COVID-19. Severe COVID-19 patients were further classified on the basis of blood gas analysis (PaO_2/FiO_2 ratio) into severe pneumonia (>300 mmHg or unavailable), mild (≤300 mmHg), moderate (≤200 mmHg), and severe ARDS (≤100 mmHg) (*1*, *18*, *76*). Acute COVID-19 severity was followed up until recovery or hospital release, thus, beyond the blood-sampling time point. Long Covid was defined as one or more persisting COVID-19–related symptoms, which could not be explained by an alternative diagnosis (*2*). In the present study, patients reporting isolated changes in

smell and/or taste were excluded from the Long Covid definition prior to analysis of the serum proteome, as pathomechanistic causes might be distinct in this patient subgroup (*19*). All reported symptoms are listed in table S1. Moreover, in accordance with the WHO clinical case definition of post COVID-19 condition (*2*), 6-month Long Covid and 12-month Long Covid were defined as persisting Long Covid symptoms at 6-month and 12-month follow-up, respectively. "Recovered Long Covid" included COVID-19 patients reporting one or more symptoms lasting for more than 4 weeks but no symptoms at 6-month follow-up (*5*, *77*).

For the Mount Sinai cohort, 6-month Long Covid symptoms were defined in the same manner as for the Zurich cohort. These included fatigue, pulmonary symptoms, chest pain, gastrointestinal symptoms, cognitive symptoms, anxiety or depression, headache, and joint and muscle pain (table S3).

Serum processing and SomaScan proteomics

Following venous blood sampling, BD Vacutainer CAT serum tubes (Becton Dickinson, Franklin Lakes, NJ) were centrifuged at 1100*g* and 4°C for 10 min, followed by aliquotation and storage at −80°C. Thirty-nine serum samples of healthy controls and 113 paired serum samples of COVID-19 patients at acute COVID-19 and 6-month follow-up were used for proteomics analysis at SomaLogic using the SomaScan platform (version 4). Protein measurements were performed using 7335 modified single-stranded aptamers (SOMAmer reagents), including 7289 aptamers specific to 6596 unique human proteins and 46 internal controls, as previously described (*20*, *78*). Protein epitope–SOMAmer reagent complexes were quantified by fluorescence using DNA-hybridization microarrays. All measurements passed manufacturer-defined quality control standards. Assay-intrinsic variation was corrected by median normalization using external references derived from 1025 healthy participants (*79*). Relative fluorescence units were $\log_{10}$-transformed for analysis.

Mass spectrometry sample preparation, data acquisition, and processing

Samples were prepared as follows: 2.5 µl of the serum/plasma was transferred to a 0.5-ml 96-deepwell plate containing 27.5 µl denaturation/reduction buffer [8 M urea, 100 mM ammonium bicarbonate (ABC) and 46 mM dithiothreitol]. The plates were subsequently centrifuged for 20 s at pulse setting (Heraeus Multifuge X3R, Rotor TX-1000), mixed, and incubated at 30°C and 500 rpm for 60 min (Eppendorf Thermomixer C). Next, 2.5 µl of 100 mM chloroacetamide (CAA) was added to each sample, and the plate was incubated in the dark at 23°C and 500 rpm for 30 min (Eppendorf Thermomixer C). The samples were then diluted with 170 µl 100 mM ABC. Subsequently, 12.5 µl trypsin solution (0.092 µg/µl) was added, and samples were incubated at 37°C and 500 rpm for 17 hours. Digestion was stopped by adding 10% formic acid (FA) to reach a final concentration of 1% FA. Peptides were desalted using an Oasis Prime HLB 96-well µElution Plate. Samples were loaded onto columns, washed twice with 400 µl 0.1% FA and once with 200 µl ddH_2O, and eluted with 50 µl 70% acetonitrile after a 1-min incubation at room temperature (RT). The liquid was evaporated with a SPD120 SpeedVac (Thermo Fisher Scientific), and samples were resuspended in 80 µl 0.1% FA, sonicated for 5 min in an ultrasonication water bath (Elmasonic P90), and centrifuged for 10 min at 3000 rpm and 4°C. Twenty microliters of the resulting samples was transferred to a 96-well plate, and 1 µl was injected.

Samples were acquired on an EASY-nLC 1200 coupled to a Thermo Orbitrap Eclipse Tribrid MS in data-independent acquisition (DIA) mode. Buffer A was 0.1% FA in water, and buffer B was 0.1% FA in 80% acetonitrile. A nanoLC column with integrated emitter from CoAnn Technologies with the following dimensions was used: 75 µm ID x 25 cm L x 365 µm OD, ReproSil-Pur120 C18 particles (1.9 µm). Further, an Acclaim PepMap 100 C18 pre-column (0.1 mm ID x 150 mm L with 5 µm particle size) was used. Flow rate was set to 200 nl/min, and the following gradient was applied: 45 min from 5% B to 31% B, 5 min from 31% B to 44% B, 1 min from 44% B to 95% B, 10 min hold at 95% B, 1 min from 95% B to 5% B, and 5 min equilibration at 5% B. For the MS1 scan, orbitrap resolution was set to 120,000, with quadrupole isolation turned on. A scan range of 380 to 980 m/z was applied, the RF lens was set to 30%, and standard automatic gain control (AGC) target with a custom maximum injection time of 50 ms was applied. For MS2 scan, orbitrap resolution was set to 15,000, with 50 fixed windows (12 m/z isolation window and 0.5 m/z overlap). Higher-energy collisional dissociation collision energy was fixed at 30%. MS/MS scan range was defined as 145 to 1450 m/z, and RF Lens was set to 30%. AGC Target was set to 1000% with custom maximum injection time of 22 ms.

Raw data was processed using DIA-NN software with default settings (*80*). Search was performed library-free with an in silico digestion and deep learning–based spectra and retention time prediction. The human fasta file was downloaded from UNIPROT (3AUP00005640Uniprot, downloaded on 7 September 2022). Downstream analysis was performed with R. For protein analysis, precursors were filtered for proteotypicity, precursor q-values were filtered with a threshold of 0.01, and protein group q-values were filtered with a threshold of 0.05. Only precursors quantified in >50% of samples were considered. Signal drifts were corrected by fitting a loess function for each precursor across the measurements (span = 0.7). Protein quantities were calculated using maxLFQ algorithm (maxLFQ function within DIA-NN R-package). For peptide level analysis of antithrombin III (gene name: *SERPINC1*), peptides with >10 observations measured per group (i.e., "no 6-month Long Covid" and "6-month Long Covid") were summarized on the basis of their identical sequences, independently of differences in oxidation.

Protein and heme quantifications by ELISA, functional assays, and complete blood counts

Total C7 (AssayPro; Cat#EC7101-1), sC5b-9 (Quidel; Cat#A020), factor Ba (Quidel; Cat#A033), CH50 Equivalent (Quidel; Cat#A018), and vWF (AssayPro; Cat#CEV20301) were measured by applying commercially available ELISAs on patient sera analyzed with the SomaScan platform (39 healthy controls, and 113 COVID-19 patients at acute COVID-19 and 6-month follow-up), according to the manufacturers' instructions. vWF measurements were calibrated against WHO controls. Factor Ba purified from human serum (CompTech; Cat#A154) was used as positive control. Six-month follow-up serum samples (n = 113) were used for measurements of PTX3 (AssayPro; Cat# EP2877-1) and heme (Sigma-Aldrich; Cat# MAK316). IL-6 serum concentrations were determined by ELISA (R&D Systems) on an Opsys Reader (Dynex). IL-6 measurements, hemoglobin, and complete blood counts were performed at an accredited laboratory of the University Hospital Zurich using serum or fresh whole blood.

C7-specific and SARS-CoV-2–specific IgG measurement

C7-specific autoantibodies were assessed using a previously described method (*81*). In brief, human C7 protein (Quidel; Cat#A405) diluted in phosphate-buffered saline (PBS) to 1 µg/ml was dispensed into high-binding 384-well SpectraPlates (Perkin Elmer) at 20 µl per well using Multidrop Combi nL (Thermo Fisher Scientific). Following coating overnight at 4°C, plates were washed three times with PBS Tween-20 0.1% (PBS-T) on a Biotek El406 device using the 192-pin manifold. Forty microliters per well of blocking buffer (PBS-T with 5% milk powder) was dispensed, and plates were blocked for 1 hour at RT. Following removal of blocking agent, serial dilutions were done of serum samples (range: 0.02 to 2.74×10^{-5}) and of a polyclonal HRP-linked rabbit anti-human anti-C7 antibody as positive control (Novus Biologicals; Cat#NBP3-00202H; range: 4 to 8.4×10^{-7} µg/ml; fig. S2I), in sample buffer (PBS-T with 1% milk powder). Following incubation at RT for 2 hours, plates were washed five times. Twenty microliters per well of secondary antibody (HRP-conjugated goat anti-human IgG; Jackson), diluted 1:4000 in sample buffer, was dispensed into wells containing plasma samples, and 20 µl/well sample

buffer without secondary antibody into wells containing positive control using a Biotek Multiflo FX dispenser. After an incubation of 1 hour at RT and three wash cycles, 20 µl/well TMB followed by 20 µl/well 0.5 M H_2SO_4 were dispensed using an eight-pin manifold on a Certus Flex device (Fritz Gyger AG). The incubation time between the two reagents was 5 min.

SARS-CoV-2 spike-specific IgG was measured with the TRABI technology, as previously described (*81*), using wild-type/Wuhan SARS-CoV-2 spike extracellular domain, spike S1, spike S2 (AcroBiosystems; Cats#S1N-C52H2, S2N-C52H5), or receptor binding domain as antigens. The inflection points of sigmoidal binding curves [$-\log_{10}EC_{50}$ or $p(EC_{50})$ values of respective sample dilutions] were determined using a custom-designed fitting algorithm, with plateau and baseline inferred from respective positive and negative controls in a plate-wise manner. Negative $p(EC_{50})$ values, reflecting nonreactive samples, were rescaled as zero. Samples included in C7-specific measurements were assayed and analyzed as duplicates, whereas SARS-CoV-2–specific measurements were conducted as unicates.

Antinuclear, anti-chemokine, and antiviral antibody quantifications

Antinuclear antibodies were measured by indirect immunofluorescence assays on HEp-2 cells using a computer-aided microscopy system (Euroimmun) at an accredited laboratory of the University Hospital Zurich, in a blinded manner, as previously reported (*22*). As positivity cutoff a dilution of 1:320 was applied. ANA patterns were classified on the basis of international consensus on ANA patterns anticell (AC) nomenclature (*82*).

Anti-chemokine antibodies were quantified by ELISA, as previously described (*51*). Anti-chemokine levels measured on 12-month follow-up samples of our cohort have been previously reported (*51*). For the present study, published data were reanalyzed in the context of 6-month Long Covid, by including only a subset of patients, for which we also had SomaScan and mass spectrometry measurements at 6-month follow-up available (n = 88, with 31 experiencing 6-month Long Covid).

For human virome-wide serological profiling by VirScan (*22*, *83*), serum samples obtained at first study visit, including 57 acute COVID-19 patients (of which 22 developed 6-month Long Covid) and 18 healthy controls were measured, as previously reported (*22*). Briefly, duplicate serum samples were normalized for total IgG and incubated with a bacteriophage library displaying linear viral epitopes with lengths of 56 amino acids. IgG–phage complexes were isolated with magnetic beads and quantified by next-generation sequencing. Epitope hit counts were obtained using *SAMtools* (*84*) upon mapping of reads to the epitope library using *Bowtie2* (*85*). Positivity for epitopes was determined on the basis of a previously published binning strategy (*86*). Data on prokaryotes, eukaryotes, and nonhuman viruses, as well as human viruses with a maximal summed epitope hit count below three, were excluded.

For quantification of CMV- and EBV-specific IgM and IgG titers, ELISAs certified for diagnostic use were applied, according to the manufacturer's instruction, on 6-month follow-up serum samples of 113 COVID-19 patients (TestLine Clinical Diagnostics; Cats#CMM096, CMG096, VCM096, VCG096).

Anti-C7 aptamer ELISA, pull-down for SDS-PAGE, and Western blot analysis

For in vitro assembly of different TCC conformations, commercially available human recombinant C5bC6 (Sigma Aldrich; Cat#204906), C7 (Quidel; Cat#A405), C8 (Sigma Aldrich; Cat#204896), and C9 (Sigma Aldrich; Cat#204910) were mixed at an equimolar concentration (200 nM for pull-down) in SOMA1 buffer with Tween-20 buffer (120 mM NaCl, 5 mM KCl, 5 mM $MgCl_2$, 40 mM HEPES, pH 7.5, 0.05% Tween-20) and incubated overnight at 4°C.

For assessing C7 aptamer (seq.2888.49) specificity, streptavidin-coated plates (Thermo Fisher Scientific) were incubated overnight with the aptamer (20 nM diluted in SBT buffer). The next day, plates were blocked with 100 µM biotin (10 min) and 3% bovine serum albumin (30 min) in SBT buffer. Upon addition of different combinations of TCCs C5bC6, C7, C8, and C9 at increasing concentrations (31.25 to 4000 ng/ml; 2 hours incubation), C7 was detected with a primary, polyclonal, full-length human C7-directed, rabbit antibody (Novus Biologicals; Cat#H00000730-D01P) and an HRP-linked secondary, goat, anti-rabbit IgG antibody (Cell Signaling Technologies; Cat#7074S). This experimental setup was repeated for measurement of monomeric C7 and of complexed C5b, C6, C7, C8, and C9 at different concentrations and in triplicates. To reduce unspecific aptamer–protein interactions, 1 µM polyanionic competitor (SomaLogic) was added, reflecting similar conditions as the SomaScan platform.

For C7 aptamer-mediated pull-down of TCC components, Pierce Streptavidin magnetic beads (Thermo Scientific) were diluted in SBT buffer and mixed with 100 nM biotinylated C7 aptamer (SomaLogic, seq.2888.49) and incubated overnight at 4°C. The next day, beads were repeatedly washed with SBT buffer and mixed with complement protein solutions at a 1:1 ratio, together with 1 µM polyanionic competitor (SomaLogic). After 2 hours of incubation, mixtures were washed with SBT buffer, and captured protein complexes were eluted from beads using 20 mM NaOH. Complement protein solutions were mixed with Laemmli sample buffer (BioRad) and incubated at 95°C for 5 min. Subsequently, samples were loaded on 4 to 20% polyacrylamide gels (BioRad). After separation, gels were stained overnight using QuickBlue Protein Stain (Lubio science). Standard Western blotting protocol was applied. Rabbit anti-human C7 polyclonal antibodies (Novus Biologicals; Cat#H00000730-D01P) and HRP-linked goat anti-rabbit IgG antibodies (Cell Signaling Technologies; Cat#7074S) were used for detection. Acquisition was performed using Pierce ECL western substrate (Thermo Scientific) and imaged with a ChemiDoc Imaging System (BioRad).

Spectral flow cytometry and fluorescence-activated cell sorting (FACS)

Frozen PBMCs were thawed by slowly adding prewarmed RPMI medium (Gibco) containing 10% fetal bovine serum (FBS; Gibco). Cells were washed and stained, as previously published (*11*, *12*). Briefly, after staining with ZombieUV Viability dye (1:400; BioLegend) and TruStain FcX (1:200) in PBS for 30 min at 4°C, surface antigens were stained with a mix of diluted fluorophore-conjugated antibodies (listed in data S6) in a 1:10 mixture of Brilliant Buffer (BD Bioscience) and FACS buffer (PBS containing 2% FBS and 2 mM EDTA) for 30 min at 4°C. Samples were acquired on a Cytek Aurora spectral flow cytometer. Gating strategy is shown in fig. S6A. Spectral flow cytometry data were analyzed using FlowJo (version 10.9.0). Dimensionality reduction and clustering were performed using the *CATALYST* (*87*) package in R (version 4.22). The dataset was reduced to 5000 randomly selected monocytes per patient to reduce required computational resources. Expression was scaled by hyperbolic arcsine transformation with a cofactor of 6000. UMAPs were calculated on CD14, CD16, CD11b, and CD36 levels using *CATALYST*. Monocytes were clustered using *PhenoGraph* clustering workflow of *Rphenograph* (*88*) package (version 0.99.1).

For sorting of monocytes using PBMC samples, staining was applied as described for the spectral flow cytometry analysis. In brief, cells were stained with fixable viability dye eFluor 780 (eBioscience) and TruStain FcX, followed by incubation with fluorescently labeled antibodies. For single-cell RNA sequencing, each sample was additionally barcoded with a hashtag antibody (data S6). For validation of monocyte–platelet aggregate staining by CD41, FACS-sorted (FACS Aria III 4L, BD Biosciences) monocytes were incubated with increasing counts of isolated platelets, with and without prestimulation of monocytes by LPS (10 ng/ml). Platelets were isolated from a healthy donor by centrifugation of EDTA blood, once at 1000*g* for 15 min and a second time at 13,000*g* for 5 min. Gating strategy is shown in fig. S6B.

Single-cell RNA sequencing

FACS-sorted PBMCs were immediately processed for further analysis by single-cell RNA sequencing. The resulting libraries were processed on an Illumina NovaSeq 6000 S1 flow

cell. Sequencing parameters were set according to 10X Genomics' recommendations, using paired-end reads with the following specifications: R1 = 26, i7 = 10, i5 = 10, and R2 = 90. An average depth of around 50,000 reads per cell was achieved. A custom reference genome was built to screen for herpesviral RNA transcripts by concatenating exogenous sequences in fasta format from NCBI to the human reference genome (build GRCh38.p13). Subsequently, the whole exogenous sequence was added to gene model (GENCODE Release 42) as an exon in GTF format, as recommended by 10X Genomics. The resulting hybrid fasta and GTF files were used as inputs to generate a 10X-compatible reference genome using *CellRanger* (v7.1.0) *mkfastq* (*89*). The *CellRanger* pipeline was used for demultiplexing, read alignment to the custom reference, barcode processing, and unique molecular identifier (UMI) counting. Custom references are listed in data S7.

Subsequent data analysis was performed in R. Single cells were mapped to a PBMC reference dataset (version 1.0.0) (*90*) using *Seurat* (version 4.3.0) (*91*), as described in the multimodal reference mapping vignette. Differential gene expression analysis of monocytes was performed using the *Seurat* package.

Gene set enrichment analysis, association analysis, and random forest prediction model

Univariate association analysis of protein-aptamer concentrations to binary 6-month Long Covid endpoints was performed using logistic regression, adjusting for patient age, sex, and hospitalization status (6-month Long Covid ~ aptamer xyz + patient age [numeric] + sex [female/male] + hospitalization during acute COVID-19 [yes/no] + intercept constant), implemented as the python Statsmodels API Logit class. Analyses were performed with respect to either acute COVID-19 or 6-month follow-up measurements. The obtained *P* values for the individual aptamers were used for further analysis, such as differential expression and ranking of protein concentrations for enrichment analysis. For gene set enrichment analysis, duplicate measurements of proteins, resulting from multiple aptamers targeting the same protein, were selected on the basis of the lowest *P* value of competing aptamers, resulting in 6408 aptamers with unique Entrez Gene ID's. Subsequently, these 6408 human proteins were ranked by multiplying univariate association $-\log_{10}$-transformed *P* values with the sign of recorded fold changes. Gene set enrichment analysis was performed using the *fgsea* R package with Reactome pathways (*26*, *92*).

Biological protein clusters were selected on the basis of protein measurement availability and maximizing uniqueness and representativeness for the respective pathways and protein groups according to the literature (table S2). Biological protein cluster associations with 6-month Long Covid were assessed by association analysis through logistic regression, with all protein levels of a specific cluster or pathway being used as model inputs, in addition to patient age, sex, and hospitalization status. Relevant *P* values in these analyses represented the log likelihood ratios (LLR) of the model as a whole. For associations, we determined a Bonferroni-corrected significance threshold of $P < 6.80 \times 10^{-6}$, corresponding to a *P* value of 0.05 divided by 7353, including all 7335 aptamer measurements and 18 biological protein clusters.

Finally, binary predictions of 6- and 12-month Long Covid were performed using a random forest classifier implemented using the *sklearn RandomForestClassifier* (*93*). Input data were obtained at acute COVID-19 or 6-month follow-up and comprised all SomaScan proteomic measurements, four protein ratios based on biological reasoning and published literature (C5/C7 complexes, C5bC6/C7 complexes, vWF/ADAMTS13, and angiopoietin-2/angiopoietin-1), and 14 general clinical variables assessed at acute COVID-19 or 6-month follow-up. We purposely excluded clinical variables on COVID-19 history, as this information may not be available for patients assessed for Long Covid diagnosis in a clinical setting. Of 7289 aptamer measurements, 61 uncorrelated measurements (with absolute Spearman correlation coefficient values below 0.3) were selected together with two uncorrelated protein ratios, and 14 clinical variables for prediction of 6-month Long Covid versus no 6-month Long Covid and healthy controls (n = 40, 73, and 39, respectively) with acute COVID-19 or 6-month follow-up data. We performed stratified (for patient age group [<60 years/>60 years] and acute COVID severity [mild/severe]) fivefold nested cross-validation using 20% of the samples for testing, and inner fivefold cross-validation for hyperparameter tuning by *GridSearch* optimizing tree depth, number of leaves, sample splits, and maximum number of features. All models were trained with balanced class weights and results were reported as mean values and standard deviations of independent test sets in terms of AUROC and relative average precision scores, reflecting the area under the precision recall curve normalized by positive class prevalence. The latter score indicates the improvement beyond a random classifier for values exceeding unity. To infer features that drove predictions, we performed post-hoc interpretability analysis using SHapley Additive exPlanations as part of the python *shap* package (*94*). We reported an agglomerated interpretation overview in terms of summed ranks of individual fivefold cross-validation interpretations. The top features obtained in this analysis and identified by visual inspection of an elbow plot were selected for reevaluation in the prediction task, as outlined above. Importantly, the identified top four features were selected on training data allowing further predictions followed by evaluation on test data.

Visualization and statistics

Descriptive statistics for the study participants are reported in numbers and percentages of total for categorical as well as median and IQR for continuous variables. Comparison of paired variables was performed using Wilcoxon's signed-rank test and comparison of nonpaired variables was performed using nonparametric two-sided Wilcoxon's rank-sum test. Split violin plots overlayed with boxplots visualize median and quartiles. Correlations were calculated using Spearman's correlation in R. Differential expression in volcano plots was calculated using *t* tests, Wilcoxon's rank-sum test, or logistic regression, as specified. *P* values were $-\log_{10}$-transformed and fold changes $\log_2$-transformed. Multiple testing was corrected by Benjamini-Hochberg (*95*) or Bonferroni correction, as specified. Statistical analysis was performed with R studio (version 2023.03.1) and R (version 4.1.2) using the packages *SomaDataIO*, *fgsea*, *pheatmap*, *corrplot* (*96*), *lmerTest* (*97*), *VennDiagram*, and *tidyverse* (*98*) as well as python (version 3.7.4) using the packages *Numpy*, *Openpyxl*, *Scipy*, *Pandas*, *Scikit-learn*, *Statsmodels* (*99*), and *Shap*. Acute COVID-19 and 6-month follow-up clinical and SomaScan data contained no missing values. ELISA validation experiments contained a total of three to seven missing values. The number of observations per analysis are listed in data S9. The present study is reported according to the STROBE (Statement for reporting cohort studies) guidelines (*100*).

REFERENCES AND NOTES

1. World Health Organization, COVID-19 Clinical management: interim guidance, 27 May 2020 (World Health Organization, 2020); https://apps.who.int/iris/handle/10665/332196.
2. World Health Organization, A clinical case definition of post COVID-19 condition by a Delphi consensus, 6 October 2021 (World Health Organization, 2021): https://www.who.int/publications/i/item/WHO-2019-nCoV-Post_COVID-19_condition-Clinical_case_definition-2021.1.
3. Global Burden of Disease Long COVID Collaborators, Estimated global proportions of individuals with persistent fatigue, cognitive, and respiratory symptom clusters following symptomatic COVID-19 in 2020 and 2021. *JAMA* **328**, 1604–1615 (2022). doi: 10.1001/jama.2022.18931; pmid: 36215063
4. D. Menges *et al*., Burden of post-COVID-19 syndrome and implications for healthcare service planning: A population-based cohort study. *PLOS ONE* **16**, e0254523 (2021). doi: 10.1371/journal.pone.0254523; pmid: 34252157
5. W. Shah, T. Hillman, E. D. Playford, L. Hishmeh, Managing the long term effects of covid-19: Summary of NICE, SIGN, and RCGP rapid guideline. *BMJ* **372**, n136 (2021). doi: 10.1136/bmj.n136; pmid: 33483331
6. C. Kedor *et al*., A prospective observational study of post-COVID-19 chronic fatigue syndrome following the first pandemic wave in Germany and biomarkers associated with symptom severity. *Nat. Commun.* **13**, 5104 (2022). doi: 10.1038/s41467-022-32507-6; pmid: 36042189
7. H. E. Davis, L. McCorkell, J. M. Vogel, E. J. Topol, Long COVID: Major findings, mechanisms and recommendations. *Nat. Rev. Microbiol.* **21**, 133–146 (2023). doi: 10.1038/s41579-022-00846-2; pmid: 36639608

8. S. Chevrier *et al*., A distinct innate immune signature marks progression from mild to severe COVID-19. *Cell Rep. Med.* **2**, 100166 (2020). doi: 10.1016/j.xcrm.2020.100166; pmid: 33521697
9. A. K. Azkur *et al*., Immune response to SARS-CoV-2 and mechanisms of immunopathological changes in COVID-19. *Allergy* **75**, 1564–1581 (2020). doi: 10.1111/all.14364; pmid: 32396996
10. C. Cervia *et al*., Systemic and mucosal antibody responses specific to SARS-CoV-2 during mild versus severe COVID-19. *J. Allergy Clin. Immunol.* **147**, 545–557.e9 (2021). doi: 10.1016/j.jaci.2020.10.040; pmid: 33221383
11. S. Adamo *et al*., Signature of long-lived memory $CD8^+$ T cells in acute SARS-CoV-2 infection. *Nature* **602**, 148–155 (2022). doi: 10.1038/s41586-021-04280-x; pmid: 34875673
12. Y. Zurbuchen *et al*., Human memory B cells show plasticity and adopt multiple fates upon recall response to SARS-CoV-2. *Nat. Immunol.* **24**, 955–965 (2023). doi: 10.1038/s41590-023-01497-y; pmid: 37106039
13. C. Cervia *et al*., Immunoglobulin signature predicts risk of post-acute COVID-19 syndrome. *Nat. Commun.* **13**, 446 (2022). doi: 10.1038/s41467-021-27797-1; pmid: 35078982
14. S. Guenther *et al*., Frequent IgG subclass and mannose binding lectin deficiency in patients with chronic fatigue syndrome. *Hum. Immunol.* **76**, 729–735 (2015). doi: 10.1016/j.humimm.2015.09.028; pmid: 26429318
15. R. C. Thompson *et al*., Molecular states during acute COVID-19 reveal distinct etiologies of long-term sequelae. *Nat. Med.* **29**, 236–246 (2023). doi: 10.1038/s41591-022-02107-4; pmid: 36482101
16. L. Perico *et al*., Immunity, endothelial injury and complement-induced coagulopathy in COVID-19. *Nat. Rev. Nephrol.* **17**, 46–64 (2021). doi: 10.1038/s41581-020-00357-4; pmid: 33077917
17. M. Merad, C. A. Blish, F. Sallusto, A. Iwasaki, The immunology and immunopathology of COVID-19. *Science* **375**, 1122–1127 (2022). doi: 10.1126/science.abm8108; pmid: 35271343
18. World Health Organization, COVID-19 Clinical management: living guidance, 25 January 2021 (World Health Organization, 2021); https://iris.who.int/bitstream/handle/10665/338882/WHO-2019-nCoV-clinical-2021.1-eng.pdf.
19. J. F. Shelton, A. J. Shastri, K. Fletez-Brant, S. Aslibekyan, A. Auton; 23andMe COVID-19 Team, The *UGT2A1/UGT2A2* locus is associated with COVID-19-related loss of smell or taste. *Nat. Genet.* **54**, 121–124 (2022). doi: 10.1038/s41588-021-00986-w; pmid: 35039640
20. L. Gold *et al*., Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLOS ONE* **5**, e15004 (2010). doi: 10.1371/journal.pone.0015004; pmid: 21165148
21. S. E. Chang *et al*., New-onset IgG autoantibodies in hospitalized patients with COVID-19. *Nat. Commun.* **12**, 5417 (2021). doi: 10.1038/s41467-021-25509-3; pmid: 34521836
22. P. Taeschler *et al*., Autoantibodies in COVID-19 correlate with antiviral humoral responses and distinct immune signatures. *Allergy* **77**, 2415–2430 (2022). doi: 10.1111/all.15302; pmid: 35364615
23. PHOSP-COVID Collaborative Group, Clinical characteristics with inflammation profiling of long COVID and association with 1-year recovery following hospitalisation in the UK: A prospective observational study. *Lancet Respir. Med.* **10**, 761–775 (2022). doi: 10.1016/S2213-2600(22)00127-8; pmid: 35472304
24. P. Taeschler *et al*., T-cell recovery and evidence of persistent immune activation 12 months after severe COVID-19. *Allergy* **77**, 2468–2481 (2022). doi: 10.1111/all.15372; pmid: 35567391
25. B. Blomberg *et al*., Long COVID in a prospective cohort of home-isolated patients. *Nat. Med.* **27**, 1607–1613 (2021). doi: 10.1038/s41591-021-01433-3; pmid: 34163090
26. M. Gillespie *et al*., The reactome pathway knowledgebase 2022. *Nucleic Acids Res.* **50**, D687–D692 (2022). doi: 10.1093/nar/gkab1028; pmid: 34788843
27. J. V. Sarma, P. A. Ward, The complement system. *Cell Tissue Res.* **343**, 227–235 (2011). doi: 10.1007/s00441-010-1034-0; pmid: 20838815
28. A. Mantovani, C. Garlanda, Humoral innate immunity and acute-phase proteins. *N. Engl. J. Med.* **388**, 439–452 (2023). doi: 10.1056/NEJMra2206346; pmid: 36724330
29. R. Würzner *et al*., Functionally active complement proteins C6 and C7 detected in C6- and C7-deficient individuals. *Clin. Exp. Immunol.* **83**, 430–437 (1991). doi: 10.1111/j.1365-2249.1991.tb05656.x; pmid: 2004484
30. R. Würzner, Modulation of complement membrane attack by local C7 synthesis. *Clin. Exp. Immunol.* **121**, 8–10 (2000). doi: 10.1046/j.1365-2249.2000.01263.x; pmid: 10886232
31. F. Bossi *et al*., C7 is expressed on endothelial cells as a trap for the assembling terminal complement complex and may exert anti-inflammatory function. *Blood* **113**, 3640–3648 (2009). doi: 10.1182/blood-2008-03-146472; pmid: 19179470
32. L. Ma *et al*., Increased complement activation is a distinctive feature of severe SARS-CoV-2 infection. *Sci. Immunol.* **6**, eabh2259 (2021). doi: 10.1126/sciimmunol.abh2259; pmid: 34446527
33. M. Huber-Lang *et al*., Generation of C5a in the absence of C3: A new complement activation pathway. *Nat. Med.* **12**, 682–687 (2006). doi: 10.1038/nm1419; pmid: 16715088
34. M. J. Krisinger *et al*., Thrombin generates previously unidentified C5 products that support the terminal complement activation pathway. *Blood* **120**, 1717–1725 (2012). doi: 10.1182/blood-2012-02-412080; pmid: 22802338
35. D. A. Lane, R. Caso, Antithrombin: Structure, genomic organization, function and inherited deficiency. *Baillieres Clin. Haematol.* **2**, 961–998 (1989). doi: 10.1016/S0950-3536(89)80054-X; pmid: 2688761
36. A. R. Rezaie, H. Giri, Anticoagulant and signaling functions of antithrombin. *J. Thromb. Haemost.* **18**, 3142–3153 (2020). doi: 10.1111/jth.15052; pmid: 32780936
37. S. Pillarisetti, Lipoprotein modulation of subendothelial heparan sulfate proteoglycans (perlecan) and atherogenicity. *Trends Cardiovasc. Med.* **10**, 60–65 (2000). doi: 10.1016/S1050-1738(00)00048-7; pmid: 11150731
38. D. Ricklin, E. S. Reis, J. D. Lambris, Complement in disease: A defence system turning offensive. *Nat. Rev. Nephrol.* **12**, 383–401 (2016). doi: 10.1038/nrneph.2016.70; pmid: 27211870
39. S. P. Jackson, R. Darbousset, S. M. Schoenwaelder, Thromboinflammation: Challenges of therapeutically targeting coagulation and other host defense mechanisms. *Blood* **133**, 906–918 (2019). doi: 10.1182/blood-2018-11-882993; pmid: 30642917
40. L. Camous *et al*., Complement alternative pathway acts as a positive feedback amplification of neutrophil activation. *Blood* **117**, 1340–1349 (2011). doi: 10.1182/blood-2010-05-283564; pmid: 21063021
41. J. Zhang, Biomarkers of endothelial activation and dysfunction in cardiovascular diseases. *Rev. Cardiovasc. Med.* **23**, 73 (2022). doi: 10.31083/j.rcm2302073; pmid: 35229564
42. J. R. Delanghe, M. R. Langlois, Hemopexin: A review of biological aspects and the role in laboratory medicine. *Clin. Chim. Acta* **312**, 13–23 (2001). doi: 10.1016/S0009-8981(01)00586-1; pmid: 11580905
43. S. Wang *et al*., S100A8/A9 in inflammation. *Front. Immunol.* **9**, 1298 (2018). doi: 10.3389/fimmu.2018.01298; pmid: 29942307
44. G. Baidildinova *et al*., Soluble platelet release factors as biomarkers for cardiovascular disease. *Front. Cardiovasc. Med.* **8**, 684920 (2021). doi: 10.3389/fcvm.2021.684920; pmid: 34235190
45. S. Kamath, A. D. Blann, G. Y. Lip, Platelet activation: Assessment and quantification. *Eur. Heart J.* **22**, 1561–1571 (2001). doi: 10.1053/euhj.2000.2515; pmid: 11492985
46. D. M. Smadja *et al*., Thrombospondin-1 is a plasmatic marker of peripheral arterial disease that modulates endothelial progenitor cell angiogenic properties. *Arterioscler. Thromb. Vasc. Biol.* **31**, 551–559 (2011). doi: 10.1161/ATVBAHA.110.220624; pmid: 21148423
47. J. Chen, D. W. Chung, Inflammation, von Willebrand factor, and ADAMTS13. *Blood* **132**, 141–147 (2018). doi: 10.1182/blood-2018-02-769000; pmid: 29866815
48. A. Marco, P. Marco, Von Willebrand factor and ADAMTS13 activity as clinical severity markers in patients with COVID-19. *J. Thromb. Thrombolysis* **52**, 497–503 (2021). doi: 10.1007/s11239-021-02457-9; pmid: 33866481
49. M. Lu, Q. Lu, Y. Zhang, G. Tian, ApoB/apoA1 is an effective predictor of coronary heart disease risk in overweight and obesity. *J. Biomed. Res.* **25**, 266–273 (2011). doi: 10.1016/S1674-8301(11)60036-5; pmid: 23554700
50. A. K. Maher *et al*., Transcriptional reprogramming from innate immune functions to a pro-thrombotic signature by monocytes in COVID-19. *Nat. Commun.* **13**, 7947 (2022). doi: 10.1038/s41467-022-35638-y; pmid: 36572683
51. J. Muri *et al*., Autoantibodies against chemokines post-SARS-CoV-2 infection correlate with disease course. *Nat. Immunol.* **24**, 604–611 (2023). doi: 10.1038/s41590-023-01445-w; pmid: 36879067
52. Y. Su *et al*., Multiple early factors anticipate post-acute COVID-19 sequelae. *Cell* **185**, 881–895.e20 (2022). doi: 10.1016/j.cell.2022.01.014; pmid: 35216672
53. M. L. Gulley, Molecular diagnosis of Epstein-Barr virus-related diseases. *J. Mol. Diagn.* **3**, 1–10 (2001). doi: 10.1016/S1525-1578(10)60642-3; pmid: 11227065
54. E. Callaway, Beyond Omicron: What's next for COVID's viral evolution. *Nature* **600**, 204–207 (2021). doi: 10.1038/d41586-021-03619-8; pmid: 34876665
55. Office for National Statistics, Prevalence of ongoing symptoms following coronavirus (COVID-19) infection in the UK: 1 June 2022, Statistical Bulletin (2022); https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/prevalenceofongoingsymptomsfollowingcoronaviruscovid19infectionintheuk/1june2022.
56. H. Krönig *et al*., ICAM G241A polymorphism and soluble ICAM-1 serum levels: Evidence for an active immune process in schizophrenia. *Neuroimmunomodulation* **12**, 54–59 (2005). doi: 10.1159/000082364; pmid: 15756053
57. S.-H. Yun, E.-H. Sim, R.-Y. Goh, J.-I. Park, J.-Y. Han, Platelet activation: the mechanisms and potential biomarkers. *BioMed Res. Int.* **2016**, 9060143 (2016). doi: 10.1155/2016/9060143; pmid: 27403440
58. L. M. Carlin *et al*., *Nr4a1*-dependent $Ly6C^{low}$ monocytes monitor endothelial cells and orchestrate their disposal. *Cell* **153**, 362–375 (2013). doi: 10.1016/j.cell.2013.03.010; pmid: 23582326
59. J. S. Lee, E.-C. Shin, The type I interferon response in COVID-19: Implications for treatment. *Nat. Rev. Immunol.* **20**, 585–586 (2020). doi: 10.1038/s41577-020-00429-3; pmid: 32788708
60. M. C. Dalakas, H. Alexopoulos, P. J. Spaeth, Complement in neurological disorders and emerging complement-targeted therapeutics. *Nat. Rev. Neurol.* **16**, 601–617 (2020). doi: 10.1038/s41582-020-0400-0; pmid: 33005040
61. A. Smith, R. J. McCulloh, Hemopexin and haptoglobin: Allies against heme toxicity from hemoglobin not contenders. *Front. Physiol.* **6**, 187 (2015). doi: 10.3389/fphys.2015.00187; pmid: 26175690
62. B.-N. Wan, S.-G. Zhou, M. Wang, X. Zhang, G. Ji, Progress on haptoglobin and metabolic diseases. *World J. Diabetes* **12**, 206–214 (2021). doi: 10.4239/wjd.v12.i3.206; pmid: 33758643
63. M. Stravalaci *et al*., Recognition and inhibition of SARS-CoV-2 by humoral innate immunity pattern recognition molecules. *Nat. Immunol.* **23**, 275–286 (2022). doi: 10.1038/s41590-021-01114-w; pmid: 35102342
64. A. Rajamanickam *et al*., Levels of complement components in children with acute COVID-19 or multisystem inflammatory syndrome. *JAMA Netw. Open* **6**, e231713 (2023). doi: 10.1001/jamanetworkopen.2023.1713; pmid: 36961465
65. E. Brunetta *et al*., Macrophage expression and prognostic significance of the long pentraxin PTX3 in COVID-19. *Nat. Immunol.* **22**, 19–24 (2021). doi: 10.1038/s41590-020-00832-x; pmid: 33208929
66. C. Phetsouphanh *et al*., Immunological dysfunction persists for 8 months following initial mild-to-moderate SARS-CoV-2 infection. *Nat. Immunol.* **23**, 210–216 (2022). doi: 10.1038/s41590-021-01113-x; pmid: 35027728
67. B. Afzali, M. Noris, B. N. Lambrecht, C. Kemper, The state of complement in COVID-19. *Nat. Rev. Immunol.* **22**, 77–84 (2022). doi: 10.1038/s41577-021-00665-1; pmid: 34912108
68. B. Yan *et al*., SARS-CoV-2 drives JAK1/2-dependent local complement hyperactivation. *Sci. Immunol.* **6**, eabg0833 (2021). doi: 10.1126/sciimmunol.abg0833; pmid: 33827897
69. J. Couzin-Frankel, Clues to Long Covid. *Science* **376**, 1261–1265 (2022). doi: 10.1126/science.add4297; pmid: 35709281
70. C. Wang *et al*., Long COVID: The nature of thrombotic sequelae determines the necessity of early anticoagulation. *Front. Cell. Infect. Microbiol.* **12**, 861703 (2022). doi: 10.3389/fcimb.2022.861703; pmid: 35449732
71. P. Ambrosino *et al*., Persistent endothelial dysfunction in post-acute COVID-19 syndrome: a case-control study. *Biomedicines* **9**, 957 (2021). doi: 10.3390/biomedicines9080957; pmid: 34440161
72. M. H. Lee *et al*., Neurovascular injury with complement activation and inflammation in COVID-19. *Brain* **145**, 2555–2568 (2022). doi: 10.1093/brain/awac151; pmid: 35788639
73. A. Constantinescu-Bercu *et al*., Analysis of thrombogenicity under flow reveals new insights into the prothrombotic state

of patients with post-COVID syndrome. *J. Thromb. Haemost.* **21**, 94–100 (2023). doi: 10.1016/j.jtha.2022.10.013; pmid: 36695401
74. E. Pretorius *et al*., Persistent clotting protein pathology in Long COVID/Post-Acute Sequelae of COVID-19 (PASC) is accompanied by increased levels of antiplasmin. *Cardiovasc. Diabetol.* **20**, 172 (2021). doi: 10.1186/s12933-021-01359-7; pmid: 34425843
75. Y. Xie, E. Xu, B. Bowe, Z. Al-Aly, Long-term cardiovascular outcomes of COVID-19. *Nat. Med.* **28**, 583–590 (2022). doi: 10.1038/s41591-022-01689-3; pmid: 35132265
76. ARDS Definition Task Force, Acute respiratory distress syndrome: The Berlin Definition. *JAMA* **307**, 2526–2533 (2012). pmid: 22797452
77. A. Nalbandian *et al*., Post-acute COVID-19 syndrome. *Nat. Med.* **27**, 601–615 (2021). doi: 10.1038/s41591-021-01283-z; pmid: 33753937
78. B. B. Sun *et al*., Genomic atlas of the human plasma proteome. *Nature* **558**, 73–79 (2018). doi: 10.1038/s41586-018-0175-2; pmid: 29875488
79. S. A. Williams *et al*., Plasma protein patterns as comprehensive indicators of health. *Nat. Med.* **25**, 1851–1857 (2019). doi: 10.1038/s41591-019-0665-2; pmid: 31792462
80. V. Demichev, C. B. Messner, S. I. Vernardis, K. S. Lilley, M. Ralser, DIA-NN: Neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **17**, 41–44 (2020). doi: 10.1038/s41592-019-0638-x; pmid: 31768060
81. M. Emmenegger *et al*., Continuous population-level monitoring of SARS-CoV-2 seroprevalence in a large European metropolitan region. *iScience* **26**, 105928 (2023). doi: 10.1016/j.isci.2023.105928; pmid: 36619367
82. E. K. L. Chan *et al*., Report of the First International Consensus on Standardized Nomenclature of Antinuclear Antibody HEp-2 Cell Patterns 2014-2015. *Front. Immunol.* **6**, 412 (2015). doi: 10.3389/fimmu.2015.00412; pmid: 26347739
83. C. Pou *et al*., The repertoire of maternal anti-viral antibodies in human newborns. *Nat. Med.* **25**, 591–596 (2019). doi: 10.1038/s41591-019-0392-8; pmid: 30886409
84. H. Li *et al*., The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009). doi: 10.1093/bioinformatics/btp352; pmid: 19505943
85. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012). doi: 10.1038/nmeth.1923; pmid: 22388286
86. M. J. Mina *et al*., Measles virus infection diminishes preexisting antibodies that offer protection from other pathogens. *Science* **366**, 599–606 (2019). doi: 10.1126/science.aay6485; pmid: 31672891
87. S. Chevrier *et al*., Compensation of signal spillover in suspension and imaging mass cytometry. *Cell Syst.* **6**, 612–620.e5 (2018). doi: 10.1016/j.cels.2018.02.010; pmid: 29605184
88. J. H. Levine *et al*., Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell* **162**, 184–197 (2015). doi: 10.1016/j.cell.2015.05.047; pmid: 26095251
89. G. X. Y. Zheng *et al*., Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017). doi: 10.1038/ncomms14049; pmid: 28091601
90. Satija Lab, PBMC CITE-seq reference, version 1.0.0, Zenodo (2023); https://doi.org/10.5281/zenodo.7779017.
91. Y. Hao *et al*., Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021). doi: 10.1016/j.cell.2021.04.048; pmid: 34062119
92. G. Korotkevich *et al*., Fast gene set enrichment analysis. bioRxiv 060012 [Preprint] (2016); https://doi.org/10.1101/060012.
93. F. Pedregosa *et al*., Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
94. S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions. arXiv:1705.07874 [cs.AI] (2017).
95. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995). doi: 10.1111/j.2517-6161.1995.tb02031.x
96. M. Friendly, Corrgrams: Exploratory displays for correlation matrices. *Am. Stat.* **56**, 316–324 (2002). doi: 10.1198/000313002533
97. A. Kuznetsova, P. B. Brockhoff, R. H. B. Christensen, lmerTest Package: Tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017). doi: 10.18637/jss.v082.i13
98. H. Wickham *et al*., Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686 (2019). doi: 10.21105/joss.01686
99. S. Seabold, J. Perktold, "Statsmodels: Econometric and Statistical Modeling with Python," SciPy 2010, 9th Python in Science Conference, Austin, Texas, 28 to 30 June 2010, pp. 92–96; https://doi.org/10.25080/Majora-92bf1922-011.
100. E. von Elm *et al*., Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *BMJ* **335**, 806–808 (2007). doi: 10.1136/bmj.39335.541782.AD; pmid: 17947786
101. C. Cervia-Hasler, S. Brueningk, T. Hoch, Persistent complement dysregulation with signs of thrombo-inflammation in active Long Covid, version 1, Mendeley Data (2024); https://doi.org/10.17632/dvf6yvrg4x.1.
102. S. Brueningk, BorgwardtLab/LongCOVID: First Release, version v0, Zenodo (2023); https://doi.org/10.5281/zenodo.10022438.
103. Mount Sinai COVID-19 Biobank data, Synapse (2022); https://www.synapse.org/#!Synapse:syn35874390/.

ACKNOWLEDGMENTS

We thank J. Muri, D. Robbiani, V. Cecchinato, M. Uguccioni, A. Guaita, E. Bächli, A. Rudiger, M. Stüssi-Helbling, and L. Huber; all members of the Mount Sinai COVID-19 Biobank Team; the SomaLogic team, including K. Jenko, L. Vahlkamp, and E. Zepko; and the Functional Genomics Center Zurich, including N. Falko; and all members of the Boyman laboratory for their support, particularly A. Maier. Details about institutional review board approvals and informed patient consent can be found in the materials and methods. **Funding:** Swiss National Science Foundation (NRP 78 Implementation Programme to C.C.-H. and O.B.; 4078P0-198431 to O.B. and J.N.; 310030-200669 and 310030-212240 to O.B.), Clinical Research Priority Program CYTIMM-Z of UZH (to O.B.), Pandemic Fund of UZH (to O.B.), Innovation grants of University Hospital Zurich (USZ) (to O.B., A.A., and M.E.), Digitalization Initiative of Zurich Higher Education Institutions Rapid-Action Call 2021.1_RAC_ID_34 (to C.C.-H.), Swiss Academy of Medical Sciences (SAMW) fellowships (323530-191220 to C.C.-H.; 323530-191230 to Y.Z.; 323530-177975 to S.A.), Young Talents in Clinical Research project grant (YTCR 08/20) by SAMW and Bangerter Foundation (to M.E.R.), Filling the Gap Program of UZH (to M.E.R.), UZH Postdoc Grant (FK-22-053 to D.M.), Botnar Research Centre for Child Health Postdoctoral Excellence Programme (#PEP-2021-1008 to S.C.B.), and European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement (813533 to K.B., B.F., and G.M.; 801076 to T.B.), Driver Grant 2017DRI17 of Swiss Personalized Health Network (to A.A. and M.E.), NOMIS Foundation, Schwyzer Winiker Stiftung and Baugarten Stiftung (coordinated by USZ Foundation, USZF27101, to A.A. and M.E.), and the Swiss canton of Grisons (to C.B.M.). **Author contributions:** Conceptualization: C.C.-H. and O.B. Methodology: C.C.-H., S.C.B., K.B., and O.B. Investigation: C.C.-H., S.C.B., T.H., B.F., G.M., R.C.T., L.C., R.M., P.W., M.E., P.T., Y.Z., M.Po., D.M., T.B., S.C.-H., S.A., M.M., A.W.C., M.Pu., P.B., J.N., A.A., M.E.R., C.B.M., N.D.B., K.B., and O.B. Visualization: C.C.-H. and O.B. Funding acquisition: C.C.-H., Y.Z., S.A., M.E.R., J.N., and O.B. Project administration: S.C. and M.E.R. Supervision: K.B. and O.B. Writing: C.C.H. and O.B. **Competing interests:** C.B.M. works as an adviser for Eliptica Ltd. C.C.-H., S.C.B., B.F., G.M., K.B., and O.B. are listed as inventors on a patent application on biomarkers in Long Covid. The other authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials or have been deposited at Mendeley (*101*) or Zenodo (*102*). Replication data were obtained directly from the authors of previously published work (*24*). Researchers can access the Mount Sinai COVID-19 Biobank data on Synapse (*103*) by registering for a free account (Synapse project ID syn35874390). There are no restrictions on access or use of these datasets. This research was funded in whole or in part by the Swiss National Science Foundation (4078P0-198431, 310030-200669, 310030-212240) and the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement (813533 and 801076), cOAlition S organizations. The author will make the Author Accepted Manuscript (AAM) version available under a CC BY public copyright license.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adg7942
Figs. S1 to S8
Tables S1 to S5
STROBE Checklist
MDAR Reproducibility Checklist
Data S1 to S9

Submitted 21 January 2023; resubmitted 1 September 2023
Accepted 24 November 2023
10.1126/science.adg7942

RESEARCH ARTICLE

RADIO ASTRONOMY

A pulsar in a binary with a compact object in the mass gap between neutron stars and black holes

Ewan D. Barr[1]*†, Arunima Dutta[1]*†, Paulo C. C. Freire[1], Mario Cadelano[2,3], Tasha Gautam[1], Michael Kramer[1], Cristina Pallanca[2,3], Scott M. Ransom[4], Alessandro Ridolfi[1,5], Benjamin W. Stappers[6], Thomas M. Tauris[1,7], Vivek Venkatraman Krishnan[1], Norbert Wex[1], Matthew Bailes[8,9], Jan Behrend[1], Sarah Buchner[10], Marta Burgay[5], Weiwei Chen[1], David J. Champion[1], C.-H. Rosie Chen[1], Alessandro Corongiu[5], Marisa Geyer[10,11]‡, Y. P. Men[1], Prajwal Voraganti Padmanabh[1,12,13], Andrea Possenti[5]

Some compact objects observed in gravitational wave events have masses in the gap between known neutron stars (NSs) and black holes (BHs). The nature of these mass gap objects is unknown, as is the formation of their host binary systems. We report pulsar timing observations made with the Karoo Array Telescope (MeerKAT) of PSR J0514–4002E, an eccentric binary millisecond pulsar in the globular cluster NGC 1851. We found a total binary mass of 3.887 ± 0.004 solar masses ($M_\odot$), and multiwavelength observations show that the pulsar's binary companion is also a compact object. The companion's mass (2.09 to 2.71 $M_\odot$, 95% confidence interval) is in the mass gap, indicating either a very massive NS or a low-mass BH. We propose that the companion formed in a merger between two earlier NSs.

Globular clusters (GCs) are dense, gravitationally bound stellar clusters. Observations show that GCs host a large number of low-mass x-ray binaries (LMXBs), consisting of a compact object accreting material from a donor star in a binary system. LMXBs are ~10^3 times more abundant per unit of stellar mass in GCs than in the disk of the Milky Way (along the Galactic plane) (*1*). This is due to the high stellar densities at the center of GCs, which increase the rate of exchange encounters in which neutron stars (NSs) acquire low-mass main-sequence (MS) binary companions. The MS stars evolve until they expand and start transferring mass to the NS, at which point an LMXB is formed.

These x-ray binaries are expected to produce millisecond pulsars (MSPs; radio-emitting NSs with spin periods P < 10 ms) in almost circ orbits around low-mass companions (*2*, *3*). There are a total of 305 pulsars known in 40 GCs (*4*), the vast majority of which are MSPs. Most of the systems in GCs are similar to the MSP population found in the Galactic plane, although their orbital eccentricities are often higher, which is thought to be a result of close encounters with other stars (*5*).

In GCs with the densest cores, any particular star or MSP is likely to experience multiple exchange encounters over its lifetime (*6*). A possible outcome is the exchange of a low-mass companion of an MSP for either a massive white dwarf (WD) or another NS, resulting in a massive, eccentric MSP binary (*7*–*10*). Observing such systems allows their component masses to be measured and can test theories of gravity (*11*). The same exchange process could also produce an MSP–black hole (BH) binary system (*12*).

The millisecond pulsar binary PSR J0514–4002E

A survey searching for MSPs in GCs (*10*) has been carried out with the MeerKAT radio telescope array in South Africa (*13*, *14*). The results of the survey (*15*) included 13 MSPs in NGC 1851, a GC located in the Southern constellation of Columba (*16*). These include three massive, eccentric MSP binaries: PSR

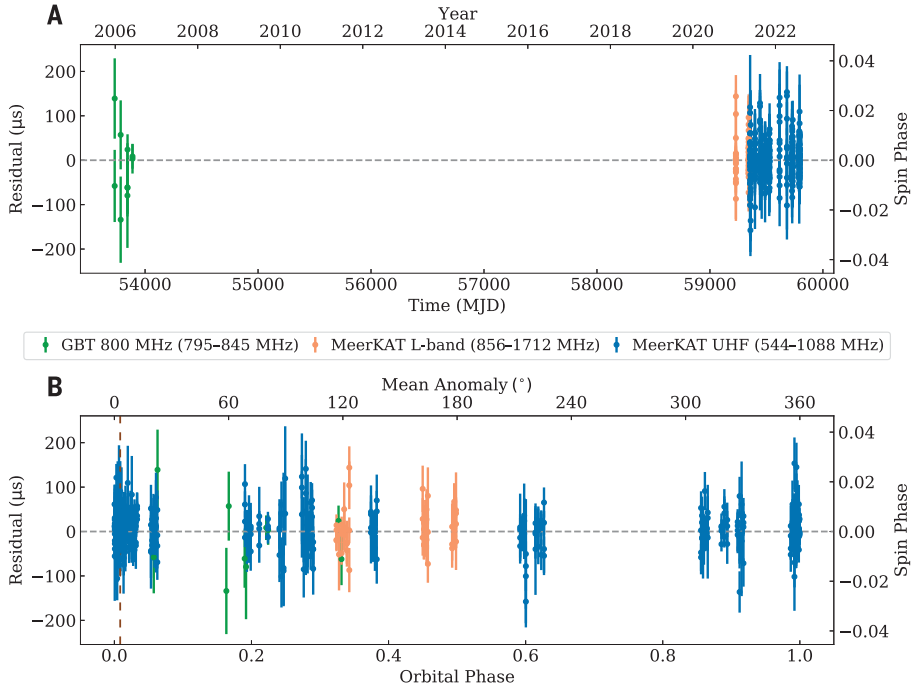


Fig. 1. Timing model fitting residuals for PSR J0514–4002E. (**A** and **B**) Residuals between the observed ToAs and the timing model presented in Table 1 are shown as (A) a function of observing epoch, in Modified Julian Date (MJD), and (B) orbital phase. An orbital phase of 0 corresponds to periastron, and superior conjunction occurs at an orbital phase of 0.008 (indicated with the brown dashed vertical line). The vertical error bars indicate 1σ uncertainties. The green points are data taken with the 800-MHz receiver on the GBT. Orange and blue points are data taken with the L-Band and UHF receivers of MeerKAT, respectively.

[1]Max-Planck-Institut für Radioastronomie, 53121 Bonn, Germany. [2]Dipartimento di Fisica e Astronomia "Augusto Righi," Università degli Studi di Bologna, 40129 Bologna, Italy. [3]Osservatorio di Astrofisica e Scienze dello Spazio di Bologna, Istituto Nazionale di Astrofisica, I-40129 Bologna, Italy. [4]National Radio Astronomy Observatory, Charlottesville, VA 22903, USA. [5]Osservatorio Astronomico di Cagliari, Istituto Nazionale di Astrofisica, I-09047 Selargius, Italy. [6]Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, The University of Manchester, Manchester M13 9PL, UK. [7]Department of Materials and Production, Aalborg University, DK-9220 Aalborg Øst, Denmark. [8]Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia. [9]Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav), Swinburne University of Technology, Hawthorn, VIC 3122, Australia. [10]South African Radio Astronomy Observatory, Observatory, Cape Town 7925, South Africa. [11]Department of Astronomy, University of Cape Town, Rondebosch, Cape Town 7700, South Africa. [12]Max Planck Institute for Gravitational Physics (Albert Einstein Institute), D-30167 Hannover, Germany. [13]Leibniz Universität Hannover, D-30167 Hannover, Germany.
***Corresponding author. Email: ebarr@mpifr-bonn.mpg.de (E.D.B.); adutta@mpifr-bonn.mpg.de (A.D.)**
†These authors contributed equally to this work.
‡Present address: Department of Mathematics and Applied Mathematics, University of Cape Town, Cape Town 7700, South Africa.

J0514−4002A (*8*, *17*), PSR J0514−4002D, and PSR J0514–4002E (*16*).

The latter MSP has a spin period (P) of 5.6 ms, an orbital period (P_b) of 7.44 days, and an orbital eccentricity (e) of 0.71 (*16*). The time required for light to cross the projected semimajor axis of the pulsar's orbit ($x \equiv a_p \sin i/c$, where a_p is the semimajor axis of the pulsar orbit, i is the orbital inclination, and c is the speed of light in vacuum) is 27.8 s (*16*). The mass function (f) is thus

$$f(m_p, m_c) \equiv \frac{(m_c \sin i)^3}{(m_p + m_c)^2} = 4\pi^2 \frac{c^3}{G} \frac{x^3}{{P_b}^2} = 0.41672 \pm 0.00022\ M_\odot \quad (1)$$

where m_c is the mass of the companion, m_p is the mass of the pulsar, and G is the gravitational constant. The unit $M_\odot$ is the mass of the Sun; we used the nominal solar mass adopted by the International Astronomical Union (*18*). All uncertainties are confidence intervals (CIs) corresponding to 68.3% confidence level, unless otherwise stated. Assuming the maximum possible inclination (an edge-on orbit, $i = 90°$) and $m_p \geq 1.17\ M_\odot$—corresponding to both the lowest NS mass measured (*19*) and a theoretical lower limit (*20*)—the mass function alone indicates $m_c \geq 1.40\ M_\odot$.

Radio timing observations

To determine the spin, astrometric, and orbital parameters for all the pulsars in NGC 1851, we conducted 24 observations of this GC using MeerKAT. The observations used either the L-band (856 to 1712 MHz) or ultrahigh frequency (UHF; 544 to 1088 MHz) receivers (*10*) and were performed between January 2021 and August 2022. Data acquisition and initial reduction were performed by using the Pulsar Timing User Supplied Equipment (PTUSE) instrument (*21*). We analyzed the resulting times of arrival (ToAs) of the pulsed signal to determine an initial phase-coherent timing model for PSR J0514–4002E (*22*). Using this model, we recovered the previously undetected signals from this pulsar in six archival observations of NGC 1851 made with the 800 MHz (795 to 845 MHz) and S-band (1.73 to 2.60 GHz) receivers on the Robert C. Byrd Green Bank Telescope (GBT) (*23*) between December 2005 and August 2006. We combined the MeerKAT and GBT ToAs (*22*) and refitted the timing parameters to determine a refined timing model. The results are listed in Table 1, and the fitting residuals between this model and the observed ToAs are shown in Fig. 1.

The timing model includes a precise measurement of the binary's rate of periastron advance $\dot{\omega} = 0.03468 \pm 0.00003°\ \text{year}^{-1}$. We obtained a consistent value (but slightly higher uncertainty) when we considered the MeerKAT data alone (fig. S2). The nonzero $\dot{\omega}$ is due to relativistic effects, with other contributions being small compared with the measurement uncertainty (supplementary text), and its value implies a high total system mass. We therefore performed an additional dense observing campaign (*22*) to search for Shapiro delay, a relativistic effect that causes light propagation delay in the system (*24*). The longer time span provided by the GBT data also enabled us to search for the Einstein delay, another relativistic effect caused by the varying time dilation experienced by the pulsar at different orbital phases. We did not detect either the Shapiro delay or Einstein delay in our timing data, setting 95% confidence upper limits of $h_3 < 1.48$ μs and $\gamma_E < 25$ ms, where h_3 is the orthometric amplitude of the Shapiro delay (*25*) and γ_E is the Einstein delay parameter. These nondetections further constrain the companion mass.

Table 1. Timing model for PSR J0514–4002E. The model was derived from the MeerKAT and GBT data, with the single GBT observation at S-band excluded during fitting (*22*). Because the proper motion and parallax cannot be measured from the timing data, we adopted the bulk proper motion and parallax for NGC 1851 derived from HST and Gaia observations (*28*). Values are reported in barycentric dynamical time (TDB), and uncertainties are 68.3% CI.

Dataset and model fitting quality	
Observation span	MJD 53731 to 59793
Number of ToAs	476
Weighted root mean square residual	28.41 μs
Reduced χ^2 value	1.019
Degrees of freedom	458
Fixed quantities	
Reference epoch	MJD 59400
Proper motion in right ascension, μ_α	2.128 milli–arc sec year^{-1}
Proper motion in declination, μ_δ	−0.646 milli–arc sec year^{-1}
Parallax	0.0858 milli–arc sec
Measured quantities	
Right ascension, α (J2000 equinox)	$05^h14^m06^s.73709 \pm 0.00017$
Declination, δ (J2000 equinox)	$-40°02'48''.0556 \pm 0.0014$
Pulse frequency, ν	178.70074989725 ± 0.000000000085 Hz
First derivative of pulse frequency, $\dot{\nu}$	$(-6.1727 \pm 0.0042) \times 10^{-15}$ Hz s^{-1}
Second derivative of pulse frequency, $\ddot{\nu}$	$(7.3 \pm 2.2) \times 10^{-26}$ Hz s^{-2}
Dispersion measure, DM	51.93061 ± 0.00057 pc cm^{-3}
Orbital period, P_b	7.4478966582 ± 0.0000000072 days
Projected semi-major axis of orbit, x	27.8192 ± 0.0050 s
Orbital eccentricity, e	0.70793232 ± 0.00000085
Epoch of periastron, T_0	MJD 59361.29117138 ± 0.00000037
Longitude of periastron, ω_0	65.317 ± 0.022°
Rate of advance of periastron, $\dot{\omega}$	0.034676 ± 0.000031° year^{-1}
Rate of variation of the orbital period, $\dot{P}_b$	$(18.1 \pm 5.6) \times 10^{-12}$ s s^{-1}
Einstein delay, γ_E	0.0111 ± 0.0084 s
Orthometric amplitude of Shapiro delay, h_3	0.02 ± 0.91 μs
Derived quantities	
Total mass*, M	3.8870 ± 0.0045 $M_\odot$
Pulse period, P	0.005595947418100 ± 0.000000000000027 s
First derivative of pulse period, $\dot{P}$	$(1.9330 \pm 0.0013) \times 10^{-19}$ s s^{-1}

*This assumes that the observed $\dot{\omega}$ is due to relativistic effects (supplementary text).

Near-ultraviolet and optical observations

If the ≥1.40 $M_\odot$ companion of PSR J0514–4002E were a MS star, it should be detectable at optical wavelengths. We searched for an optical source using archival Hubble Space Telescope (HST) observations with the Wide Field Camera 3 (WFC3) in the F275W and F336W filters. This filter combination is particularly sensitive to blue stars such as blue straggler stars (BSSs) and WDs (*26*, *27*). Bright BSSs are common in GCs, formed through stellar collisions or mass-transfer in a binary system; they can in principle have masses compatible with those predicted for the companion of PSR J0514–4002E.

No optical source was detected at the position of PSR J0514–4002E (Fig. 2, A and B). The closest stellar source to the pulsar position is a star with a color-magnitude position consistent with those of BSSs (Fig. 2C), which is offset

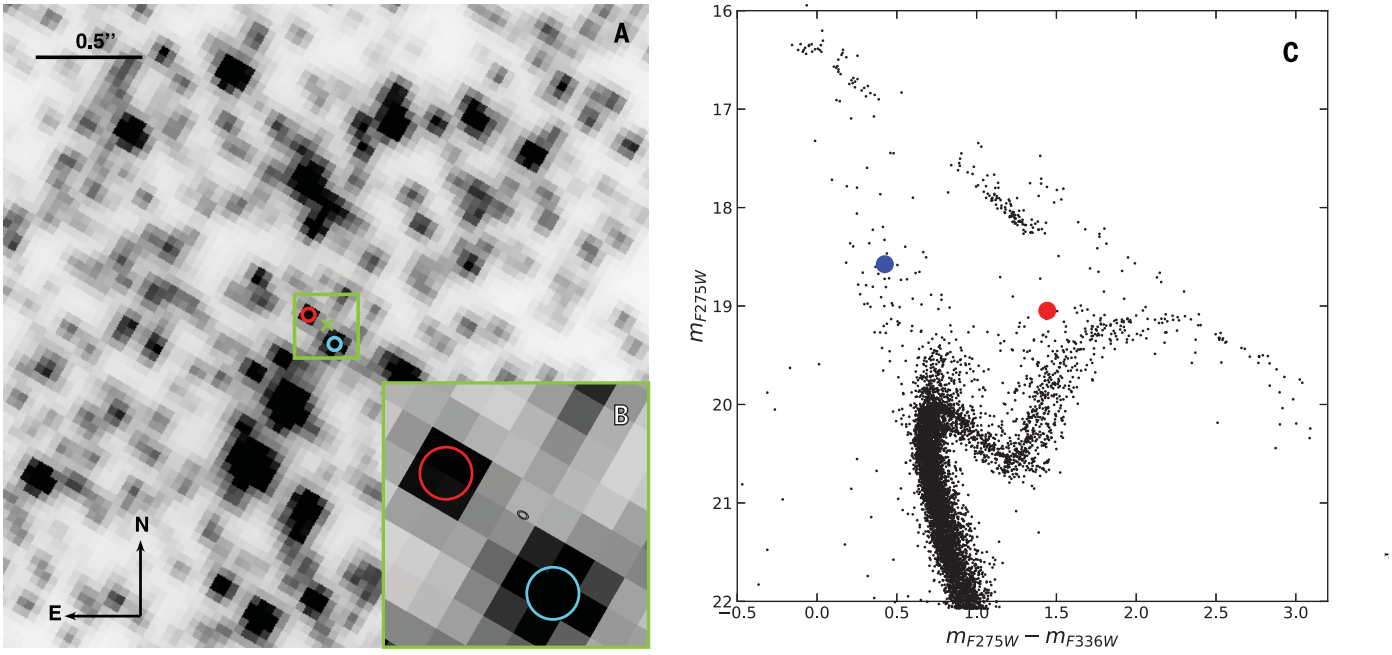


Fig. 2. Stars near the position of PSR J0514–4002E. (**A**) HST image in the F275W filter of the 4″ by 4″ region surrounding the radio timing position of PSR J0514–4002E (green cross). The blue and red circles indicate the two stars discussed in the main text; the circle sizes are twice the astrometric uncertainty. (**B**) A 0.3″ by 0.3″ zoom of the green box in (A). The radio timing position is indicated with a black ellipse with size and orientation equal to the 95% CI on the pulsar's position, relative to the International Celestial Reference System. (**C**) Color-magnitude diagram of NGC 1851 derived from the HST images. *m* is apparent magnitude, with subscripts indicating the filter used. The black dots indicate all the cluster stars within the sampled field of view. The blue and red circles indicate the stars marked in (A) and (B), which we interpret as a BSS and red giant, respectively. The photometric uncertainty for both the magnitude and color of the two marked stars is smaller than the symbol size (*22*).

by 90 milli–arc sec, which is more than six times the astrometric precision (14 milli–arc sec). This excludes a physical association because the orbital separation between the pulsar and companion calculated from our timing model, located at the cluster distance of 11.66 kpc (*28*), is $<10^{-3}$ milli–arc sec. Assuming that the optical source is a hydrogen-burning star, we estimate its mass as ~1.2 $M_\odot$ (*22*), which is lower than the lower limit on m_c from the mass function. Another star, located 100 milli–arc secfrom the pulsar position, is a red giant. Red giant branch stars in old GCs such as NGC 1851 have masses of about 0.7 to 0.8 $M_\odot$ (*29*), which is also lower than the minimum companion mass. We therefore conclude that the companion of PSR J0514–4002E is not detectable in the HST images.

Total mass of the system

The nondetection of the companion of PSR J0514–4002E in the HST images implies that it must be a compact object. There is no measurable excess pulse dispersion around superior conjunction, which could be produced by ionized gas emanating from either a main sequence or a giant star companion (*22*). We therefore interpret the measured rate of periastron advance ($\dot{\omega}$) as being of purely relativistic origin, with negligible contributions from the spins of the binary components (supplementary text). Assuming general relativity (GR), the total mass of the system is then (*30*, *31*)

$$M \equiv m_p + m_c = \frac{c^3}{G}\left[\frac{\dot{\omega}}{3}\left(1-e^2\right)\right]^{3/2}\left(\frac{P_b}{2\pi}\right)^{5/2} = 3.887 \pm 0.004\ M_\odot \quad (2)$$

This is 1.0 $M_\odot$ more massive than PSR J1913+1102, the highest-mass double neutron star (DNS) system known in the Milky Way, which has a total mass of 2.8887 ± 0.0006 $M_\odot$ (*32*). It is also larger than the total mass of GW190425, the heaviest DNS merger detected by using gravitational waves, at >99.5% probability [(*33*), their figure 5].

Nature of the companion

Combining the measured total mass and mass function, the limiting case of an edge-on orbit (i = 90°) sets $m_p \leq 2.04\ M_\odot$ and $m_c \geq M_\odot$ (Fig. 3). Smaller inclination angles (i < 90°) would imply a smaller m_p and larger m_c. This companion mass is far too high for a WD; the upper mass limit for a rigidly rotating WD is about 1.47 $M_\odot$ (*34*). Adopting the minimum pulsar mass discussed above, $m_p \geq 1.17\ M_\odot$, we constrain $i \geq 42.9°$ and $m_c \leq 2.71\ M_\odot$.

We did not detect additional relativistic effects that could allow the individual masses to be determined. However, our upper limits on the Shapiro delay and the Einstein delay provide additional constraints on the masses and orbital inclination. To quantify these, we performed Bayesian estimation of the component masses based on the goodness of fit of the model to the observational data (χ^2) calculated over a grid of total masses and orbital inclinations (*22*). The variation of the orbital period is contaminated by the acceleration of the system within the GC. We assumed that all other relativistic effects are as predicted by GR and adopted $m_p \geq 1.17\ M_\odot$ as above (*22*). We found Bayesian posteriors of $M = 3.887 \pm 0.004\ M_\odot$, $m_p = 1.53^{+0.18}_{-0.20}\ M_\odot$, $m_c = 2.35^{+0.20}_{-0.18} M_\odot$, and $i = 52^{+6°}_{-5}$ (all median values with 68.3% confidence limits) (*22*). This value of M is identical to that calculated from Eq. 2, but through a different method. The 95% probability limits are $i < 62°$, $m_p < 1.79\ M_\odot$, and therefore, $m_c > 2.09\ M_\odot$.

The companion mass is therefore likely to be in the mass gap for compact objects (*35*). It is higher than the largest precisely measured pulsar masses, $m_p = 2.08 \pm 0.07\ M_\odot$ for PSR J0740+6620 (*36*) and $m_p = 2.01 \pm 0.04\ M_\odot$ for PSR J0348+0432 (*37*). It is simultaneously below the observed minimum mass of BHs in

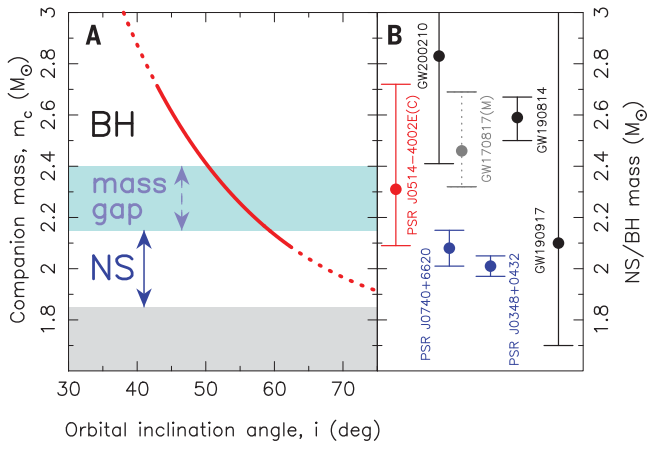


Fig. 3. Companion mass of PSR J0514–4002E. (**A**) The derived companion mass at different orbital inclinations for the PSR J0514–4002E system. The solid red curve indicates solutions within the 95% CI; the dotted segments indicate lower and higher masses that are consistent with the mass function but excluded by our Bayesian model or the adopted minimum pulsar mass $m_p \geq 1.17\ M_\odot$, respectively. The gray shaded region is ruled out by the mass function and the total mass (Eqs. 1 and 2). Depending on the (unknown) NS equation of state (*49*, *50*), the light blue–shaded mass gap corresponds to either massive NSs or light BHs. (**B**) Inferred companion mass of PSR J0514–4002E (red) compared with the largest observed masses of radio pulsars (blue), low-mass components of gravitational-wave mergers (black), and the total postmerger remnant mass of GW170817 (gray, assuming no energy and mass loss after the inspiral, so an absolute upper limit). Source names followed by (C) or (M) indicate companion star mass and remnant mass of the merger product, respectively. The masses, uncertainties, and references are listed in table S2.

Galactic x-ray binaries, which is about 5 $M_\odot$ (*38*, *39*).

If the companion were a massive NS, it might also be a radio pulsar. We searched for radio pulsations from the companion, assuming the full allowed range of mass ratios, but did not detect any (*22*). We therefore cannot determine whether the companion is a massive NS or a low-mass BH.

Formation of the system

The combination of the location in a dense GC (where stellar exchange encounters often occur), the highly eccentric orbit, the fast spin of the pulsar, and the large companion mass indicates that the PSR J0514–4002E system is the product of a secondary exchange encounter. We propose that an earlier low-mass companion transferred mass to this pulsar, increasing the pulsar spin rate, before being replaced by the present high-mass companion in an exchange encounter. However, a more complicated evolution with multiple exchange encounters is also possible. We therefore cannot infer the nature of the companion from binary evolution models.

If the mass of the primary in PSR J0514–4002E is in the range of 1.25 to 1.55 $M_\odot$, which spans the four measurements of pulsar masses in GCs (*9*, *11*, *17*, *40*), then the corresponding value of m_c (2.34 to 2.63 $M_\odot$) overlaps with the range of masses of remnants from mergers of DNSs, such as the merger product of GW170817 (Fig. 3). We suggest that the companion could potentially have formed in such a merger event, before becoming part of the current PSR J0514–4002E system, regardless of whether it is an NS or BH. Although the probability of DNS mergers is low in GCs that have not undergone core collapse (*41*), NGC 1851 has a dense core (supplementary text) that makes a DNS merger in the progenitor of the PSR J0514–4002E system more probable. A DNS has been observed in a GC (PSR B2127+11C, in M15) with a calculated merger time of 217 million years (*11*), implying that merger remnants are likely to be present in GCs.

Our derived companion mass overlaps with the mass estimates derived from gravitational waves for the lighter components of the BH+BH or BH+NS merger candidates GW190814 (*42*), GW190917, and GW200210 (*43*). The lighter component of GW190814 has previously been interpreted as the product of an earlier merger (*44*) that later acquired a more massive BH companion through exchange encounters in GCs (and then merged in the GW190814 event).

If the companion of PSR J0514–4002E is a light BH formed in such a merger, it would acquire a spin parameter χ_c of 0.6 to 0.875 during the merger (*45*), where χ_c is the dimensionless BH spin angular momentum. An NS rotating at the maximum theoretical rate would have a similar χ_c immediately after merger (*46*), although we expect that this would decrease rapidly after formation because of electromagnetic torque. Assuming a magnetic field of 10^9 G, the spin parameter would become $\lesssim$0.3 (corresponding to the fastest known MSPs) after ~30 million years, so we regard a fast-spinning NS companion as unlikely.

A BH companion with $0.6 < \chi_c < 0.875$ would induce relativistic spin-orbit coupling, causing the orbital plane of the binary to precess around the total angular momentum vector, an effect known as Lense-Thirring precession. We calculated that Lense-Thirring precession would cause a variation of the projected semimajor axis of the pulsar's orbit ($\dot{x}$) of $\lesssim 1.7 \times 10^{-13}$ (supplementary text). This is slightly smaller than the effect size that would be detectable in our data; our 1σ uncertainty on the measured $\dot{x}$ is 2.0×10^{-13}. We therefore cannot differentiate between a NS and a BH companion.

REFERENCES AND NOTES

1. G. W. Clark, *Astrophys. J.* **199**, L143 (1975).
2. M. A. Alpar, A. F. Cheng, M. A. Ruderman, J. Shaham, *Nature* **300**, 728–730 (1982).
3. T. M. Tauris, E. P. J. van den Heuvel, *Physics of Binary Star Evolution. From Stars to X-ray Binaries and Gravitational Wave Sources* (Princeton Univ. Press, 2023).
4. P. C. C. Freire, Pulsars in globular clusters; https://www3.mpifr-bonn.mpg.de/staff/pfreire/GCpsr.html (accessed 10 November 2023).
5. E. S. Phinney, *Philos. Trans. Phys. Sci. Eng.* **341**, 39–75 (1992).
6. F. Verbunt, P. C. C. Freire, *Astron. Astrophys.* **561**, A11 (2014).
7. T. A. Prince, S. B. Anderson, S. R. Kulkarni, A. Wolszczan, *Astrophys. J.* **374**, L41 (1991).
8. P. C. Freire, Y. Gupta, S. M. Ransom, C. H. Ishwara-Chandra, *Astrophys. J.* **606**, L53–L56 (2004).
9. R. S. Lynch, P. C. C. Freire, S. M. Ransom, B. A. Jacoby, *Astrophys. J.* **745**, 109 (2012).
10. A. Ridolfi *et al*., *Mon. Not. R. Astron. Soc.* **504**, 1407–1426 (2021).
11. B. A. Jacoby *et al*., *Astrophys. J.* **644**, L113–L116 (2006).
12. K. Kremer *et al*., *Phys. Rev. Lett.* **120**, 191103 (2018).
13. J. Jonas, MeerKAT Team, in *Proceedings of MeerKAT Science: On the Pathway to the SKA, Stellenbosch, 25–27 May 2016* (PoS, 2016).
14. F. Camilo, *Nat. Astron.* **2**, 594 (2018).
15. E. D. Barr, TRAPUM Discoveries; http://www.trapum.org/discoveries/ (accessed 10 November 2023).
16. A. Ridolfi *et al*., *Astron. Astrophys.* **664**, A27 (2022).
17. A. Ridolfi, P. C. C. Freire, Y. Gupta, S. M. Ransom, *Mon. Not. R. Astron. Soc.* **490**, 3860–3874 (2019).
18. A. Prša *et al*., *Astron. J.* **152**, 41 (2016).
19. J. G. Martinez *et al*., *Astrophys. J.* **812**, 143 (2015).
20. Y. Suwa, T. Yoshida, M. Shibata, H. Umeda, K. Takahashi, *Mon. Not. R. Astron. Soc.* **481**, 3305–3312 (2018).
21. M. Bailes *et al*., *Publ. Astron. Soc. Aust.* **37**, e028 (2020).
22. Materials and methods are available as supplementary materials.
23. P. C. C. Freire, S. M. Ransom, Y. Gupta, *Astrophys. J.* **662**, 1177–1182 (2007).
24. I. I. Shapiro, *Phys. Rev. Lett.* **13**, 789–791 (1964).
25. P. C. C. Freire, N. Wex, *Mon. Not. R. Astron. Soc.* **409**, 199–212 (2010).
26. S. Raso *et al*., *Astrophys. J.* **895**, 15 (2020).
27. J. Chen *et al*., *Nat. Astron.* **5**, 1170–1177 (2021).
28. M. Libralato *et al*., *Astrophys. J.* **934**, 150 (2022).
29. D. Valcin, J. L. Bernal, R. Jimenez, L. Verde, B. D. Wandelt, *J. Cosmol. Astropart. Phys.* **2020**, 002 (2020).
30. H. P. Robertson, *Ann. Math.* **39**, 101 (1938).
31. J. H. Taylor, J. M. Weisberg, *Astrophys. J.* **253**, 908 (1982).
32. R. D. Ferdman *et al*., *Nature* **583**, 211–214 (2020).
33. B. P. Abbott *et al*., *Astrophys. J. Lett.* **892**, L3 (2020).
34. S. C. Yoon, N. Langer, *Astron. Astrophys.* **435**, 967–985 (2005).
35. T. B. Littenberg, B. Farr, S. Coughlin, V. Kalogera, D. E. Holz, *Astrophys. J. Lett.* **807**, L24 (2015).
36. E. Fonseca *et al*., *Astrophys. J. Lett.* **915**, L12 (2021).
37. J. Antoniadis *et al*., *Science* **340**, 448, 1233232 (2013).
38. C. D. Bailyn, R. K. Jain, P. Coppi, J. A. Orosz, *Astrophys. J.* **499**, 367–374 (1998).
39. F. Özel, D. Psaltis, R. Narayan, J. E. McClintock, *Astrophys. J.* **725**, 1918–1927 (2010).
40. A. Corongiu *et al*., *Astron. Astrophys.* **671**, A72 (2023).
41. C. S. Ye *et al*., *Astrophys. J. Lett.* **888**, L10 (2020).
42. R. Abbott *et al*, *Astrophys. J. Lett.* **896**, L44 (2020).
43. The LIGO Scientific Collaboration *et al*, arXiv:2111.03606 [gr-qc] (2021).
44. W. Lu, P. Beniamini, C. Bonnerot, *Mon. Not. R. Astron. Soc.* **500**, 1817–1832 (2021).
45. S. Bernuzzi, *Gen. Relativ. Gravit.* **52**, 108 (2020).

46. K.-W. Lo, L.-M. Lin, *Astrophys. J.* **728**, 12 (2011).
47. A. Dutta, E. Barr, ewanbarr/NGC1851E_Additional_Materials. Zenodo (2023); https://doi.org/10.5281/zenodo.10123123.
48. E. Barr, A. Dutta, Radio data for PSR J0514-4002E. Keeper (2023); https://doi.org/10.17617/4.e3.
49. M. Shibata, E. Zhou, K. Kiuchi, S. Fujibayashi, *Phys. Rev. D* **100**, 023015 (2019).
50. E. R. Most, L. J. Papenfort, L. R. Weih, L. Rezzolla, *Mon. Not. R. Astron. Soc.* **499**, L82–L86 (2020).

ACKNOWLEDGMENTS

The MeerKAT telescope is operated by the South African Radio Astronomy Observatory (SARAO), which is a facility of the National Research Foundation, an agency of the Department of Science and Innovation. SARAO acknowledges the ongoing advice and calibration of GPSs by the National Metrology Institute of South Africa (NMISA) and the time space reference systems department of the Paris Observatory. Parts of this research were conducted by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav), through project CE170100004. MeerTime data are housed on the OzSTAR supercomputer at Swinburne University of Technology. The OzSTAR program receives funding in part from the Astronomy National Collaborative Research Infrastructure Strategy (NCRIS) allocation provided by the Australian Government. PTUSE was developed with support from the Australian SKA Office and Swinburne University of Technology. The authors also acknowledge Max-Planck-Institut für Radioastronomie funding to contribute to MeerTime infrastructure. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities. The Green Bank Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities. TRAPUM observations use the FBFUSE and APSUSE computing clusters for data acquisition, storage, and analysis. These clusters were funded, installed, and operated by the Max-Planck-Institut für Radioastronomie and the Max-Planck-Gesellschaft. This work has made use of data from the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC; https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement. **Funding:** A.C., A.P., A.R., and M.Bu. acknowledge resources from the research grant "iPeska" (principal investigator, A.P.) funded under the INAF national call Prin-SKA/CTA approved with the Presidential Decree 70/2016. A.P. and A.R. were supported by the Ministero degli Affari Esteri e della Cooperazione Internazionale–Direzione Generale per la Promozione del Sistema Paese–Progetto di Grande Rilevanza ZA18GR02. S.M.R. is a CIFAR Fellow and is supported by NSF Physics Frontiers Center awards 1430284 and 2020265. E.D.B., A.D., P.C.C.F., M.K., V.V.K., N.W., J.B., W.C., D.J.C., C.-H.R.C., Y.P.M., and P.V.P. acknowledge support from the Max-Planck Society. B.W.S. acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement 694745). C.-H.R.C. acknowledges support from the Deutsches Zentrum für Luft- und Raumfahrt (DLR) grant NS1 under contract 50 OR 2214. C.P. and M.C. acknowledge funding from Italian MIUR throughout the PRIN-2017 grant awarded to the Light-on-Dark project (principal investigator, F. R. Ferraro) through contract PRIN-2017K7REXT. M.G. acknowledges funding from the South African Research Chairs Initiative of the Department of Science and Technology and the National Research Foundation of South Africa. **Author contributions:** Conceptualization: A.D., A.P., B.W.S., M.Ba., M.K., N.W., P.C.C.F., S.M.R., and V.V.K. Data curation: A.D., A.R., E.D.B., M.Ba., M.G., P.V.P., S.M.R., Y.P.M., and V.V.K. Formal analysis: A.D., A.R., C.-H.R.C., C.P., D.J.C., M.Bu., M.C., N.W., P.C.C.F., T.G., and V.V.K. Funding acquisition: A.P., M.Ba., and M.K. Investigation: A.D., A.R., C.-H.R.C., E.D.B., M.Bu., S.B., and V.V.K. Methodology: A.D., E.D.B., M.Ba., P.C.C.F., and S.M.R. Project administration: A.D., A.P., A.R., B.W.S., E.D.B., M.Ba., M.Bu., M.K., and P.C.C.F. Resources: A.R., B.W.S., E.D.B., J.B., M.Ba., M.Bu., M.K., P.C.C.F., and C.-H.R.C. Software: A.R., D.J.C., E.D.B., M.Ba., M.G., P.C.C.F., P.V.P., S.B., W.C., and S.M.R. Supervision: A.P., B.W.S., D.J.C., E.D.B., M.K., N.W., P.C.C.F., S.M.R., and V.V.K. Validation: A.D., C.-H.R.C., N.W., P.C.C.F., T.M.T., and V.V.K. Visualization: A.D., E.D.B., M.C., P.C.C.F., T.M.T., and V.V.K. Writing – original draft: A.D., A.R., C.-H.R.C., C.P., E.D.B., M.C., N.W., P.C.C.F., T.M.T., and V.V.K. Writing – review and editing: A.C., A.D., A.P., B.W.S., C.-H.R.C., C.P., D.J.C., E.D.B., M.Ba., M.C., M.G., M.K., N.W., P.C.C.F., P.V.P., S.M.R., T.G., T.M.T., and V.V.K. **Competing interests:** There are no competing interests to declare. **Data and materials availability:** Our measured ToAs, timing model, and simulation code are available at https://github.com/ewanbarr/NGC1851E_Additional_Materials and archived at Zenodo (*47*). Folded pulsar timing archives from MeerKAT and the GBT, and dedispersed pulsar search mode observations from MeerKAT, are available at https://keeper.mpdl.mpg.de/d/a59d968eccf148239f1d and archived at the Max Planck Digital Library (*48*). The optical and near-ultraviolet data are available on the MAST archive at https://mast.stsci.edu/search/ui/#/hst, using the search terms "Target=NGC 1851", "Proposal ID=12311,13297", and "Filters / Gratings=F275W,F336W".

SUPPLEMENTARY MATERIAL

Submitted 15 December 2022; accepted 28 November 2023
10.1126/science.adg3005

ELECTROPHOTOCATALYSIS

Electrophotocatalytic perfluoroalkylation by LMCT excitation of Ag(II) perfluoroalkyl carboxylates

Brandon M. Campbell[1], Jesse B. Gordon[1], Elaine Reichert Raguram[1], Miguel I. Gonzalez[1], Kristopher G. Reynolds[1], Matthew Nava[1,2], Daniel G. Nocera[1]*

Molecular Ag(II) complexes are superoxidizing photoredox catalysts capable of generating radicals from redox-reticent substrates. In this work, we exploited the electrophilicity of Ag(II) centers in $[Ag(bpy)_2(TFA)][OTf]$ and $Ag(bpy)(TFA)_2$ (bpy, 2,2′-bipyridine; OTf, $CF_3SO_3^-$) complexes to activate trifluoroacetate (TFA) by visible light–induced homolysis. The resulting trifluoromethyl radicals may react with a variety of arenes to forge $C(sp^2)–CF_3$ bonds. This methodology is general and extends to other perfluoroalkyl carboxylates of higher chain length ($R_FCO_2^-$; R_F, CF_2CF_3 or $CF_2CF_2CF_3$). The photoredox reaction may be rendered electrophotocatalytic by regenerating the Ag(II) complexes electrochemically during irradiation. Electrophotocatalytic perfluoroalkylation of arenes at turnover numbers exceeding 20 was accomplished by photoexciting the Ag(II)–TFA ligand-to-metal charge transfer (LMCT) state, followed by electrochemical reoxidation of the Ag(I) photoproduct back to the Ag(II) photoreactant.

Photoredox methodologies have emerged as indispensable tools in organic synthesis, using photons as a clean and selective source of energy to drive challenging chemical reactions under relatively mild conditions (*1*). Of the various photoredox strategies, ligand-to-metal charge transfer (LMCT) processes have been exploited for the generation of reactive intermediates from inert substrates (*2*, *3*). As LMCT processes involve the transfer of electron density to the metal center, typically only electrophilic metals in higher-valent oxidation states engage in LMCT photochemistry. A variety of complexed metal ions, including Cu(II) (*4*, *5*), Co(III) (*6*, *7*), Fe(III) (*8*, *9*), Ni(III) (*10*–*12*), V(V) (*13*), and Ce(IV) (*14*–*16*) centers, photochemically generate open-shell intermediates upon LMCT excitation (Fig. 1), including chlorine, azidyl, alkoxy, acyl, and alkyl radicals generated from the homolysis of M–Cl (*8*, *11*, *12*, *16*), $–N_3$ (*5*, *9*), –OR (*15*), –C(O)R (*7*), and –R (R, alkyl group) (*6*) bonds. In addition, photolysis of $M–O_2CR$ bonds is known to induce decarboxylation of the generated carboxyl radical, which can provide a source of alkyl or aryl radicals (*4*, *14*, *17*, *18*). Noticeably absent from the palette of electrophilic metals is Ag(II). Although Ag(I) is a strong oxidant [$E°$ = 0.799 V versus normal hydrogen electrode (NHE)], Ag(II) with its d^9 electronic configuration wields even greater oxidative power ($E°$ = 1.980 V versus NHE) (*19*) by virtue of the hole in its 4d subshell (*20*). As schematically shown in Fig. 1, we envisioned that filling this hole by visible light excitation of a LMCT transition would position the underused Ag(II) metal center as a potent photooxidant for the activation of challenging substrates.

We targeted trifluoroacetate (TFA) because it is an ideal source of the pharmaceutically relevant trifluoromethyl group (*21*), which imparts dramatically enhanced pharmacokinetic properties to drug molecules (*22*). As compared with common CF_3 sources, such as the Ruppert–Prakash or Umemoto reagents (*23*), TFA is an ideal CF_3 source owing to its low cost and high annual production (*24*). However, the demanding oxidation potential of TFA [>2.2 V versus saturated calomel electrode (SCE) (*25*)] necessitates forcing conditions for the liberation of its CF_3 group. Thus, previous methods for the decarboxylation of TFA have required either high-energy ultraviolet (UV) irradiation (TiO_2 catalyst, λ_{exc} < 365 nm) (*26*, *27*), high temperatures [T > 120°C with Ag salts (*28*) or T > 140°C with Cu salts (*29*)], or harsh chemical oxidants such as XeF_2 (*30*). Other strategies rely on the preactivation of TFA or trifluoroacetic anhydride with exogenous reactants, such as pyridine *N*-oxides (*31*), sulfoxides (*32*), or hypervalent iodine reagents (*33*), to effectively lower the redox potential of TFA within range of traditional photocatalysts. Although functionalization of alkyl carboxylates with *N*-hydroxyphthalimide to form *N*-acyloxyphthalimides allows for the reductive fragmentation of the N–O bond under mild conditions to liberate CO_2 and an alkyl radical, when TFA is employed, the strong electron-withdrawing CF_3 group biases the reductive fragmentation toward the generation of an N-centered radical and the TFA anion leading to amination rather than the desired trifluoromethylation (*34*). Direct Kolbe-type electrolysis of TFA requires large operating cell potentials and the use of oxidatively resistant substrates; even with unactivated arenes such as benzene, competitive electrochemical oxidation of the arene leads to trifluoroacetoxylation side products (PhO_2CCF_3) in addition to the desired trifluoromethylated aromatics ($PhCF_3$) (*35*). LMCT excitation of well-defined Ag(II)–TFA complexes with visible light would represent a milder and more chemoselective pathway for harnessing the trifluoromethyl group from the redox-reticent TFA substrate.

In this work, we report the synthesis, isolation, and structural characterization of $[Ag(bpy)_2][OTf]_2$ and $[Ag(bpy)_n(O_2CCF_3)_m][OTf]_{2-m}$ complexes (bpy, 2,2′-bipyridine; OTf, $CF_3SO_3^-$). Photolysis of these Ag(II)–TFA complexes with light at excitation wavelengths (λ_{exc}) > 400 nm resulted in rapid decarboxylation to generate trifluoromethyl radicals with the concomitant extrusion of CO_2. We extended this methodology to perfluoroalkyl carboxylates of higher chain length ($R_FCO_2^-$; R_F, CF_2CF_3, or $CF_2CF_2CF_3$), demonstrating the generality of this platform. The suite of $Ag(II)–CO_2R_F$ complexes were shown to be competent perfluoroalkylating agents of arenes. By performing the photochemistry under an applied potential [E_{appl} > $E°$(Ag(II/I)], the system may be turned over catalytically. The electrophotocatalytic perfluoroalkylation of a variety

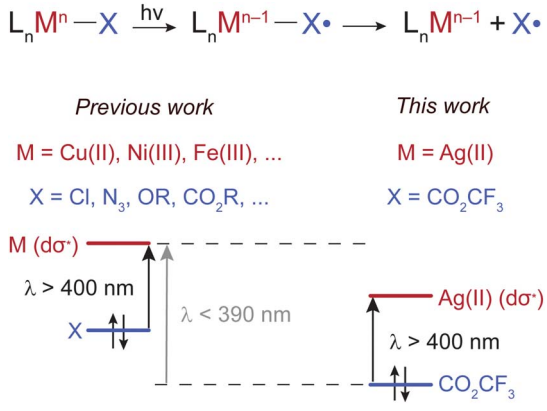


Fig. 1. LMCT chemistry of molecular compounds. Comparison of LMCT photochemistry of typical metal (M) centers versus that of the highly electrophilic Ag(II) center.

[1]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, United States.
[2]Department of Chemistry and Biochemistry, University of California at Los Angeles, Los Angeles, CA 90095, United States.
*Corresponding author. Email: dnocera@fas.harvard.edu

of (hetero)arenes was demonstrated with low Ag loadings and turnover numbers exceeding 20. Although electrophotocatalytic C–H trifluoromethylation has been accomplished previously by using the Langlois reagent (NaO_2SCF_3) as a specialized CF_3 source owing to its accessible oxidation potential (*36*), the described Ag(II)-mediated electrophotocatalytic system enabled the direct use of TFA. Whereas the work described here highlights the activation of TFA, our results suggest that molecular Ag(II) complexes are superior, yet underused superoxidizing catalysts that are capable of activating a variety of inert substrates with visible light.

Synthesis and characterization of Ag(II) complexes

Addition of AgO to a solution of bpy in aqueous triflic acid (50% v/v) gave a deep red solution that furnished $[Ag(bpy)_2][OTf]_2$ (**1**$[OTf]_2$) as a red microcrystalline powder upon precipitation with water. The tetrafluoroborate salt **1**$[BF_4]_2$ was prepared similarly by replacing triflic acid with terafluoroboric acid. The synthesis was performed on a multigram scale (8 mmol), and **1**$[OTf]_2$ was isolated in 64% yield. We note that **1**$[OTf]_2$ has been synthesized previously, requiring either a multistep electrochemical synthesis or one-pot procedures that gave impure product or low yields, and the compound has not been structurally characterized (*37*). Distinctive among other known Ag(II) molecular complexes is the solubility of **1**$[OTf]_2$ or **1**$[BF_4]_2$ in organic solvents, which allowed us to avoid aqueous solution and thus circumvent the oxidation of water. Dissolution of the isolable microcrystalline powder in acetonitrile (MeCN) followed by vapor diffusion of benzene resulted in the formation of deep red crystals. Crystal structure analysis revealed Ag(II) to reside in a tetragonally distorted bpy ligand field (Fig. 2A). **1**$[OTf]_2$ was characterized by an electron paramagnetic resonance (EPR) signal that was an axial doublet (fig. S1) and an effective magnetic moment of 1.97 μ_B, consistent with a d^9 ground state.

Treatment of a slurry of **1**$[OTf]_2$ in dichloromethane (CH_2Cl_2) with bis(triphenylphosphine) iminium trifluoroacetate ([PPN][TFA]) led to the introduction of TFA into the primary coordination sphere of Ag(II). **1**$[OTf]_2$ reacted with one equivalent of [PPN][TFA] in the presence of an additional equivalent of bpy to furnish a solid, which may be crystallized as black needles by layering CH_2Cl_2 solutions with pentane at −36°C. The crystal structure established $[Ag(bpy)_2(O_2CCF_3)][OTf]$ (**2**[OTf]) (Fig. 2B), obtained in 48% yield, as a rare example of Ag(II) in a distorted trigonal bipyramidal ligand field in which TFA binds through essentially a κ^1 coordination mode [d(Ag–O1) = 2.4814(15) Å; d(Ag–O2) = 2.7104(14) Å]. The bpy ligands in **2**[OTf] distorted to assume a nearly linear axial N1–Ag–N4 bond angle of 178.65(5)° to accommodate TFA in the primary coordination sphere. The relative similarity in the C21–O1 and C21–O2 bond distances in the TFA ligand point to delocalization of the negative charge on the carboxylate group. When a slurry of **1**$[OTf]_2$ in CH_2Cl_2 was treated with excess [PPN][TFA], the orange solid $Ag(bpy)(O_2CCF_3)_2$ (**3**) (Fig. 2C), in which one bpy ligand in **1**$[OTf]_2$ is substituted by two monodentate TFA ligands, was obtained in 59% yield. The neutral square planar coordination geometry draws similarities to Ag(4,4′-dimethyl-bpy)$(NO_3)_2$, in which nitrate anions serve as anionic ligands (*38*). **3** packs in the crystalline lattice as sets of symmetrically equivalent dimers, forming diamond core–like structures with weak apical Ag···O interactions. Both **2**[OTf] and **3** showed no resolvable ^{1}H nuclear magnetic resonance (NMR) signals for bpy owing to paramagnetic broadening. The compounds displayed axial EPR signals consistent with a paramagnetic doublet ground state (fig. S1), which was consistent with effective magnetic moments of 2.06 (**2**[OTf]) and 1.78 (**3**) μ_B. **3** was ^{19}F NMR silent, whereas **2**[OTf]

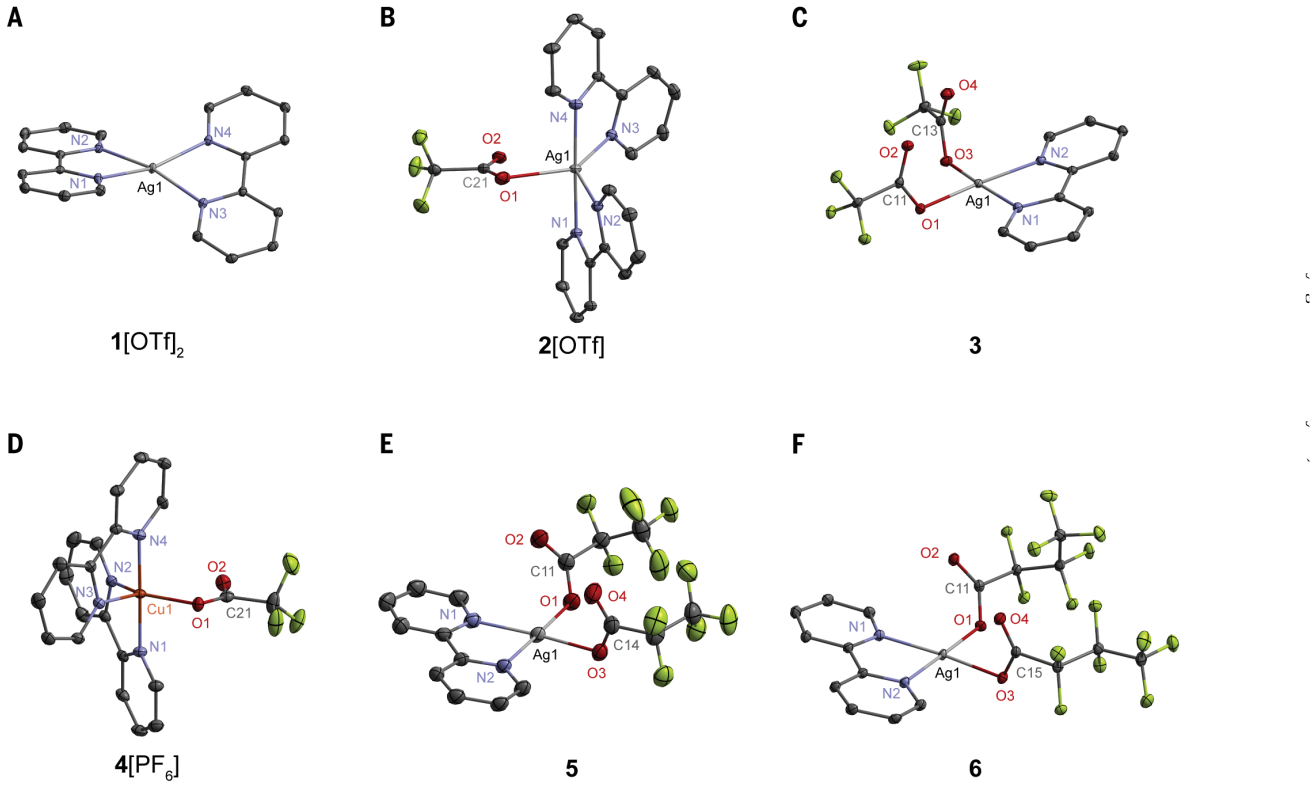


Fig. 2. Crystal structures of Ag(II) complexes and a Cu(II) congener. Atomic displacement parameter plot of (**A**) **1**$[OTf]_2$ at 15 K, (**B**) **2**[OTf] at 15 K, (**C**) **3** at 15 K, (**D**) **4**$[PF_6]$ at 100 K, (**E**) **5** at 100 K, and (**F**) **6** at 15 K, drawn at 50% probability level as ascertained by single-crystal x-ray diffraction analysis. Color scheme for structures **1**–**6**: Ag, light gray; Cu, dark orange; F, yellow green; O, red; N, blue; C, dark gray. Counterions, hydrogen atoms, and solvent molecules are omitted for clarity. Selected crystallographic bond metrics are provided in table S3.

displayed a single resonance corresponding to unbound triflate.

LMCT photogeneration of CF_3 radical from Ag(II)–TFA complexes

LMCT photochemistry of **2**[OTf] and ***3*** was accessed by using visible light. Photolysis (λ_{exc} = 450 nm) of **2**[OTf] in MeCN at 21.3°C resulted in the complete bleaching of all visible absorption features within 30 min, whereas ***3*** exhibited growth of a band at 700 nm that disappeared upon continued irradiation (Fig. 3, A and B). Under identical conditions, **1**$[OTf]_2$ was found to be photostable as compared with **2**[OTf] and ***3*** (fig. S2). Compound **2**[OTf] converted to ***3*** with excess TFA, and similarly, the addition of excess bpy to ***3*** resulted in conversion to **2**[OTf] (fig. S3). Thus, in the case of ***3***, we believe that there is an initial equilibrium established between **2**[OTf] and ***3***, thus accounting for the initial appearance of the 700-nm absorption band in Fig. 3B. The loss of absorption bands in the visible spectral region and the disappearance of the initial axial EPR signal of **2**[OTf] and ***3*** with irradiation (fig. S4) were consistent with the photoreduction of the Ag(II) center to form the d^{10} Ag(I). Additionally, the ^{1}H NMR signals of Ag(I) bpy complexes appeared over the time course of photolysis. For ***3***, clean conversion to the trigonal planar Ag(I)(bpy)(TFA) complex (*39*) was observed after photolysis by means of ^{1}H NMR and UV-visible (vis) spectroscopy (fig. S5). Concomitant to the formation of Ag(I), photoinduced cleavage of the Ag(II)–TFA bond should furnish the TFA radical (TFA•), decarboxylation of which to yield CO_2 and CF_3• is highly exergonic (*40*).

The decarboxylation of TFA• to CO_2 was observed by steady-state and time-resolved vibrational spectroscopy. Figure 4A shows the steady-state infrared (IR) spectrum of photolyzed solutions of ***3***. Signals associated with the TFA ligand in ***3*** (asymmetric and symmetric carboxyl stretches at 1693 and 1408 cm^{-1}, respectively, and a CF_3 signal at 1196 cm^{-1}) disappeared upon irradiation of solutions of ***3*** with the concomitant growth of the vibrational signature of CO_2 (2341 cm^{-1}). A similar decarboxylation process was observed for photolyzed solutions of **2**[OTf] (fig. S6). Figure 4C shows the time-resolved infrared (TRIR) spectrum upon laser excitation of ***3***. On a nanosecond timescale, bleach signals were observed owing to the disappearance of the asymmetric carboxyl and CF_3 stretching vibrations of TFA with the concomitant growth of CO_2. The time evolution of the vibrational signatures in Fig. 4C approached the instrument response function, placing a lower limit of $k > 2.5 \times 10^7\ s^{-1}$ for TFA• decarboxylation.

The decarboxylation of TFA• to CF_3• was further established with ^{19}F NMR spectroscopy. Photolysis of **2**[OTf] generated CF_3• radicals, which reacted with bpy ligands in the absence of a substrate to produce CF_3-derivatized bpy, and with solvent to produce fluoroform (in MeCN), fluoroform-D (in CD_3CN), and chlorotrifluoromethane (in CH_2Cl_2) (fig. S7). We note that heating (60°C) either **2**[OTf] or ***3*** in CD_3CN solutions protected from ambient light did not produce Ag(I) or CF_3•, as ascertained by ^{1}H and ^{19}F NMR spectroscopy, respectively (fig. S8).

The importance of Ag(II) in mediating the challenging decarboxylation of TFA through LMCT excitation was revealed by a comparative study that used the less electrophilic Cu(II) center in the same ligand field afforded by [Cu$(bpy)_2$(TFA)][PF_6] (**4**[PF_6]). The solid-state structure of **4**[PF_6] (Fig. 2D) was similar to that of **2**[OTf], with the TFA ligand engaging in a shorter contact to the Cu center [*d*(Cu–O1) = 2.0703(19) Å]. Irradiation of **4**[PF_6] with λ_{exc} = 370 nm in CD_3CN resulted in no observable changes as monitored by UV-vis–near infrared (NIR), ^{19}F NMR, and Fourier transform infrared (FTIR) spectroscopies (fig. S9).

Reactivity of photogenerated perfluoroalkyl radicals

The CF_3• radical readily reacted with arenes to form C(sp^2)–CF_3 linkages. Photolysis of **2**[OTf] in the presence of excess benzene (75 equiv) resulted in the formation of trifluorotoluene ($PhCF_3$) (36%) and a mixture of TFA and trifluoroacetic acid (TFAH) (52%), as was confirmed by spiking samples with either [PPN][TFA] or TFAH (fig. S10). The generation of approximately half an equivalent of $PhCF_3$ for each equivalent of **2**[OTf] suggested that after radical addition of CF_3• to benzene, the resultant arene radical was oxidized by a second equivalent of Ag(II), with TFA acting as a base to furnish $PhCF_3$ and TFAH (Fig. 5). As shown by the action spectrum in Fig. 3C, the photoreaction quantum yield of **2**[OTf], as ascertained by ferrioxalate actinometry, was appreciable; the increase and leveling of the quantum yield with wavelength (4.2, 6.8 and 14.9% at 450, 405 and 340 nm, respectively) established the energy of the LMCT transition to occur at ~350 nm. Similar photochemistry and quantum efficiency trends were observed for ***3***. We note that the isolation of **2**[OTf] (or ***3***) is not necessary to perform perfluoroalkylations. The **1**$[OTf]_2$ complex competently binds $R_FCO_2^-$ in situ to provide entry into the LMCT photochemistry. Irradiating (λ_{exc} = 450 nm) a mixture containing Na(TFA) as the limiting reagent (12.5 mM) and 2 equiv of **1**$[OTf]_2$ in combination with one equivalent of benzene for 1 hour produced $PhCF_3$ in 27% yield, as detected by ^{19}F NMR spectroscopy and gas chromatography. The remainder of the mass balance can be ascribed to H-atom abstraction of the solvent by CF_3• to produce fluoroform as a byproduct, as well as the trifluoromethylation of bpy ligands instead of the benzene substrate (fig. S7). The side reactions of CF_3• may be minimized by performing the photolysis with excess substrate; the $PhCF_3$ yield increased to 78% upon the use of 20 equiv of benzene (fig. S11). No $PhCF_3$ was observed upon irradiation of **1**$[OTf]_2$ with benzene in the absence of a TFA source (fig. S12).

As shown in Fig. 6, a variety of arenes as well as heteroarenes such as 1-methylpyrazole are compatible with the described reaction conditions, and trifluoromethylated aromatic compounds may be produced in moderate to high

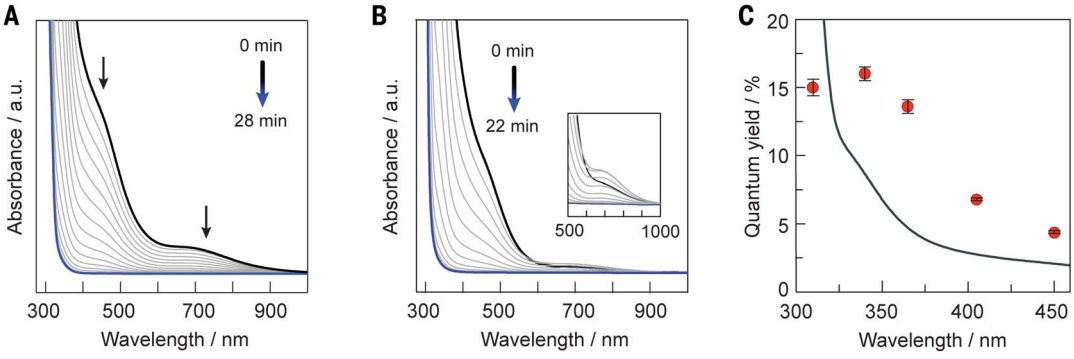


Fig. 3. LMCT photochemistry of Ag(II) complexes. Steady-state photolysis of 1-mM solutions of (**A**) **2**[OTf] and (**B**) **3** in MeCN under 450-nm irradiation at 21.3°C. Spectra were recorded at 2-min intervals. a.u., arbitrary units. (**C**) Action spectrum of **2**[OTf]; quantum yields (red circles) are superimposed on the UV spectrum of **2**[OTf] (black line), with error bars determined from the least-squares fit of the photolytic conversion at different time points.

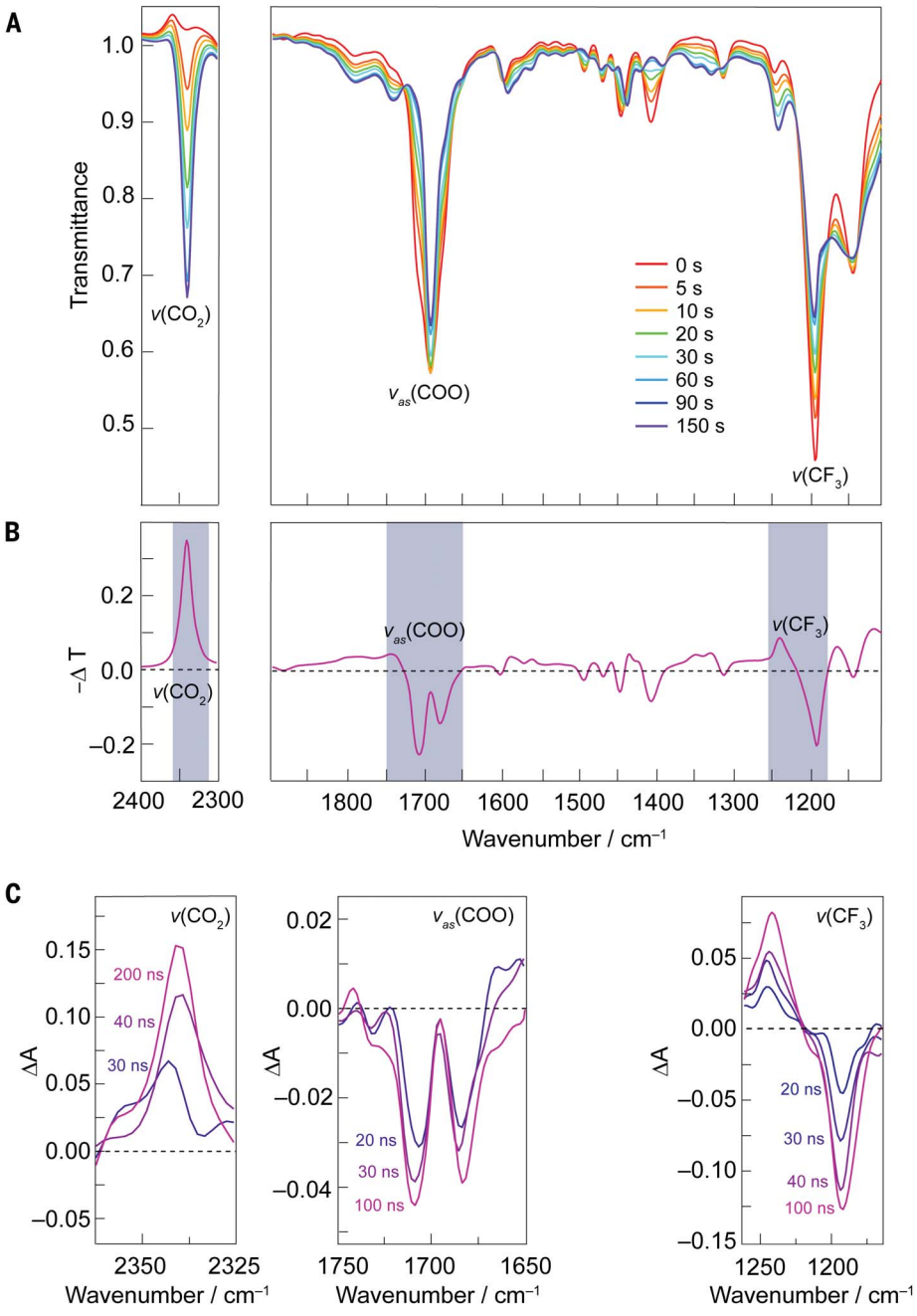


Fig. 4. Steady-state and TRIR spectra of photolyzed solutions of 3. (**A**) Steady-state FTIR spectral changes of a photolyzed solution (λ_{exc} = 467) of **3** in CD_3CN. (**B**) Difference spectrum of FTIR traces in (A) recorded at 150 and 0 s of irradiation. (**C**) TRIR spectra measured in the energy ranges indicated by the shaded boxes in (B). ν_{as}(COO), asymmetric carboxyl stretch.

yields with LMCT irradiation of **1**$[OTf]_2$ in the presence of TFA. The use of an inorganic base (K_2HPO_4) did not substantially alter reaction yields. The observed product ratios for chlorobenzene, methyl benzoate, 4-chloroanisole, methyl 4-chlorobenzoate, 1-methylpyrazole, and benzamide reflect the radical nature of the reaction. Notably, presaturating the reaction mixture with O_2 prior to irradiation led to trifluoromethoxylation, as evidenced by the production of (trifluoromethoxy)benzene when O_2 is present. The trifluoromethoxy group is an important motif in pharmaceutical chemistry, yet trifluoromethoxylation reactions remain underdeveloped, and relatively few radical-based trifluoromethoxylations have been reported (*41*, *42*). The mechanism of this novel O_2/Ag(II)-mediated radical trifluoromethoxylation is currently under study.

To further demonstrate the generality of the current approach, perfluoro groups beyond the parent CF_3 group may be installed. Such higher-order perfluoroalkylations are important to pharmaceutical chemistry, though with far fewer demonstrations as compared with trifluoromethylations (*43*). Single crystals of Ag(bpy)$(O_2CCF_2CF_3)_2$ (**5**) and Ag(bpy)$(O_2CCF_2CF_2CF_3)_2$ (**6**) were obtained, and these compounds were shown to be structural analogs to **3** (Fig. 2, E and F). Photoinduced perfluoroalkylations of benzene could be performed by irradiating solutions of **1**$[OTf]_2$ with the sodium salts of pentafluoropropionate and heptafluorobutyrate (Fig. 6).

Electrophotocatalysis

The photoredox reaction may be rendered electrophotocatalytic by regenerating **1**[OTf] electrochemically during irradiation (Fig. 5). Cyclic voltammetry revealed a reversible one-electron redox wave for **1**$[OTf]_2$ at E^o = 930 mV versus ferrocenium/ferrocene (Fc^+/Fc) (fig. S13). The trifluoromethylation of benzene was thus achieved with 21.2 turnovers upon performing the photoredox reaction (λ_{exc} = 440 nm) with 2.5 mol % **1**$[OTf]_2$, 0.5 equiv of bpy, 1 equiv of benzene, and 10 equiv of Na(TFA) in a three-electrode divided cell under an applied bias of 1.3 V versus Fc^+/Fc. Bis(trifluoromethylated) benzene products were also observed in 11% yield. Of note is the absence of phenyl trifluoroacetate (PhO_2CCF_3) side product in the reaction mixture, which can form from the direct oxidation of benzene (*35*). This highlights the mildness of this electrophotocatalytic method, as oxidation of arenes is avoided in contrast to direct electrolysis methods. Additionally, the lack of this side-product is consonant with the TRIR measurements, which establish that the short-lived TFA radical will decarboxylate before diffusing to arene substrates. Extension of the electrophotocatalytic method to the pentafluoroethylation and heptafluoropropylation of benzene resulted in similar reaction efficiency (27.2 and 20.8 turnovers, respectively). Application of the electrophotocatalysis to other arenes resulted in their successful, catalytic trifluoromethylation with moderate turnover numbers. Any potential anodic of the Ag(II/I) couple but cathodic of the substrate oxidation potential may be applied. Figure S14 shows the permissible potential window for electrophotocatalysis of substrates with oxidation potentials more anodic than that of TFA. In the case of chloroanisole, the reaction was run at an applied potential of 1.1 V versus Fc^+/Fc to avoid direct oxidation of the electron-rich substrate.

We have detailed a new electrophotoredox method to deliver perfluoroalkyl radicals to arenes using the LMCT photochemistry of Ag(II) coupled to Ag(I) oxidation. The homolysis of the Ag(II)–O_2CR_F bond smoothly leads to Ag(I) and the perfluoroalkyl radical, which may be harnessed for C(sp^2)–R_F bond formations. The high electrophilicity of the Ag(II) center is essential for generating perfluoroalkyl radicals from

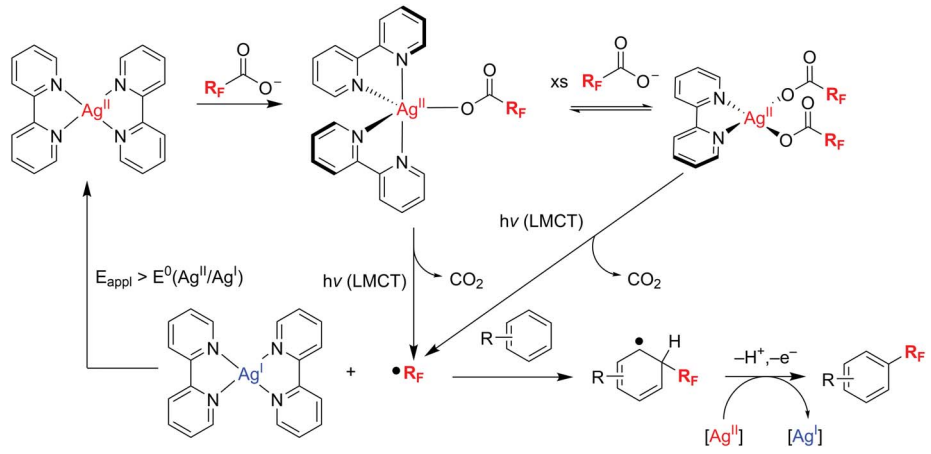


Fig. 5. Electrophotocatalytic reaction mechanism. Mechanism of Ag(II)-mediated perfluoroalkylation of arenes via LMCT excitation under an applied potential to engender a catalytic cycle.

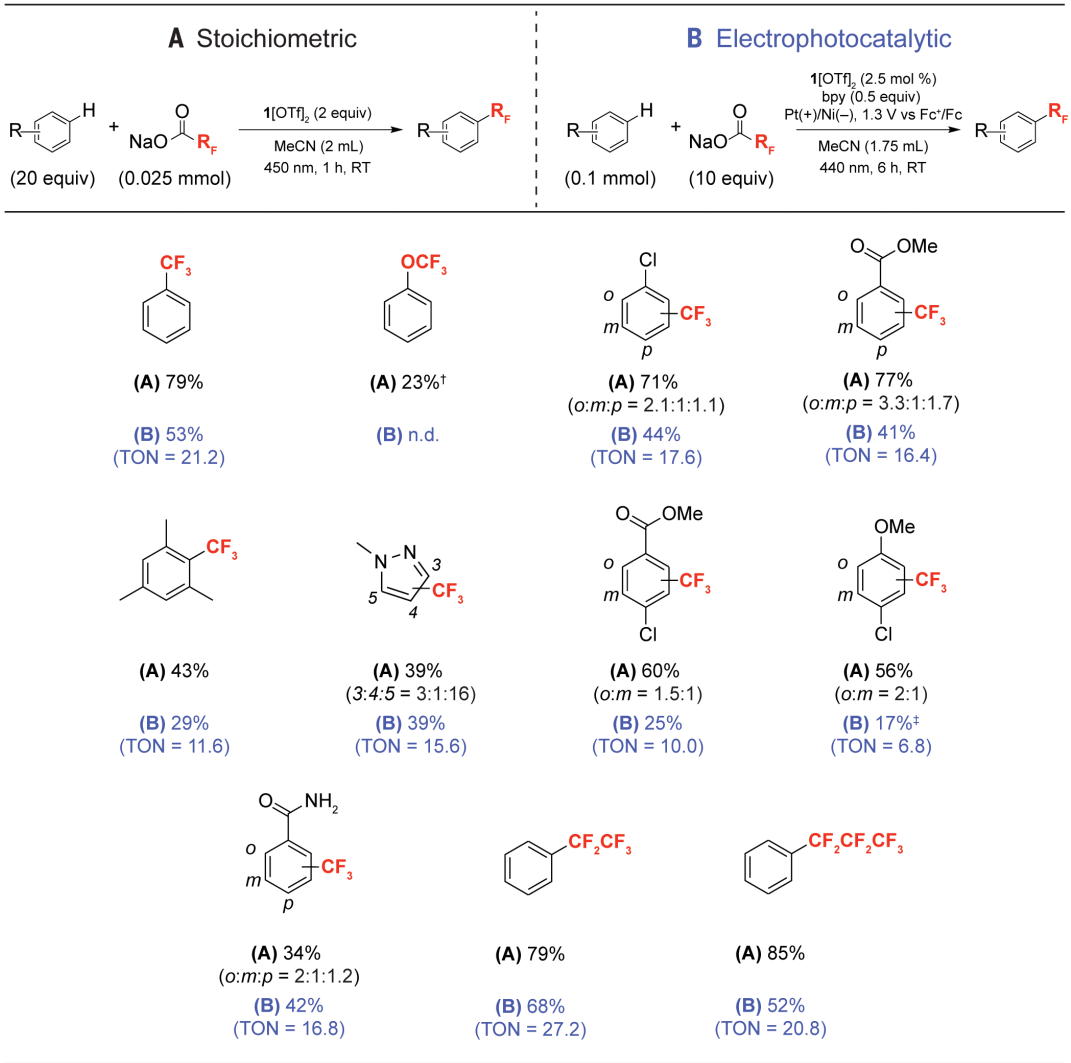


Fig. 6. Ag(II)-mediated trifluoromethylation of arenes.* *Yields for trifluoromethylated products determined by ^{19}F NMR spectroscopy. †Solution was saturated with O_2 prior to irradiation; reaction time, 6 hours. ‡Electrophotocatalysis performed at 1.1 V versus Fc^+/Fc. TON, turnover number.

their corresponding carboxylates under mild photochemical conditions (*44*, *45*). The electrophotoredox method described here is direct and requires only visible light to drive the LMCT activation of TFA or higher perfluoroalkyl carboxylates, without the need for the prefunctionalization of TFA or the use of forcing potentials for Kolbe electrolysis. By performing the LMCT photochemistry under electrochemical conditions in which Ag(I) is reoxidized to Ag(II), an electrophotocatalytic cycle may be established where perfluoroalkyl carboxylates are directly used as perfluoroalkylating agents under mild conditions. Considering the inertness of perfluoroalkyl carboxylates toward oxidation, the chemistry described here may be generalized to a variety of inert substrates that are capable of ligating to a Ag(II) center.

REFERENCES AND NOTES

1. M. H. Shaw, J. Twilton, D. W. C. MacMillan, *J. Org. Chem.* **81**, 6898–6926 (2016).
2. F. Juliá, *ChemCatChem* **14**, e202200916 (2022).
3. Y. Abderrazak, A. Bhattacharyya, O. Reiser, *Angew. Chem. Int. Ed.* **60**, 21100–21115 (2021).
4. A. Reichle *et al*., *Chem. Commun. (Cambridge)* **58**, 4456–4459 (2022).
5. A. Hossain *et al*., *Angew. Chem. Int. Ed.* **57**, 8288–8292 (2018).
6. M. E. Weiss, L. M. Kreis, A. Lauber, E. M. Carreira, *Angew. Chem. Int. Ed.* **50**, 11125–11128 (2011).
7. M. Ociepa, O. Baka, J. Narodowiec, D. Gryko, *Adv. Synth. Catal.* **359**, 3560–3565 (2017).
8. M. I. Gonzalez *et al*., *J. Am. Chem. Soc.* **144**, 1464–1472 (2022).
9. K.-J. Bian, S.-C. Kao, D. Nemoto Jr., X.-W. Chen, J. G. West, *Nat. Commun.* **13**, 7881 (2022).
10. S. K. Kariofillis, A. G. Doyle, *Acc. Chem. Res.* **54**, 988–1000 (2021).
11. S. J. Hwang *et al*., *J. Am. Chem. Soc.* **137**, 6472–6475 (2015).
12. S. J. Hwang *et al*., *Organometallics* **34**, 4766–4774 (2015).
13. S. M. Aliwi, C. H. Bamford, *J. Chem. Soc., Faraday Trans. I* **71**, 1733–1743 (1975).
14. S. Shirase *et al*., *J. Am. Chem. Soc.* **142**, 5668–5675 (2020).
15. A. Hu, J.-J. Guo, H. Pan, Z. Zuo, *Science* **361**, 668–672 (2018).
16. Q. Yang *et al*., *Science* **372**, 847–852 (2021).
17. G. A. Lutovsky, S. N. Gockel, M. W. Bundesmann, S. W. Bagley, T. P. Yoon, *Chem* **9**, 1610–1621 (2023).
18. S. B. Beil, T. Q. Chen, N. E. Intermaggio, D. W. C. MacMillan, *Acc. Chem. Res.* **55**, 3481–3494 (2022).
19. G. V. Zhutaeva, N. A. Shumilova, Copper, Silver, and Gold, in Standard Potentials in Aqueous Solutions, A. J. Bard, R. Parsons, J. Jordan, Eds. (Marcel Dekker, 1985), pp. 294–313.
20. W. Grochala, Z. Mazej, *Philos. Trans. A Math. Phys. Eng. Sci.* **373**, 20140179 (2015).
21. O. A. Tomashenko, V. V. Grushin, *Chem. Rev.* **111**, 4475–4521 (2011).
22. K. Müller, C. Faeh, F. Diederich, *Science* **317**, 1881–1886 (2007).
23. K. A. McReynolds *et al*., *J. Fluor. Chem.* **131**, 1108–1112 (2010).
24. S. E. López, J. Salazar, *J. Fluor. Chem.* **156**, 73–100 (2013).
25. C. Depecker, H. Marzouk, S. Trevin, J. Devynck, *New J. Chem.* **23**, 739–742 (1999).
26. J. Lin *et al*., *Nat. Commun.* **8**, 14353 (2017).
27. C. Lai, T. E. Mallouk, *J. Chem. Soc. Chem. Commun.* **17**, 1359–1361 (1993).
28. G. Shi, C. Shao, S. Pan, J. Yu, Y. Zhang, *Org. Lett.* **17**, 38–41 (2015).
29. D. Staveness, I. Bosque, C. R. J. Stephenson, *Acc. Chem. Res.* **49**, 2295–2306 (2016).
30. Y. Tanabe, N. Matsuo, N. Ohno, *J. Org. Chem.* **53**, 4582–4585 (1988).
31. J. W. Beatty, J. J. Douglas, K. P. Cole, C. R. J. Stephenson, *Nat. Commun.* **6**, 7919 (2015).
32. D. Yin, D. Su, J. Jin, *Cell Rep. Phys. Sci.* **1**, 100141 (2020).
33. B. Yang, D. Yu, X.-H. Xu, F.-L. Qing, *ACS Catal.* **8**, 2839–2843 (2018).
34. L. J. Allen, P. J. Cabrera, M. Lee, M. S. Sanford, *J. Am. Chem. Soc.* **136**, 5607–5610 (2014).
35. J. Qi *et al*., *J. Am. Chem. Soc.* **145**, 24965–24971 (2023).
36. Y. Qiu, A. Scheremetjew, L. H. Finger, L. Ackermann, *Chemistry* **26**, 3241–3246 (2020).
37. W. G. Thorpe, J. K. Kochi, *J. Inorg. Nucl. Chem.* **33**, 3958–3962 (1971).
38. S. Kandaiah, E. M. Peters, M. Jansen, *Z. Anorg. Allg. Chem.* **634**, 2483–2486 (2008).
39. E. Szłyk, R. Szczęsny, A. Wojtczak, *Dalton Trans.* **39**, 1823–1830 (2010).
40. G. Dos Passos Gomes, A. Wimmer, J. M. Smith, B. König, I. V. Alabugin, *J. Org. Chem.* **84**, 6232–6243 (2019).
41. Y. Ouyang, X.-H. Xu, F.-L. Qing, *Angew. Chem. Int. Ed.* **61**, e202114048 (2022).
42. S. Dix, P. Golz, J. R. Schmid, S. Riedel, M. N. Hopkinson, *Chemistry* **27**, 11554–11558 (2021).
43. L. I. Panferova *et al*., *Angew. Chem. Int. Ed.* **54**, 5218–5222 (2015).
44. During the revision process for this manuscript, alkene hydrotrifluoromethylation by an iron catalyst irradiated with 390-nm light to decarboxylate trifluoroacetate was reported (*45*).
45. K.-J. Bian *et al*., *Nat. Chem.* **15**, 1683–1692 (2023).

ACKNOWLEDGMENTS

Funding: National Science Foundation CHE-2243724 (D.G.N.); National Science Foundation Graduate Research Fellowship Program under Grant DGE-2140743 and the Herchel Smith Graduate Fellowship Program at Harvard University (B.M.C.); Arnold and Mabel Beckman Foundation for an Arnold O. Beckman Postdoctoral Fellowship (M.I.G.); National Institute of Health F32GM147975 (J.B.G.); Advanced Photon Source, an Office of Science User Facility, operated for the U.S. Department of Energy (Office of Science) by Argonne National Laboratory under DOE contract no. DE-AC02-06CH11357; Major Research Instrumentation Program of the National Science Foundation award no. 2216066 (Harvard Department of Chemistry and Chemical Biology x-ray facility). **Author contributions:** Conceptualization: B.M.C., M.N., and D.G.N.; Methodology: B.M.C., J.B.G., E.R.R., M.I.G., and K.G.R.; Investigation: B.M.C., J.B.G., E.R.R., M.I.G., and K.G.R.; Funding acquisition: D.G.N.; Writing – original draft: B.M.C. and D.G.N.; Writing – review and editing: B.M.C., J.B.G., E.R.R., M.I.G., K.G.R., M.N., and D.G.N. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Metrical data and supplementary crystallographic data for the solid-state structures are available from the Cambridge Crystallographic Data Centre under reference numbers CCDC 2289548 to 2289554. These data can be obtained free of charge at www.ccdc.cam.ac.uk/data_request/cif, by emailing data_request@ccdc.cam.ac.uk, or by contacting The Cambridge Crystallographic Data Centre at 12 Union Road, Cambridge CB2 1EZ, UK; fax: +44 1223 336033. All other data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adk4919
Materials and Methods
Supplementary Text
Figs. S1 to S49
Tables S1 to S3
References (*46*–*59*)

Submitted 24 August 2023; resubmitted 30 September 2023
Accepted 4 December 2023
Published online 14 December 2023
10.1126/science.adk4919

WATER CHEMISTRY

The mutual neutralization of hydronium and hydroxide

Alon Bogot[1], Mathias Poline[2], MingChao Ji[2], Arnaud Dochain[2], Ansgar Simonsson[2], Stefan Rosén[2], Henning Zettergren[2], Henning T. Schmidt[2], Richard D. Thomas[2], Daniel Strasser[1]*

Mutual neutralization of hydronium (H_3O^+) and hydroxide (OH^-) ions is a very fundamental chemical reaction. Yet, there is only limited experimental evidence about the underlying reaction mechanisms. Here, we report three-dimensional imaging of coincident neutral products of mutual-neutralization reactions at low collision energies of cold and isolated ions in the cryogenic double electrostatic ion-beam storage ring (DESIREE). We identified predominant H_2O + OH + H and 2OH + H_2 product channels and attributed them to an electron-transfer mechanism, whereas a minor contribution of H_2O + H_2O with high internal excitation was attributed to proton transfer. The reported mechanism-resolved internal product excitation, as well as collision-energy and initial ion-temperature dependence, provide a benchmark for modeling charge-transfer mechanisms.

Ionization and neutralization mechanisms are fundamental to the understanding of various chemical environments, ranging from low-temperature chemistry in the interstellar medium (*1*–*3*) to atmospheric chemistry (*4*–*6*), acid-base phenomena (*7*–*11*), and high-temperature artificial plasma (*12*, *13*). In particular, the formation and destruction of "water ions" attracts a broad interest in isolated or gas-phase environments (*14*–*22*), as well as in aqueous solution (*7*–*11*, *23*–*26*). One of the most fundamental ionization-neutralization equilibria is $H_3O^+ + OH^- \rightleftharpoons 2H_2O$, the mutual neutralization by proton transfer (PT) between hydronium and hydroxide ions and the reverse reaction of water autoionization. Nevertheless, there is only limited direct experimental evidence about the mechanisms of these reactions that determine the pH of pure water (*8*–*10*). In contrast to the seeming simplicity of this textbook reaction, mutual neutralization (MN) of molecular ions can exhibit a rich variety of products. Owing to the large difference between typical ionization potentials and electron affinities of neutral species, the excess energy in MN reactions allows multiple competing product channels. In the case of isolated hydronium and hydroxide ions, up to ~9.77 eV of excess energy can be available for excitation of the pair of neutral water molecules produced by a PT mechanism (*7*, *24*, *27*). This excess energy can be released as kinetic energy or facilitate internal excitation as well as fragmentation of the molecular products. Furthermore, the ionization potential of a neutral hydronium radical was estimated to be 4.3 eV (*28*), about 2.5 eV higher than the 1.8-eV electron affinity of the OH radical (*29*). Thus, the alternative electron transfer (ET)–based MN mechanism should be considered in the case of isolated ions. Assuming a simple Coulombic attraction between the isolated ionic reactants and weakly interacting neutral radicals, we could estimate a crossing of the ionic and neutral potentials at a large ~6-Å distance. This is consistent with the tentative assignment of predominant ET, as this distance is appreciably larger than expected for PT reactions to occur (*20*). Although such simple estimates have been shown to capture the essence of ET in cation-anion reactions with atomic species (*30*–*36*), they should be regarded with caution when applied to molecular systems, as recently demonstrated in a combined experimental and theoretical study of Coulomb explosions ignited by ET from a neutral to a dication reactant (*37*). Thus, bond-forming mechanisms such as PT that are expected to occur at closer range can often compete with ET and lead to nontrivial product distributions (*37*, *38*).

Early theoretical simulations considered MN of isolated hydronium and hydroxide by a PT mechanism and predicted an asymmetric internal-energy distribution, in which one of the water products is internally hot (*21*). With the increase of computational resources, the PT mechanism was extensively explored by ab initio simulations of liquid water (*9*, *10*, *23*, *39*, *40*). Although the subnanometer dimensions and ultrafast timescale of MN reactions have so far prevented a direct experimental study in the liquid phase, pioneering merged-beams experiments measured the rates of MN reactions for several $(H_2O)_nH_3O^+ + (H_2O)_mOH^-$ cluster systems (*20*). In this way, they systematically explored the transition from an isolated ion system to solvated ions in the liquid. Coe and co-workers attributed the observed 10-fold increase in the MN rate constant due to microsolvation of the ions by three water molecules to a transition from a purely ET mechanism in the case of isolated ions to the very different, PT-based mechanism that is expected to dominate MN in liquid water (*20*). Nevertheless, these early experiments were performed using ion sources with little control over the rovibrational or electronic excitations of the different ion species. Furthermore, no information could be provided about the final MN products, making it difficult to come to firm conclusions on the underlying mechanism or to benchmark high-level theoretical simulations. Recent merged atomic-ion beam experiments, coupled with three-dimensional (3D) imaging of the coincident neutral MN products, make it possible to study the MN mechanism in unprecedented detail in low-energy (sub–electron volt) collisions (*30*, *31*). In the case of molecular ions, in addition to the collision energy, such detailed studies also require control of the ion's internal excitation. This effect can be achieved in the recently developed cryogenic ion storage devices (*32*, *33*, *41*–*48*). In particular, the double electrostatic storage ring experiment DESIREE allows internally hot ions to relax to temperatures close to the ambient 13 K environment while circulating in the ring (*43*, *49*). A common section of the two rings allows merging beams of cations and anions at low center-of-mass (CM) energy to provide detailed information about the MN mechanism by 3D imaging of coincident MN products (*32*–*34*).

The stability of the isolated hydronium radical (H_3O), the neutral product of ET to H_3O^+, has been the subject of much theoretical and experimental debate. Although early studies suggested that the radical is metastable (*26*, *28*, *50*–*54*), recent experiments point to its rapid decay (*16*, *55*). H_3O^+ charge-exchange studies point to a subnanosecond dissociation of the unstable intermediate [H_3O] complex to H_2O + H (*14*–*16*), in agreement with the predicted very low (~0.1 eV) barrier toward this product channel (*22*, *52*). Furthermore, dissociative recombination of H_3O^+ with slow electrons exhibited two additional product channels: a dominant OH + H + H and a minor OH + H_2 (*17*–*19*), where the latter channel overcomes a ~0.8-eV barrier on the [H_3O] ground state (*22*, *52*). On the basis of this experimental evidence, we can consider three distinct product channels for the MN of isolated hydronium and hydroxide ions. Figure 1 shows the energies of (i) the $2H_2O$ channel, produced through effectively a PT mechanism with 9.77 eV of excess energy; (ii) the 2OH + H_2 channel with 4.06 eV of excess energy, produced only through an ET mechanism; and (iii) the H_2O + OH + H channel with an excess energy of 4.67 eV that could be produced either by the breakup of [H_3O] in an ET mechanism or the breakup of a hot H_2O product of PT dynamics. In contrast to dissociative recombination, the electron affinity of the OH radical lowers the maximal neutral product potential that could be reached in MN, such that the 2OH + 2H channel is energetically closed. Nevertheless, one could in principle consider also exotic mechanisms such as concerted ET and structural rearrangement

[1]Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel. [2]Department of Physics, Stockholm University; SE-10691 Stockholm, Sweden.
*Corresponding author. Email: strasser@huji.ac.il

dynamics (*56*), potentially yielding energetically higher channels that are indicated in Fig. 1. These high-lying product channels with correspondingly lower excess energies include the formation of electronically excited OH* and production of atomic O(^{1}D), as well as four-body dissociation of the reactive complex.

Here, by using 3D imaging of the coincident neutral MN products, we disentangled the PT and ET contributions to the energetically possible product channels shown in Fig. 1. In addition, imaging the kinetic energy release (KER) provided information about the internal excitation of the neutral products, and the long storage times in the DESIREE storage rings allowed probing of the effect of internal parent-ion temperatures and collision energy on the MN reaction.

Results and discussion

The neutral signal yield was dominated by single neutral products that were attributed primarily to collisions of the stored ions with the residual gas. Despite the low ~10^{-14}-mbar residual-gas pressure, the yield of neutral MN products was more than two orders of magnitude lower. Nevertheless, it could be efficiently disentangled from the residual-gas background by observing coincident neutral fragments. The MN events were found to predominantly result in three-body breakup. As we show in detail in the supplementary materials (SM), a combined analysis of the total CM momentum conservation and KER in the recorded three-body breakup events allowed us to identify the predominant H_2O + OH + H channel. Figure 2A shows the background-subtracted KER spectrum, obtained at the lowest possible collision energies by analyzing the measured recoil of triple-coincidence neutral fragments and assigning the product H_2O + OH + H masses according to total CM momentum conservation. The distribution was well represented by two Gaussian peaks, where the main peak at ~4.5 e V coincided with the dashed vertical line that indicates the excess energy available for this channel. This fact suggested that the majority of the excess energy was released as KER and was not deposited in the internal excitation of the molecular products. The minor peak was situated ~1.3 eV below the maximal available energy and could be tentatively assigned to some internal product excitation. However, as the lowest electronically excited product state of OH* requires ~4 eV (*57*) and a bimodal rovibrational excitation distribution was unlikely, we considered the possibility of background due to the minor 2OH + H_2 channel. As we show in detail in the SM, CM conservation allowed us to distinguish the light hydrogen from the heavy fragments and to exclude and subtract any random coincidence background contributions. Nevertheless, it was not sufficient to clearly distinguish between the two possible three-body breakup channels, both of which included one light and two heavy fragments that could mainly be distinguished by their KER. In contrast to Fig. 2A, the KER derived by analysis of all triple coincidences as the 2OH + H_2 channel markedly exceeded the maximal possible KER for that channel (SM). We therefore restricted the events considered as 2OH + H_2 to the low-recoil events that would contribute to the minor peak in Fig. 2A. The resulting background-subtracted KER spectrum is shown in Fig. 2B. The KER peak coincided with the ~4.05 eV of excess energy that is available for the H_2 + 2OH channel, indicated by the dashed vertical line. Similar analysis showed that energetically possible contributions from the higher-lying channels could be neglected, as they would result in nonphysical KER above the available excess energy (see SM for more details). We therefore assigned the minor peak in Fig. 2A to 2OH + H_2 breakup in which the available energy was released primarily as KER. We could also conclude that 75 ± 7% of the three-body breakups was ascribed to the predominant H_2O + OH + H channel. Furthermore, for both three-body product channels, the KER coincided with the maximal available excess energy in the MN reaction, indicating a relatively low internal excitation of the molecular products.

In addition to the triple coincidences, coincident pairs of neutral products were also detected. Analysis of these double-coincidence

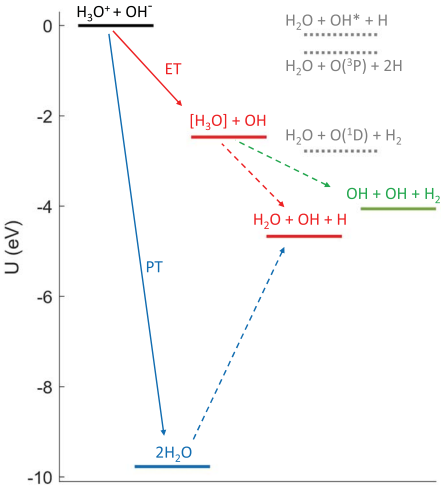
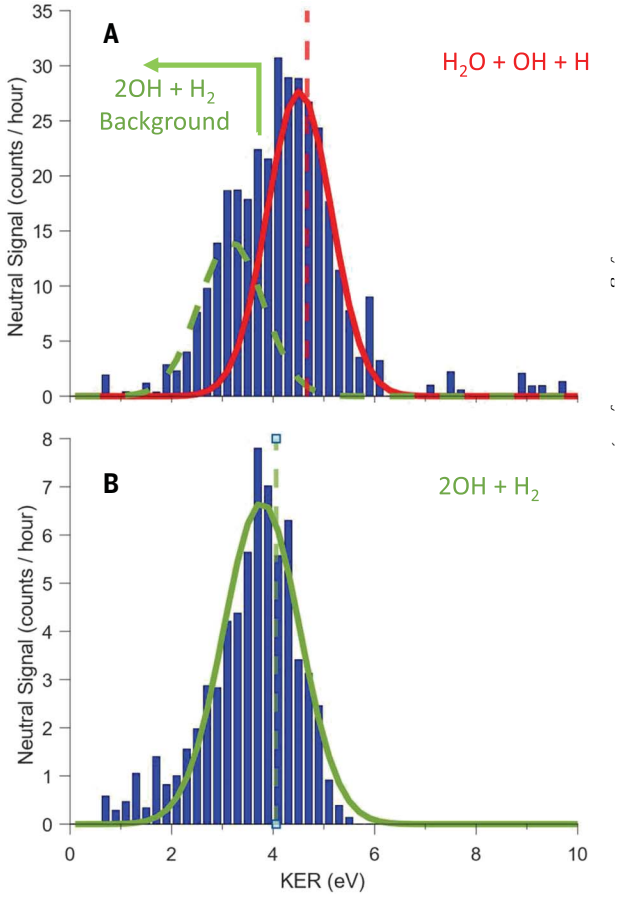


Fig. 1. Energetics of the H_3O^+ + OH^- mutual neutralization reaction. Arrows indicate the possible electron-transfer (ET) and proton-transfer (PT) initiated pathways toward the energetically accessible product channels. In addition to ground-state products, an energetically accessible product channel with electronically excited OH* ($^2\Sigma^+$) is shown, as well as atomic O(^{1}D) formation and four-body breakup high-lying channels.

Fig. 2. Coincident three-body product channel analysis. (**A**) Background-subtracted KER spectrum, where the KER was derived from the measured 3D product recoil distances assuming an H_2O + OH + H breakup and requiring the appropriate CM momentum conservation. The peak fitted by the red Gaussian curve was assigned to the H_2O + OH + H channel, and the lower recoil peak fitted with the dotted green Gaussian curve was attributed to a background contribution due to the 2OH+H_2 channel. (**B**) Analysis of events with low-recoil events (contributing to KER <3.5 eV in the H_2O + OH + H analysis), assuming a 2OH + H_2 breakup. In both panels, dashed vertical lines indicate the maximal available energy in the respective product channel.



events must take into consideration all the different scenarios that could result in the detection of two neutral products. In the following, we show that the proton-transfer MN mechanism, resulting in two neutral water molecule products, was responsible for only a small fraction of the detected neutral pairs. The finite detection efficiency of the microchannel-plate (MCP) detector resulted in a considerable background from partially detected three-body events. Furthermore, one has to consider partial detection of the predominant three-body breakup channels due to the detector resolution and its finite active area. As both three-body channels included a pair of similar-mass products, the total CM conservation requirement had a finite capacity to exclude this background contribution from real two-body events. Nevertheless, we could estimate the contribution of partial detection on the basis of the detected triple-coincidence signal and a detailed simulation of the experimental geometry, assuming that the three-body breakup plane was isotropically distributed in the moving frame of reference. Additional details about the procedure can be found in the SM.

Assuming that the neutral pair corresponded to the two-body $H_2O + H_2O$ channel, the square of the distance between the fragments as the fragments hit the detector (R_{3D}) could be directly related to the KER (see materials and methods in the SM). The blue bars in Fig. 3 show the spectrum obtained for coincident pairs with equal and opposite recoil, according to total CM momentum conservation in the $H_2O + H_2O$ channel. The gray region in Fig. 3 shows the background contribution that was estimated on the basis of events in which the total CM momentum of the neutral fragments did not match the total momentum of the parent ions, while scaling the overall random coincidence background yield based on measured random events with nonphysical KER above 10 eV (see SM for more details). The red region in Fig. 3 shows the evaluated contribution of the above-mentioned partially detected three-body breakup events to the two-body spectrum. In the general case, it is difficult to estimate the exact detection efficiency of MCP detectors, as it depends not only on the so-called open-area ratio but also on the velocity of the detected particles as well as on the readout method and operation temperature (*58*, *59*). However, as the spectral shape of the partial detection background fitted the shape of the low R_{3D} part of the measured two-body spectrum, the relative count rate was found to be consistent with a ~30% detection efficiency for each individual neutral fragment, lower than the optimal open-area efficiency that could be reached at room-temperature operation of MCP detectors but consistent with independent estimates of the detector efficiency (*60*, *61*). Similarly, part of the spectrum corresponding to R_{3D} > 32 mm was well reproduced by the background. However, as the maximal KER in the three-body breakup channels was at most ~4.67 eV, the high-recoil events in the region indicated by cyan color in Fig. 3 could not be accounted for by background contributions (see SM for more technical details). These high-recoil events could not be attributed to formation by ET of a metastable H_3O and an OH radical because it allowed only ~2.47 eV of excess energy. It could therefore be assigned only to the $2H_2O$ channel with ~5 eV KER. The dotted vertical line shows the 9.77 eV of excess energy for this process, indicating that as opposed to the three-body breakup channels, the two water molecules carried substantial internal energy. The internal energy could be calculated by subtracting the measured KER from the total excess energy, and the internal energy spectrum of the $2H_2O$ molecules, after subtraction of the background and partial detection events, is shown in the inset of Fig. 3. In contrast with both three-body channels, the PT mechanism yielded high internal excitation of the water products. For comparison, the dashed vertical line in the inset of Fig. 3 indicates the dissociation energy of a single H_2O molecule. The $2H_2O$ internal energy distribution dropped abruptly above the energy required to dissociate one H_2O molecule, suggesting a possible contribution of PT to the $H_2O + OH + H$ channel. By taking into account the simulated detector acceptance of each event and the finite detection efficiency, we concluded that the probability for production of two water molecules by the PT mechanism corresponded to at least 2 ± 0.5% of the total MN yield.

The CM ion-ion collision energy was controlled by an offset potential (U_0) applied to a drift tube located in the merged section of the stored cation and anion beams. Nevertheless, the finite collision angles resulted in an effective instrumental collision energy spread (see materials and methods in the SM for more details). Within the experimental error, the observed rates for the different MN channels exhibited similar collision energy dependencies. Figure 4 shows the total rate of MN events as a function of collision energy, neglecting the finite instrumental spread. Background counts were estimated and subtracted on the basis of total CM conservation. As could be expected for a barrierless process, the MN rate followed a power-law dependence as a function of collision energy (*35*, *62*), illustrated by the $\propto E_{CM}^{-0.5}$ dashed blue line in Fig. 4. For nominal collision energies of a few tens of milli–electron volts, the measured yields fell below the power law. The red curve shows the expected effect of a fitted ~80-meV instrumental collision energy spread, in agreement with earlier DESIREE studies (*32*–*34*, *63*). The strong dependence on collision energy suppressed the relative contribution from outside the drift tube region, in which the collision energy was ~30 eV.

The long storage times in DESIREE allowed the observation of ions that were formed in hot ion sources as they cooled in the ambient 13 K temperature environment (*43*). Although vibrational relaxation was expected to occur on the millisecond time scale, the decay of rotations could take considerably longer times (*19*, *64*, *65*). The hydroxide anion has a permanent dipole moment; nevertheless, the decay of its lowest-lying J = 1, 2, and 3 rotational states was observed to proceed on the respective 193-, 21- and 5-s time scales of storage in a cryogenic environment (*42*, *43*). For H_3O^+, the C_{3V} symmetry of the vibrational ground state could also result in slow decay for high-lying rotational states

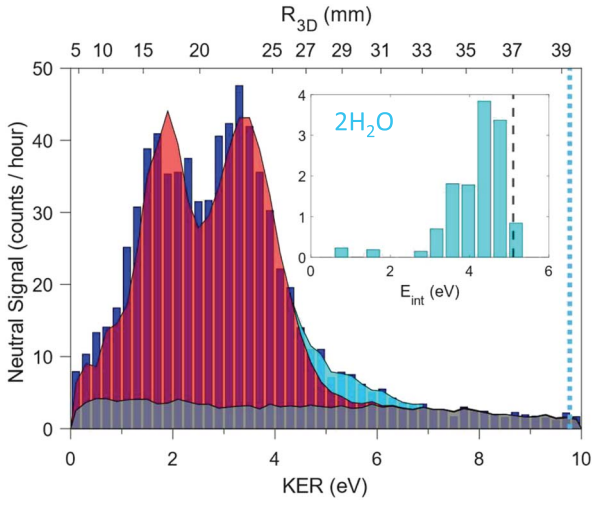


Fig. 3. Coincident two-body product channel analysis. Blue bars show the total KER spectrum, where the KER was derived from the measured 3D distance between the fragments as they hit the detector, assuming an $H_2O + H_2O$ breakup and requiring the appropriate CM momentum conservation. Gray region indicates the background contribution (evaluated from events not conserving CM momentum). The red region indicates the contribution of three-body breakup events to the two-body spectrum (evaluated from the measured three-body events). Cyan region indicates the contribution attributed to $H_2O + H_2O$ events, and the dotted cyan line indicates the maximal 9.77 eV KER for this channel. Inset shows the background-subtracted internal energy distribution in events that were attributed to the $H_2O + H_2O$ channel, and the dashed black line in the inset indicates the dissociation energy of a single water molecule.

(*64*, *65*), and dissociative recombination measurements of stored H_3O^+ ions with slow electrons showed the persistent effect of a high rotational temperature at long storage times (*19*, *64*, *65*). It was therefore valuable to observe the dependence of the MN rate on the ion storage time. The yield of MN events depended on the bimolecular reaction rate constant, which could in turn depend on the internal excitation of the reactants. The yield also scales with the number of stored cations times the number of stored anions. Separate experiments, in which only a cation beam or only an anion beam was injected into the storage device, were performed to estimate the decay of ions in the storage ring. The green squares in the upper panel of Fig. 5 show the rate of single neutral products impinging on the center of the detector while storing only the hydroxide anion beam. The decay was well described by a single exponential decay of the number of stored anions, fitted with a long ~220-s lifetime. Similarly, the red crosses show the neutral yield from a stored hydronium cation beam, fitted with a shorter ~80-s lifetime. The blue circles show the total rate of measured MN events as a function of storage time. In contrast to the single exponential decay of the individually stored ions as a function of storage time, manifested in linear decay on the logarithmic scale of Fig. 5, the MN rate exhibited a more complex dependence indicating a change in the effective rate constant as a function of storage time in the 13K environment. The lower panel of Fig. 5 shows the relative change in the MN reaction rate as a function of storage time, evaluated by dividing the measured MN rate by the product of the separately measured rates of neutrals from the anion and cation beams. This evaluation is based on the assumption that the neutralization rates are constant on the relevant time scales and that they are similar under the single-beam and merged-beams conditions. This assumption was supported by the single exponential decay of the individual ion beams as well as the low, less than one event per second, rate of MN events that could be expected to have a negligible effect on the lifetimes of the individual cation and anion beams. The relative reaction rate constant was fitted with a ~15-s decay to an asymptotic value that was a factor of 3 lower at long storage times. Although thermal equilibrium with the 13 K environment was not expected, the observed decay was attributed to the radiative cooling of highly rotationally excited H_3O^+ ions that were expected to form in the electron cyclotron resonance (ECR) ion source (*19*).

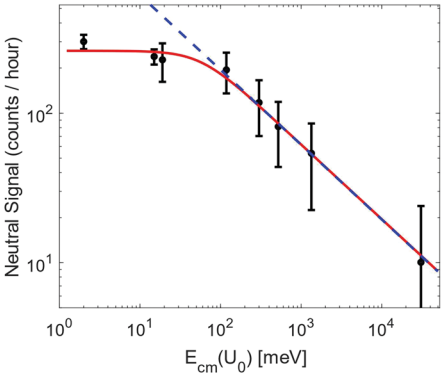


Fig. 4. Total MN signal as a function of collision energy. Black points represent the MN signal as a function of collision energy. Error bars represent the 1σ statistical uncertainty of the measurement. The dashed blue line shows the fitted $\propto E_{cm}^{-0.5}$ trend, neglecting the contribution of the finite spread due to collision angles that becomes significant below ~80 meV. The red solid line shows the same trend, including the instrumental collision energy spread.

Discussion

Analysis of the coincident neutral fragments originating from low-energy $H_3O^+ + OH^-$ collisions showed evidence for both electron-transfer and proton-transfer mechanisms. Although the predominant channel, $H_2O + OH + H$, could in principle originate from both ET and PT mechanisms, the three-body $2OH + H_2$ channel could originate solely from ET, and the two-body $2H_2O$ could only be produced by PT. The kinetic energy spectra indicated that most of the excess energy was released as KER in the case of $2OH + H_2$, with little internal product excitation. The width of the KER peaks that was ~1 eV could be attributed to a combination of the instrumental resolution, initial internal excitation of the ionic reactants, and final excitation of the neutral products. The same fact was observed in the dominant $H_2O + OH + H$ channel, suggesting that it also originated from ET. The low internal excitation in the $H_2O + OH + H$ channel was consistent with the effectively barrierless dissociation of the $[H_3O]$ intermediate (*22*, *52*). Furthermore, the predicted ~0.8-eV barrier toward H_2 elimination on the $[H_3O]$ ground state combined with the low internal excitation may suggest a more direct dissociation involving nonadiabatic dynamics on an electronically excited potential (*22*, *52*). In comparison, water molecules formed by H_3O^+ charge exchange dissociation or dissociative recombination with slow electrons could exhibit higher internal excitation of ~1.5 eV or >3 eV, respectively (*15*, *19*). Nevertheless, it is reasonable that different $[H_3O]$ potentials were populated by the different mechanisms, resulting in different energy partitioning and different product channels. In contrast to the ET mechanism, the two water molecules formed by PT exhibited a high internal excitation. Considering the prediction of Stillinger *et al.* for an asymmetric product excitation (*21*), in which the proton acceptor contains most of the internal excitation, it is possible that some of the PT products dissociated. Consequently, we proposed that the observed ~2% branching fraction toward the $2H_2O$ channel is only a lower limit for the contribution of the PT mechanism that might also contribute to the predominant $H_2O + OH + H$ channel.

We concluded that the mutual neutralization of isolated $H_3O^+ + OH^-$ is a valuable benchmark system for testing theoretical modeling of competing ET and PT reactions. The detailed product information provided here is expected to serve as a test for high-level ab initio simulations of the competing ET and PT dynamics. Follow-up investigations—for example, of the $H_5O_2^+$, Zundel form of the water cation—will provide additional experimental insight about the effect of solvation (*20*). An accurate theoretical modeling framework for the isolated ion pair system will shed light on the

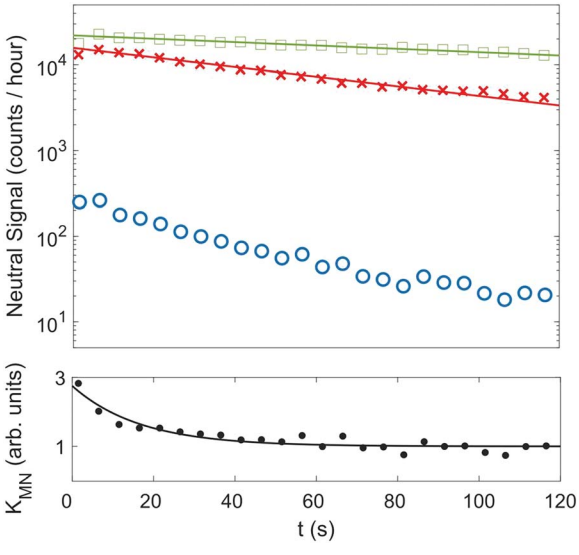


Fig. 5. Rate of neutral yields as a function of ion storage time. (Upper panel) Green squares show the rate of neutrals from the anion beam, and red crosses indicate the neutral yield from the separately measured cation beam. Both are well fitted with single exponential decays of the respective ions as a function of storage time. Blue circles indicate the measured rate of MN events as a function of storage time, showing a nontrivial decay, attributed to the radiative cooling of rotationally hot H_3O^+ ions. (Lower panel) Time evolution of the MN rate constant, evaluated by dividing the MN event yield by the product of neutral yields from the individual cation and anion beams.

coexistence and competition of ET and PT mechanisms in different MN reactions, as well as other charge-transfer reactions of isolated or solvated systems (*37*, *66*, *67*). Furthermore, the observed dependence of the MN rate on the internal temperature of the ionic reactants must be considered when applying rate constants measured using hot ions for the modeling of MN reactions in the ionosphere (*4*–*6*), planetary, and interstellar medium environments (*1*–*3*).

REFERENCES AND NOTES

1. T. J. Millar, C. Walsh, T. A. Field, *Chem. Rev.* **117**, 1765–1795 (2017).
2. T. P. Snow, V. M. Bierbaum, *Annu. Rev. Anal. Chem.* **1**, 229–259 (2008).
3. M. Alagia *et al*., *Rend. Lincei Sci. Fis. Nat.* **24**, 53–65 (2013).
4. L. A. Capone *et al*., *Nature* **293**, 45–46 (1981).
5. H. L. Macintyre, M. J. Evans, *Atmos. Chem. Phys.* **10**, 7409–7414 (2010).
6. N. S. Shuman, D. E. Hunton, A. A. Viggiano, *Chem. Rev.* **115**, 4542–4570 (2015).
7. H. Sato, F. Hirata, *J. Phys. Chem. A* **102**, 2603–2608 (1998).
8. M. Moqadam *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **115**, E4569–E4576 (2018).
9. P. L. Geissler, C. Dellago, D. Chandler, J. Hutter, M. Parrinello, *Science* **291**, 2121–2124 (2001).
10. A. Hassanali, M. K. Prakash, H. Eshet, M. Parrinello, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20410–20415 (2011).
11. J. V. Coe, *Int. Rev. Phys. Chem.* **20**, 33–58 (2001).
12. S. I. Krasheninnikov *et al*., *Phys. Plasmas* **4**, 1638–1646 (1997).
13. Y. L. Wei *et al*., *Rev. Sci. Instrum.* **85**, 103503 (2014).
14. J. E. Mann, Z. Xie, J. D. Savee, J. M. Bowman, R. E. Continetti, *J. Chem. Phys.* **130**, 041102 (2009).
15. J. E. Mann, Z. Xie, J. D. Savee, J. M. Bowman, R. E. Continetti, *J. Phys. Chem. A* **117**, 7256–7266 (2013).
16. P. Hvelplund, S. B. Nielsen, S. Panja, J. O. P. Pedersen, E. Uggerud, *Int. J. Mass Spectrom.* **281**, 52–54 (2009).
17. L. H. Andersen *et al*., *Phys. Rev. Lett.* **77**, 4891–4894 (1996).
18. M. J. Jensen *et al*., *Astrophys. J.* **543**, 764–774 (2000).
19. H. Buhr *et al*., *Phys. Rev. Lett.* **105**, 103202 (2010).
20. B. Plastridge, M. H. Cohen, K. A. Cowen, D. A. Wood, J. V. Coe, *J. Phys. Chem.* **99**, 118–122 (1995).
21. F. H. Stillinger, T. A. Weber, *J. Chem. Phys.* **76**, 4028–4036 (1982).
22. A. E. Ketvirtis, J. Simons, *Phys. Rev. A* **103**, 6552–6563 (1999).
23. E. Perlt, M. von Domaros, B. Kirchner, R. Ludwig, F. Weinhold, *Sci. Rep.* **7**, 10244 (2017).
24. C. Bai, J. Herzfeld, *J. Phys. Chem. B* **121**, 4213–4219 (2017).
25. M. F. Lin *et al*., *Science* **374**, 92–95 (2021).
26. A. L. Sobolewski, W. Domcke, *J. Phys. Chem. A* **106**, 4158–4167 (2002).
27. P. Maksyutenko, T. R. Rizzo, O. V. Boyarkin, *J. Chem. Phys.* **125**, 181101 (2006).
28. G. I. Gellene, R. F. Porter, *J. Chem. Phys.* **81**, 5570–5576 (1984).
29. J. R. Smith, J. B. Kim, W. C. Lineberger, *Phys. Rev. A* **55**, 2036–2043 (1997).
30. T. Launoy *et al*., *Astrophys. J.* **883**, 85 (2019).
31. N. de Ruette *et al*., *Phys. Rev. Lett.* **121**, 083401 (2018).
32. M. Poline *et al*., *Phys. Rev. A* **105**, 062825 (2022).
33. G. Eklund *et al*., *Phys. Rev. A* **103**, 032814 (2021).
34. G. Eklund *et al*., *Phys. Rev. A* **102**, 012823 (2020).
35. A. Dochain, V. M. Andrianarijaona, X. Urbain, *Phys. Rev. A* **108**, 042809 (2023).
36. A. F. Schmidt-May *et al*., *Phys. Rev. A* **108**, 042810 (2023).
37. K. Gope, E. Livshits, D. M. Bittner, R. Baer, D. Strasser, *Sci. Adv.* **8**, eabq8084 (2022).
38. J. Roithová, C. L. Ricketts, D. Schröder, S. D. Price, *Angew. Chem. Int. Ed.* **46**, 9316–9319 (2007).
39. E. Brodskaya, A. P. Lyubartsev, A. Laaksonen, *J. Phys. Chem. B* **106**, 6479–6487 (2002).
40. A. A. Hassanali, J. Cuny, V. Verdolino, M. Parrinello, *Philos. Trans. A Math. Phys. Eng. Sci.* **372**, 20120482 (2014).
41. A. P. O'Connor *et al*., *Phys. Rev. Lett.* **116**, 113002 (2016).
42. C. Meyer *et al*., *Phys. Rev. Lett.* **119**, 023202 (2017).
43. H. T. Schmidt *et al*., *Phys. Rev. Lett.* **119**, 073001 (2017).
44. S. Jinno *et al*., *Nucl. Instruments Methods Phys. Res. A* **532**, 477–482 (2004).
45. S. Jinno *et al*., *Nucl. Instrum. Methods Phys. Res. A* **572**, 568–579 (2007).
46. S. P. Moller, *Proc. IEEE Part. Accel. Conf.* **1**, 1027–1029 (1998).
47. J. U. Andersen *et al*., *Hyperfine Interact.* **146–147**, 283–291 (2003).
48. U. V. Pedersen, M. Hyde, S. P. Møller, T. Andersen, *Phys. Rev. A* **64**, 012503 (2001).
49. H. T. Schmidt *et al*., *Rev. Sci. Instrum.* **84**, 055115 (2013).
50. K. S. E. Niblaeus, B. O. Roos, P. E. M. Siegbahn, *Chem. Phys.* **25**, 207–213 (1977).
51. P. W. McLoughlin, G. I. Gellene, *J. Phys. Chem.* **96**, 4396–4404 (1992).
52. M. Luo, M. Jungen, *Chem. Phys.* **241**, 297–303 (1999).
53. I. Ogurtsov, N. Gorinchoy, I. Balan, *J. Mol. Struct.* **838**, 107–111 (2007).
54. W. J. Griffiths, F. M. Harris, J. H. Beynon, *Int. J. Mass Spectrom. Ion Process.* **77**, 233–239 (1987).
55. S. Panja, P. Hvelplund, S. B. Nielsen, E. Uggerud, *Phys. Chem. Chem. Phys.* **11**, 6125–6131 (2009).
56. D. R. Weinberg *et al*., *Chem. Rev.* **112**, 4016–4093 (2012).
57. K. P. Huber, G. Herzberg, *Molecular Spectra and Molecular Structure* (Springer, 1979), vol. 4.
58. G. A. Rochau *et al*., *Rev. Sci. Instrum.* **77**, 10E323 (2006).
59. P. Roth, G. W. Fraser, *Nucl. Instrum. Methods Phys. Res. A* **439**, 134–137 (2000).
60. R. D. Thomas *et al*., *Rev. Sci. Instrum.* **82**, 065112 (2011).
61. M. H. Stockett *et al*., *Nat. Commun.* **14**, 395 (2023).
62. R. E. Olson, *J. Chem. Phys.* **56**, 2979–2984 (1972).
63. M. Poline *et al*., *Phys. Rev. A* **106**, 012812 (2022).
64. V. V. Melnikov, S. N. Yurchenko, J. Tennyson, P. Jensen, *Phys. Chem. Chem. Phys.* **18**, 26268–26274 (2016).
65. T. Oka, *J. Mol. Spectrosc.* **379**, 111482 (2021).
66. E. Livshits, I. Luzon, K. Gope, R. Baer, D. Strasser, *Commun. Chem.* **3**, 49 (2020).
67. I. Luzon, E. Livshits, K. Gope, R. Baer, D. Strasser, *J. Phys. Chem. Lett.* **10**, 1361–1367 (2019).
68. A. Bogot *et al*., Data Underlying Figures in Mutual Neutralization of Hydronium and Hydroxide, Zenodo (2024); https://doi.org/10.5281/zenodo.10070062.

ACKNOWLEDGMENTS

Funding: This work was performed at the Swedish National Infrastructure, DESIREE (Swedish Research Council contract nos. 2017-00621 and 2021-00155). H.Z. and H.T.S. thank the Swedish Research Council for individual project grants (with contract nos. 2020-03437 and 2022-02822). R.D.T., H.T.S., and H.Z. acknowledge the project grant "Probing charge- and mass-transfer reactions on the atomic level" (2018.0028) from the Knut and Alice Wallenberg Foundation. R.D.T. acknowledges support by the Air Force Office of Scientific Research under award no. FA9550-19-1-7012. This publication is based upon work from COST Action CA18212 - Molecular Dynamics in the GAS phase (MD-GAS), supported by COST (European Cooperation in Science and Technology). D.S. and A.B. acknowledge support from ISF grant 674/21 and the Minerva Center for Making Bonds by Fragmentation. **Author contributions:** Investigation: A.B., M.P., M.J., A.D., A.S., S.R., H.Z., H.T.S., R.D.T., and D.S. Funding acquisition: D.S., H.T.S, H.Z., R.D.T. Supervision: D.S., H.Z., H.T.S., R.T., and D.S. Writing – original draft: A.B. and D.S. Writing – review and editing: A.B., M.P., A.D. H.Z., H.T.S., R.D.T., and D.S. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data underlying the figures are deposited in Zenodo (*68*) All data needed to evaluate the conclusions in the paper are present in the paper and/or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adk1950
Materials and Methods
Supplementary Text
Figs. S1 to S14
References (*69*–*77*)

Submitted 31 August 2023; accepted 13 December 2023
10.1126/science.adk1950

QUANTUM OPTICS

Logical states for fault-tolerant quantum computation with propagating light

Shunya Konno[1]†, Warit Asavanant[1,2]*, Fumiya Hanamura[1], Hironari Nagayoshi[1], Kosuke Fukui[1], Atsushi Sakaguchi[2], Ryuhoh Ide[1], Fumihiro China[3], Masahiro Yabuno[3], Shigehito Miki[3,4], Hirotaka Terai[3], Kan Takase[1,2], Mamoru Endo[1,2], Petr Marek[5], Radim Filip[5], Peter van Loock[6], Akira Furusawa[1,2]*

To harness the potential of a quantum computer, quantum information must be protected against error by encoding it into a logical state that is suitable for quantum error correction. The Gottesman-Kitaev-Preskill (GKP) qubit is a promising candidate because the required multiqubit operations are readily available at optical frequency. To date, however, GKP qubits have been demonstrated only at mechanical and microwave frequencies. We realized a GKP state in propagating light at telecommunication wavelength and verified it through homodyne measurements without loss corrections. The generation is based on interference of cat states, followed by homodyne measurements. Our final states exhibit nonclassicality and non-Gaussianity, including the trident shape of faint instances of GKP states. Improvements toward brighter, multipeaked GKP qubits will be the basis for quantum computation with light.

Quantum computers whose operations are based on quantum mechanics outperform classical computers in certain tasks, with recent demonstrations of quantum advantage being reported (*1*, *2*). Within a few years, noisy intermediate-scale quantum computers are expected to be available. To go beyond that, however, fault-tolerant quantum processors are required. To mitigate the errors in quantum computation, quantum error correction (QEC) is a promising strategy. QEC allows the detection and correction of errors without disturbing the quantum information by having additional redundancy in the system. QEC that uses bosonic systems has attracted attention owing to their large Hilbert spaces, which are capable of redundantly encoding qubits, and various bosonic codes such as binomial codes (*3*), cat codes (*4*), and Gottesman-Kitaev-Preskill (GKP) codes (*5*) are being explored.

In terms of wave functions, GKP states have a periodic sharp-peaked grid structure in both of their quadratures, which allows for the detection and correction of small loss and displacement errors that are dominant in the optical systems. Advantageously, the basic qubit-level operations on the GKP codes (also called Clifford operations), including entangling gates and error syndrome measurements, can be done using only linear operations. Moreover, universal quantum computation can be achieved by additional light squeezers (*6*), which have been already demonstrated in optical systems, and specifically designed GKP qubit states (*7*, *8*). The key technologies for fault-tolerant quantum computation using GKP qubits are based on quantum teleportation (*9*). Quantum teleportation provides the capability for linear operations with Bell measurement and feedforward displacement that correspond to syndrome measurement and recovery operation, respectively (*10*, *11*). Quantum teleportation technology is highly developed in the optical system (*12*, *13*), which makes it a promising platform for GKP states. The experimental demonstrations of the GKP qubits to date, however, are in stationary systems that can be coupled to qubits (*14*, *15*) providing strong nonlinearity. Conversely, linear operations are not naturally available and have to be constructed by engineering the nonlinearity of the system (*16*, *17*), which limits the scalability to a large-scale multimode interaction. This is in contrast to the optical system, in which linear operations can be simply done with commercially available beamsplitters and multimode linear operations can be easily implemented (*18*, *19*).

In addition to the linearity of the system, there are a few key properties required for the GKP states to be useful in actual quantum computation. The physical platform for using GKP states should allow large-scale and fast operations, because slow implementation of syndrome measurements and corrections increases the error rates and thus entails an additional overhead. Moreover, for applications such as quantum communication or interconnecting quantum computers using quantum internet, GKP states encoded in a propagating wave can be the core elements. Given this, a propagating electromagnetic wave system in the optical regime is a promising candidate. By multiplexing of degrees of freedom such as time (*20*, *21*) or frequency (*22*) in the propagating wave, a large-scale quantum computation platform has already been demonstrated. Also, terahertz-bandwidth light source (*23*), 43-GHz optical homodyne measurement (*24*), and high-speed nonlinear feedforward (*25*), which are key components to high-speed optical quantum computation and error correction, have been demonstrated; this means that we can expect near-term optical quantum computation with a clock frequency of at least a few gigahertz, which surpasses other physical systems by several orders of magnitude. Despite these appealing features, however, the actual optical generation of GKP qubits in a propagating optical system has remained elusive because propagating electromagnetic systems lack viable strong nonlinearity, and even if we try to obtain nonlinearity by using a system such as cavity QED (*26*), a complex arrangement would be required to realize a complex quantum state. By realizing GKP states in the optical system, we can overcome the limitations of multimode linear operations and scalability of the GKP approach in the nonlinear systems (*14*, *15*). Therefore, the realization of GKP states in the propagating wave is a key to practical quantum computation and is the final main ingredient of the fault-tolerant universal quantum computer using optical systems.

There are many research articles and proposals on how to generate optical GKP states [for example, see (*27*–*30*)]. Our method uses some of these concepts and generates the GKP state in a propagating wave system based on the two-mode interference between cat states and a single-mode projection obtained through an optical homodyne measurement (*28*, *29*). This method benefits from the fact that cat states are fairly well known and are experimentally accessible in the optical system (*31*). The required nonlinearity in the GKP state generation is introduced off-line by photon number measurements used in the generation of the cat states. The generated state is characterized through homodyne measurements and is reconstructed with quantum tomography. No corrections for experimental imperfection are used in either the homodyne measurements or quantum tomography. Homodyne measurement is a linear measurement that is required in the syndrome measurements and operations of GKP states. Homodyne measurement of the GKP states, however, has not been implemented in previous experimental demonstrations (*14*, *15*). Although the cat-state generation in this work is probabilistic, given the combination of the cutting-edge photon-number resolving detector (*32*) and methods for high-rate cat-state generation (*33*), the whole process can have a high success rate. Also, for a more complex multistep generation, feedforward displacement based

[1]Department of Applied Physics, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. [2]Optical Quantum Computing Research Team, RIKEN Center for Quantum Computing, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. [3]Advanced ICR Research Institute, National Institute of Information and Communications Technology, 588-2 Iwaoka, Nishi, Kobe 651-2492, Japan. [4]Graduate School of Engineering, Kobe University, 1-1 Rokko-dai, Nada, Kobe 657-0013, Japan. [5]Department of Optics, Palacky University, 17. listopadu 1192/12, 77146 Olomouc, Czech Republic. [6]Institute of Physics, Johannes-Gutenberg University of Mainz, Staudingerweg 7, 55128 Mainz, Germany.
*Corresponding author. Email: warit@alice.t.u-tokyo.ac.jp (W.A.); akiraf@ap.t.u-tokyo.ac.jp (A.F.)
†Present address: NTT Device Technology Laboratories, NTT Corporation, Atsugi-shi, Kanagawa 243-0198, Japan.

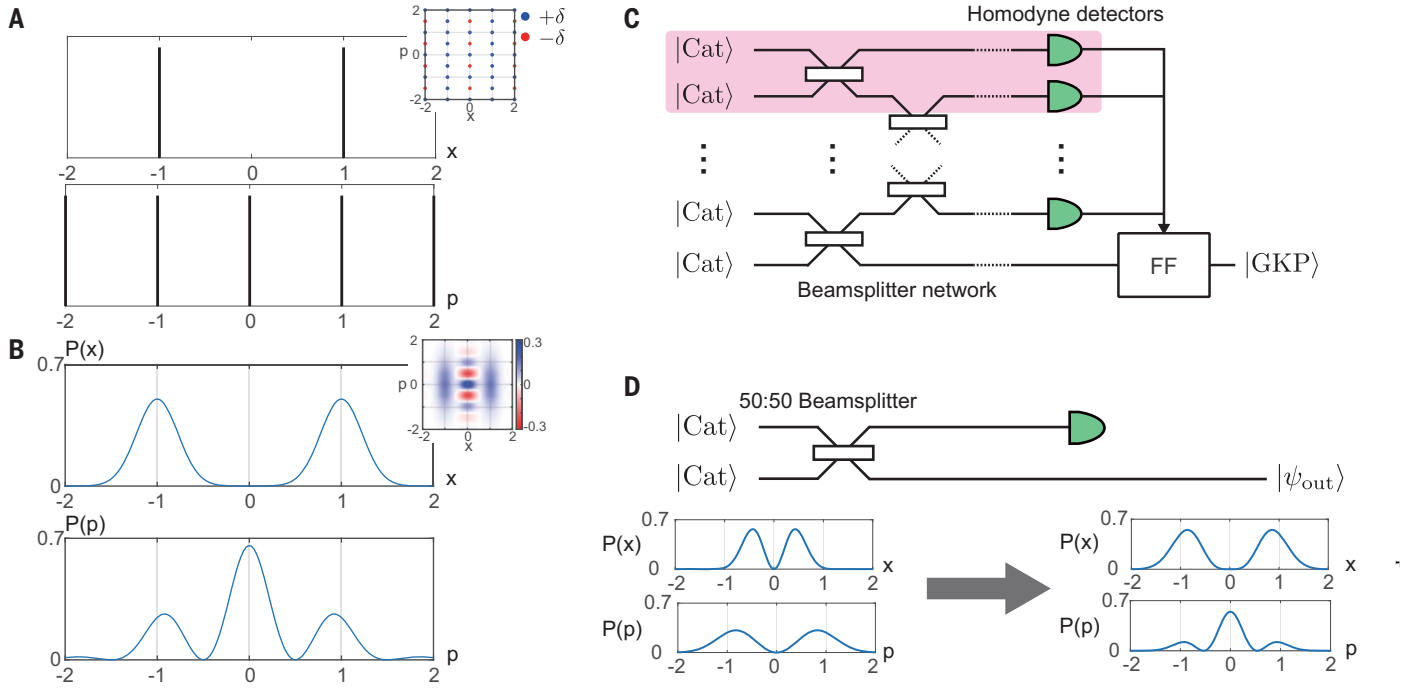


Fig. 1. GKP state and its generation method in the optical system. (**A** and **B**) Quadrature distributions of ideal (infinite energy) (A) and approximated GKP |1⟩ states (B) with 5 dB of squeezing (the units are defined with reduced Planck's constant $\hbar = 1$, and the quadratures are normalized to $\sqrt{\pi}$). The insets show the Wigner functions. δ, Delta function peak. (**C**) General generation method in the optical system based on the interference of cat states using a beamsplitter network and implementing homodyne measurements and feedforward operations (FFs) that incorporate squeezing and displacement operations. Light red shading indicates the single step of the generation method. (**D**) Single step of the generation method. Note that FFs with Gaussian operations are omitted in this figure and are replaced by conditioning in this experiment. The formation of peaks at the stabilizer values is visible.

on the results of the homodyne measurements can be done to remove the need for the homodyne conditioning (*29*), which makes the process after cat-state generation semideterministic.

GKP state and generation methodology

The GKP state can be defined in several ways; in this work, we consider a GKP state with square lattice structure. When we consider the basic observables of light—quadrature operators $\hat{x}$ and $\hat{p}$ ($[\hat{x}, \hat{p}] = i$)—the Pauli operators on the phase-space grid are $\hat{\tilde{X}} = \exp(-i\sqrt{\pi}\hat{p})$ and $\hat{\tilde{Z}} = \exp(i\sqrt{\pi}\hat{x})$, which correspond to displacement in x and p by $\sqrt{\pi}$, respectively. Then, the GKP states are defined in the logical space that is stabilized, that is, invariant, under the operators $\hat{S}_x = \hat{\tilde{X}}^2 = \exp(-i2\sqrt{\pi}\hat{p})$ and $\hat{S}_p = \hat{\tilde{Z}}^2 = \exp(i2\sqrt{\pi}\hat{x})$, which are displacements by $2\sqrt{\pi}$ in the x and p directions, respectively (*5*). Alternatively, for any ideal GKP states in the logical space, we have $\langle \hat{S}_x \rangle = \langle \hat{S}_p \rangle = 1$. This means that ideal GKP states should have periodicity of $2\sqrt{\pi}$ in both x and p. Defining these two stabilizers does not uniquely determine the states; they determine a logical space in which the information is encoded, and we have to consider another stabilizer to define our state. We can define our basis for the logical space, that is, the eigenstates for $\hat{\tilde{Z}}$, as $|0_L\rangle \propto \sum_k |x = 2k\sqrt{\pi}\rangle$ and $|1_L\rangle \propto \sum_k |x = (2k+1)\sqrt{\pi}\rangle$, where the sum k is over all integers. These two states have $\hat{\tilde{Z}}$ and $-\hat{\tilde{Z}}$ as their stabilizers, respectively, making one of the two GKP code stabilizers $\hat{S}_p$ redundant. Note that $\pm\hat{\tilde{Z}}$ commutes with both $\hat{S}_x$ and $\hat{S}_p$. Because these states are unnormalizable, they are usually approximated by replacing the position eigenstates with squeezed states, and the whole states are enveloped with a Gaussian envelope for the symmetry in both the x and p quadrature (*34*). Figure 1, A and B, shows the grid quadrature distribution of an ideal and approximated GKP state.

We generated the peak structure of the GKP states by using the interference of kitten states and homodyne measurements (*28*, *29*). Figure 1C shows a diagram of the general generation method. As an example, let us consider a single interference step (shown in Fig. 1D). If we interfere two cat states $|\psi_{cat}\rangle \propto |i\alpha\rangle - |-i\alpha\rangle$, where $|i\alpha\rangle$ and $|-i\alpha\rangle$ are coherent states with α being a real number, the output two-mode state is $(|i\sqrt{2}\alpha\rangle + |-i\sqrt{2}\alpha\rangle)|0\rangle - |0\rangle(|-i\sqrt{2}\alpha\rangle + |i\sqrt{2}\alpha\rangle)$. If we measure the first mode and condition it at $x = 0$, we will approach $|\psi_{out}\rangle \propto |-i\sqrt{2}\alpha\rangle - 2|0\rangle + |i\sqrt{2}\alpha\rangle$. The output state now approaches the central part of the GKP state in the phase space (up to additional squeezing), and the process can be repeated so that a larger part of the GKP state is synthesized. The coefficients of the peaks in this method can be easily shown to be binomial coefficients that approach Gaussian distribution for a large number of peaks (*28*). To make the width of the peaks small and the distance between the peaks correct as a GKP state, one can either implement the squeezing operation at the end or start with squeezed cat states instead of an unsqueezed cat state (*28*, *29*). Such Gaussian operations only shape the GKP state and do not increase the non-Gaussian grid aspects. Although we use ordinary cat states for our experimental demonstration, squeezing operations on a cat state (*6*) and the high-rate generation of squeezed cat states (*33*) are both being explored in optics.

Experimental results

Figure 2 shows the experimental system. The master laser of the system is a continuous-wave laser with a wavelength of 1545.32 nm and a second-harmonic generator for the generation of 772.66-nm light. The kitten states for the interference are generated by using photon subtraction on squeezed light (*35*). The squeezed light sources are optical parametric oscillators (OPOs). We detected the photon in the photon subtraction using superconducting nanostrip single-photon detectors (SNSPDs). The coincidence detection at both SNSPDs heralded success in the interference. We measured one of the

modes in the x basis using a homodyne detector so that the GKP state was generated in the other mode. The signal of the SNSPDs acted as the measurement trigger for the oscilloscope, and the electrical signals of the homodyne detectors were measured in real time. We measured the quadrature of the other mode at various phases and used the collected data to reconstruct the state by quantum tomography. No correction for optical losses or any experimental imperfections were done on the measurement results or during the tomography process. Note that in this work, we only implemented a single step, and iterations need to be done in the future to realize high-quality GKP states (*28*, *29*). Also, we replaced the additional displacement operation for correcting the effects of homodyne measurement with conditioning of the homodyne measurement outcomes. See fig. S1 for details on the experimental systems (*36*).

For GKP states to be useful, they must be able to detect and correct the error displacement in both the x and p quadratures. Although it is prohibited by the uncertainty principle for both quadratures to be sharply defined, locally we can have a Wigner function with sharp peaks. If we consider the Wigner function in Fig. 3 and consider the center positive peak, where all nearby positive regions are considered as a part of the peak, the variances in each quadrature are $\langle\Delta^2\hat{x}\rangle$=1.45±0.03, $\langle\Delta^2\hat{p}\rangle$ = 0.070±0.001, and $\langle\Delta^2\hat{x}\rangle\langle\Delta^2\hat{p}\rangle$ = 0.100±0.002, which is well within the sub-Planck regime (*37*) ($\langle\Delta^2\hat{x}\rangle\langle\Delta^2\hat{p}\rangle < 0.25$). A sub-Planck regime that occurs simultaneously in both quadratures is a regime that cannot be reached

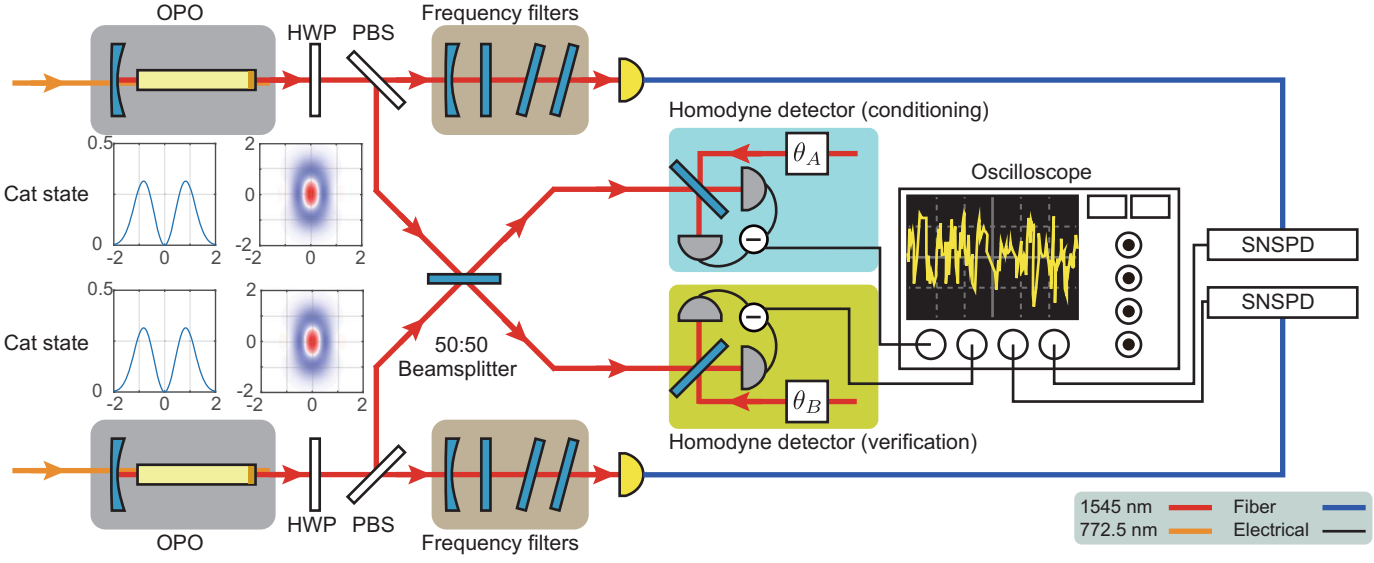


Fig. 2. Experimental setup. Schematic of the experimental setup used in this work. The insets show the ideal states from the photon subtraction with the quadrature distribution in p. The phases of the two homodyne measurements are given by θ_A and θ_B, where θ_A is set to 0° for measuring and conditioning of x quadrature and θ_B is set to 0°, 30°, 60°, 90°, 120°, or 150° for collecting quadrature values for quantum tomography. HWP, half-wave plate; PBS, polarization beamsplitter.

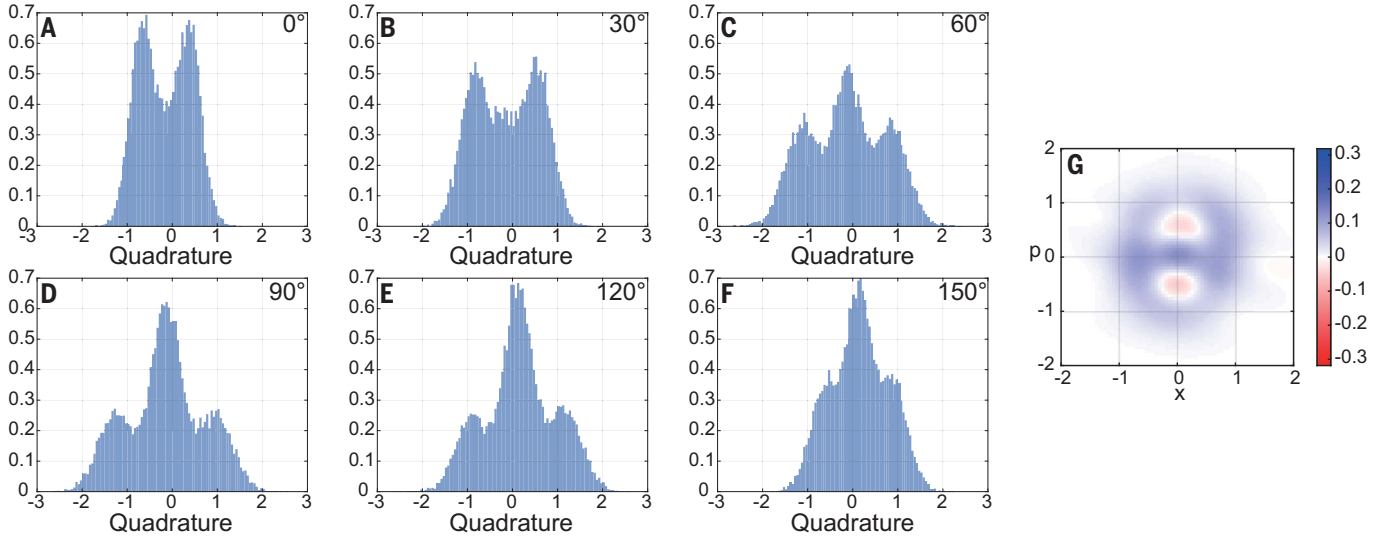


Fig. 3. Quadrature distributions and reconstructed Wigner function. The experimentally measured quadrature distributions are normalized as a probability density function. No corrections were performed on the obtained distributions. (**A** to **F**) Quadrature distributions of the generated states. The homodyne phases are provided in each panel. (**G**) The reconstructed Wigner function. The Gaussian corrections, which correspond to a linear symplectic transformation in phase space, are implemented on the reconstructed Wigner function [see (*36*)]. The color scale bar indicates the values of Wigner function.

with a simple Gaussian state, and this sub-Planck structure is qualitative evidence that our approach can synthesize the sharp delta-function peak structure of the GKP state. Note that the value $\langle \Delta^2 \hat{x} \rangle$ is limited by the amplitude of the initial kitten state and can be improved by using a cat state with a bigger amplitude.

Regarding the quantitative evaluation, we assessed the stabilizers of the generated state. As we mentioned, for our specific state $|1_L\rangle$, the two stabilizers become $\hat{S}_x$ and $\hat{S}_{|1\rangle} = -\hat{Z}$. We calculated the average value of the stabilizers from the generated states, and they are $\langle \hat{S}_x \rangle$=0.170±0.003 and $\langle \hat{S}_{|1\rangle} \rangle = 0.216 \pm 0.006$, respectively. Although these values are still far from unity, they clearly surpass the values that can be achieved with classical states. Furthermore, their collective properties overcome the limits set by Gaussian states. See the supplementary materials for a more detailed analysis and discussion.

Conclusions and discussion

The effective squeezing of our optical GKP state is estimated to be about 2.5 dB, which is, at present, lower than those reported in the trapped ion (5.5 and 7.3 dB) (*14*) and microwave (between 7.4 and 9.5 dB) (*15*). Presently, the limiting factor is the optical loss in the system and the lack of iteration in our demonstration. The mode mismatch between two OPOs is also a relevant source of imperfection, but it can be removed by properly tailoring the temporal mode (*36*). In addition, we also need to make a cat state with larger amplitude and concatenate the process with deeper circuit depth, as depicted in Fig. 1, to increase this effective squeezing. The generalized photon subtraction (*33*) is a promising approach for this because the generated cat states are squeezed and the success rate can be three to six orders of magnitude greater than that of the conventional method. Furthermore, it is possible to use the time-domain multiplexing technique to increase the circuit depth in a hardware-efficient way (*31*). Our present setup has a success rate of about 10 Hz, which is as expected from the experimental parameters (*36*). In addition to the aforementioned generalized photon subtraction, improving the squeezed light source and the photon-number resolving detector can increase the success rate. Regarding the former, improvement can be achieved by replacing the OPO in this work with a terahertz-bandwidth optical parametric amplifier (*23*). Regarding the latter, although the photon counter in this experiment is based on SNSPDs without photon-number resolution, with proper technique, SNSPDs can be used for counting photons with high speed (timing jitter of ~20 ps) (*38*). Using these SNSPDs would allow an even higher generation rate of the cat state for GKP-state generation.

Finally, we briefly comment on how optical GKP states generated in the manner described in this work can be used in actual quantum computation. Because the GKP states generated by this method are a propagating wave, they can be coupled into the various types of optical quantum processors seamlessly without requiring any additional processes such as wavelength conversion. Multimode linear operations (*18*, *19*), quantum teleportation (*12*, *13*), and squeezing operation (*6*) on a propagating optical field have already been demonstrated and implemented, and these technologies can be used to implement syndrome measurement and recovery operations on a GKP state.

REFERENCES AND NOTES

1. H. Wang *et al*., *Phys. Rev. Lett.* **123**, 250503 (2019).
2. H.-S. Zhong *et al*., *Science* **370**, 1460–1463 (2020).
3. M. H. Michael *et al*., *Phys. Rev. X* **6**, 031006 (2016).
4. P. T. Cochrane, G. J. Milburn, W. J. Munro, *Phys. Rev. A* **59**, 2631–2634 (1999).
5. D. Gottesman, A. Kitaev, J. Preskill, *Phys. Rev. A* **64**, 012310 (2001).
6. Y. Miwa *et al*., *Phys. Rev. Lett.* **113**, 013601 (2014).
7. S. Konno *et al*., *Phys. Rev. Res.* **3**, 043026 (2021).
8. B. Q. Baragiola, G. Pantaleoni, R. N. Alexander, A. Karanjai, N. C. Menicucci, *Phys. Rev. Lett.* **123**, 200502 (2019).
9. N. C. Menicucci, *Phys. Rev. Lett.* **112**, 120504 (2014).
10. B. W. Walshe, R. N. Alexander, N. C. Menicucci, B. Q. Baragiola, *Phys. Rev. A* **104**, 062427 (2021).
11. W. Asavanant, K. Fukui, A. Sakaguchi, A. Furusawa, *Phys. Rev. A* **107**, 032412 (2023).
12. A. Furusawa *et al*., *Science* **282**, 706–709 (1998).
13. N. Lee *et al*., *Science* **332**, 330–333 (2011).
14. C. Flühmann *et al*., *Nature* **566**, 513–517 (2019).
15. P. Campagne-Ibarcq *et al*., *Nature* **584**, 368–372 (2020).
16. Y. Y. Gao *et al*., *Nature* **566**, 509–512 (2019).
17. X. Pan *et al*., *Phys. Rev. X* **13**, 021004 (2023).
18. W. Asavanant *et al*., *Phys. Rev. Appl.* **16**, 034005 (2021).
19. M. V. Larsen, X. Guo, C. R. Breum, J. S. Neergaard-Nielsen, U. L. Andersen, *Nat. Phys.* **17**, 1018–1023 (2021).
20. W. Asavanant *et al*., *Science* **366**, 373–376 (2019).
21. M. V. Larsen, X. Guo, C. R. Breum, J. S. Neergaard-Nielsen, U. L. Andersen, *Science* **366**, 369–372 (2019).
22. O. Pfister, *J. Phys. At. Mol. Opt. Phys.* **53**, 012001 (2019).
23. T. Kashiwazaki *et al*., *Appl. Phys. Lett.* **119**, 251104 (2021).
24. A. Inoue *et al*., *Appl. Phys. Lett.* **122**, 104001 (2023).
25. A. Sakaguchi *et al*., *Nat. Commun.* **14**, 3817 (2023).
26. V. Magro, J. Vaneecloo, S. Garcia, A. Ourjoumtsev, *Nat. Photonics* **17**, 688–693 (2023).
27. I. Tzitrin, J. E. Bourassa, N. C. Menicucci, K. K. Sabapathy, *Phys. Rev. A* **101**, 032315 (2020).
28. H. M. Vasconcelos, L. Sanz, S. Glancy, *Opt. Lett.* **35**, 3261–3263 (2010).
29. D. J. Weigand, B. M. Terhal, *Phys. Rev. A* **97**, 022341 (2018).
30. K. Fukui *et al*., *Phys. Rev. Lett.* **128**, 240503 (2022).
31. W. Asavanant, A. Furusawa, *Optical Quantum Computers: A Route to Practical Continuous Variable Quantum Information Processing* (AIP Publishing LLC, 2023).
32. M. Endo *et al*., *Opt. Express* **31**, 12865–12879 (2023).
33. K. Takase, J. Yoshikawa, W. Asavanant, M. Endo, A. Furusawa, *Phys. Rev. A* **103**, 013710 (2021).
34. T. Matsuura, H. Yamasaki, M. Koashi, *Phys. Rev. A* **102**, 032408 (2020).
35. M. Dakna, T. Anhut, T. Opatrný, L. Knöll, D.-G. Welsch, *Phys. Rev. A* **55**, 3184–3194 (1997).
36. See supplementary materials.
37. W. H. Zurek, *Nature* **412**, 712–717 (2001).
38. M. Endo *et al*., *Opt. Express* **29**, 11728–11738 (2021).
39. S. Konno *et al*., Propagating Gottesman-Kitaev-Preskill states encoded in an optical oscillator. Dryad (2023); https://doi.org/10.5061/dryad.t76hdr86j.

ACKNOWLEDGMENTS

Funding: This work was partly supported by the Japan Science and Technology (JST) Agency (Moonshot R&D) grants JPMJMS2064 and JPMJMS2066; by the UTokyo Foundation; and by donations from the Nichia Corporation of Japan. S.K. acknowledges funding from the Japan Society for the Promotion of Science (JSPS) KAKENHI (no. 21J11615). W.A. acknowledges funding from JSPS KAKENHI (no. 23K13040). M.E. acknowledges funding from JST (JPMJPR2254). W.A. and M.E. acknowledge support from the Research Foundation for OptoScience and Technology. P.v.L. acknowledges funding from the Federal Ministry of Education and Research (BMBF) in Germany (QR.X, PhotonQ, and QuKuK) the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 429529648 – TRR 306 QuCoLiMa ("Quantum Cooperativity of Light and Matter"), from the European Union and BMBF through QuantERA (ShoQC), and from the European Union's HORIZON Research and Innovation Actions (CLUSTEC). P.M. acknowledges support from grant no. 22-08772S of the Czech Science Foundation (GACR) and the European Union's HORIZON Research and Innovation Actions under grant agreement no. 101080173 (CLUSTEC). R.F. acknowledges project 21-13265X of the Czech Science Foundation. P.M. and R.F. acknowledge EU H2020-WIDESPREAD-2020-5 project NONGAUSS (951737) under the Coordination and Support Action (CSA). **Author contributions:** S.K. led the experiment with supervision from W.A., K.T., M.E., and A.F. S.K. collected the experimental data. S.K. and W.A. analyzed the data. W.A. visualized the data for the manuscript. Theoretical discussions and interpretations of the data were done by W.A., P.v.L., R.F., P.M., F.H., H.N., K.F., A.S., and R.I. The criteria in the supplementary materials was developed by P.M., W.A., and R.F. H.T., S.M., M.Y., and F.C. provided the SNSPDs that were used in this experiment. W.A. wrote the manuscript and the supplementary materials with the help of P.v.L., A.F., P.M., R.F., and the other authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The quadrature obtained and used in this paper is available at Dryad (*39*). Additional data are available in the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adk7560
Materials and Methods
Supplementary Text
Figs. S1 to S4
References (*40*–*44*)

Submitted 8 September 2023; accepted 16 November 2023
10.1126/science.adk7560

BIOGEOGRAPHY

The global distribution of plants used by humans

S. Pironon[1,2]†‡*, I. Ondo[1,2]‡, M. Diazgranados[1,3], R. Allkin[1], A. C. Baquero[2], R. Cámara-Leret[4], C. Canteiro[1], Z. Dennehy-Carr[1,5], R. Govaerts[1], S. Hargreaves[1], A. J. Hudson[6,7], R. Lemmens[8], W. Milliken[6], M. Nesbitt[1,9,10], K. Patmore[1], G. Schmelzer[8], R. M. Turner[1], T. R. van Andel[8,11], T. Ulian[6,12], A. Antonelli[1,13,14]§, K. J. Willis[1,14]§

Plants sustain human life. Understanding geographic patterns of the diversity of species used by people is thus essential for the sustainable management of plant resources. Here, we investigate the global distribution of 35,687 utilized plant species spanning 10 use categories (e.g., food, medicine, material). Our findings indicate general concordance between utilized and total plant diversity, supporting the potential for simultaneously conserving species diversity and its contributions to people. Although Indigenous lands across Mesoamerica, the Horn of Africa, and Southern Asia harbor a disproportionate diversity of utilized plants, the incidence of protected areas is negatively correlated with utilized species richness. Finding mechanisms to preserve areas containing concentrations of utilized plants and traditional knowledge must become a priority for the implementation of the Kunming-Montreal Global Biodiversity Framework.

Biodiversity provides essential goods and services that sustain human life and well-being (e.g., food, medicines, materials, fuel) (*1*, *2*). The balance between humanity's needs and the protection of the natural environment is nevertheless fragile, as increased consumption of resources, global trade, land- and sea-use change, and socio-economic inequalities are having a marked influence on biodiversity (*3*, *4*). To minimize biodiversity loss, conservation biologists have focused on identifying and prioritizing regions of high species richness, endemism, and threat (*5*). The "biodiversity hotspot" concept (*6*) assumes that species diversity is spatially congruent with the contributions that it provides to people and therefore, protecting areas with the largest concentrations of threatened species will also protect humanity indirectly (*5*). Moreover, as biodiversity is most concentrated where human cultural diversity is highest, it is assumed that high biocultural diversity is associated with high concentrations of species used by humans (*7*). Yet, these assumptions lack empirical support, leading to growing calls for better integration of human–nature interactions into conservation planning and implementation (*3*, *8*–*10*), as highlighted by the recently adopted Kunming-Montreal Global Biodiversity Framework (GBF) and the 2022 assessment report on the sustainable use of wild species of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) (*2*).

Plants are essential structuring components of ecosystems and human livelihoods (*9*, *11*). Although the geography of terrestrial plant diversity has been extensively investigated globally (*6*, *12*, *13*), our understanding of the distribution of ecosystem services and societal benefits provided by plants is incipient, despite the importance of this information for decision-makers and local stakeholders in supporting the sustainable development agenda (*14*, *15*). Recent modeling efforts have been dedicated to the global distribution of nature's contributions to people, including water quality, crop pollination, and carbon stocks (*16*, *17*). However, the extent to which these contributions relate to species diversity remains largely unknown, hampering progress toward a more sustainable management of biodiversity. Assessing the global diversity and distribution of plant species used by people is thus critical to better understand, manage, and preserve both the intrinsic and instrumental values of biodiversity (*18*).

The global distribution of utilized plant species richness and endemism

Most plant species can potentially be useful to people, but only a fraction of plant diversity is currently known to be used. Here we consider utilized plants as vascular terrestrial species for which material and nonmaterial benefits to humans have been reported and made publicly accessible (*19*, *20*). By extracting information from 12 databases containing plant uses (table S1) (*21*), we identified 35,687 utilized species and assembled >11 million georeferenced occurrence records to map their global distribution (i.e., native and introduced ranges) (figs. S1 and S2) (*19*). We built species distribution models for each utilized species and stacked the resulting maps to assess the global distribution of their potential richness (figs. S3 to S6) (*19*). We find the highest concentrations of utilized plant species in the tropics (Fig. 1), but several temperate areas also contain high native (e.g., China, the Himalayas; fig. S7) and introduced richness (e.g., Western Europe, Eastern USA; Fig. 1). Despite large discrepancies in the sampling of species geographic records (fig. S1) (*22*), these results match our estimates using coarser but more complete independent distribution data from the World Checklist of Vascular Plants (WCVP) (*23*) (fig. S8), which provides additional support for our predictions.

Distribution patterns in species richness do not systematically match those of other biodiversity indices considered important for conservation such as rarity or threat (*5*, *6*). Therefore, we also estimated the distribution of utilized plant species richness weighted by each species' range size (i.e., weighted endemism) to identify areas with high concentrations of rare and potentially irreplaceable species. We find that many areas with high richness of utilized plant species also exhibit high endemism (e.g., Mesoamerica, Gulf of Guinea, Southern Africa, the Himalayas, Southeast Asia; Fig. 1 and fig. S8). Other areas also to emerge as exceptional centers of endemic utilized plant species include California, Macaronesia, Madagascar, the Eastern Mediterranean, the Western Ghats, Sri Lanka, Eastern Australia, and the Pacific islands. Conversely, concentrations of endemic utilized species are relatively low across temperate areas. This confirms that the high species richness observed in some temperate regions is due to a high concentration of well-surveyed, widely distributed, and often-introduced plant species of economic importance (*22*, *24*). Overall, the distribution of utilized plant endemicity mirrors patterns observed across all vascular plants, with higher endemism in areas with insularity and high topographic and environmental heterogeneity (*25*, *26*).

The latitudinal distribution of utilized plant species and their different uses

To refine our understanding of the geographic patterns underpinning the diversity of utilized plant species, we disaggregated plant-use reports into 10 use categories, adapted from the Economic Botany Data Collection Standards: human food (including beverages and additives), vertebrate food (forage and fodder), invertebrate food (e.g., plants feeding honey bees or silkworms), materials (e.g., wood, fiber), fuels (e.g.,

[1]Royal Botanic Gardens, Kew, Richmond, Surrey, UK. [2]UN Environment Programme World Conservation Monitoring Centre (UNEP-WCMC), Cambridge, UK. [3]International Plant Science Center, New York Botanical Garden, New York, NY, USA. [4]Department of Systematic and Evolutionary Botany, University of Zurich, Switzerland. [5]Herbarium, School of Biological Sciences, University of Reading, Whiteknights, UK. [6]Royal Botanic Gardens, Kew, Wakehurst, Ardingly, UK. [7]Botanic Gardens Conservation International, Richmond, UK. [8]Wageningen University and Research, Wageningen, Netherlands. [9]Department of Geography, Royal Holloway, University of London, Egham, UK. [10]Institute of Archaeology, University College London, London, UK. [11]Naturalis Biodiversity Center, Leiden, Netherlands. [12]Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy. [13]Gothenburg Global Biodiversity Centre, Department of Biological and Environmental Sciences, University of Gothenburg, Gothenburg, Sweden. [14]Department of Biology, University of Oxford, Oxford, UK.
*Corresponding author. Email: s.pironon@kew.org
†Present address: School of Biological and Behavioural Sciences, Queen Mary University of London, London, UK.
‡These authors contributed equally to this work.
§These authors contributed equally to this work.

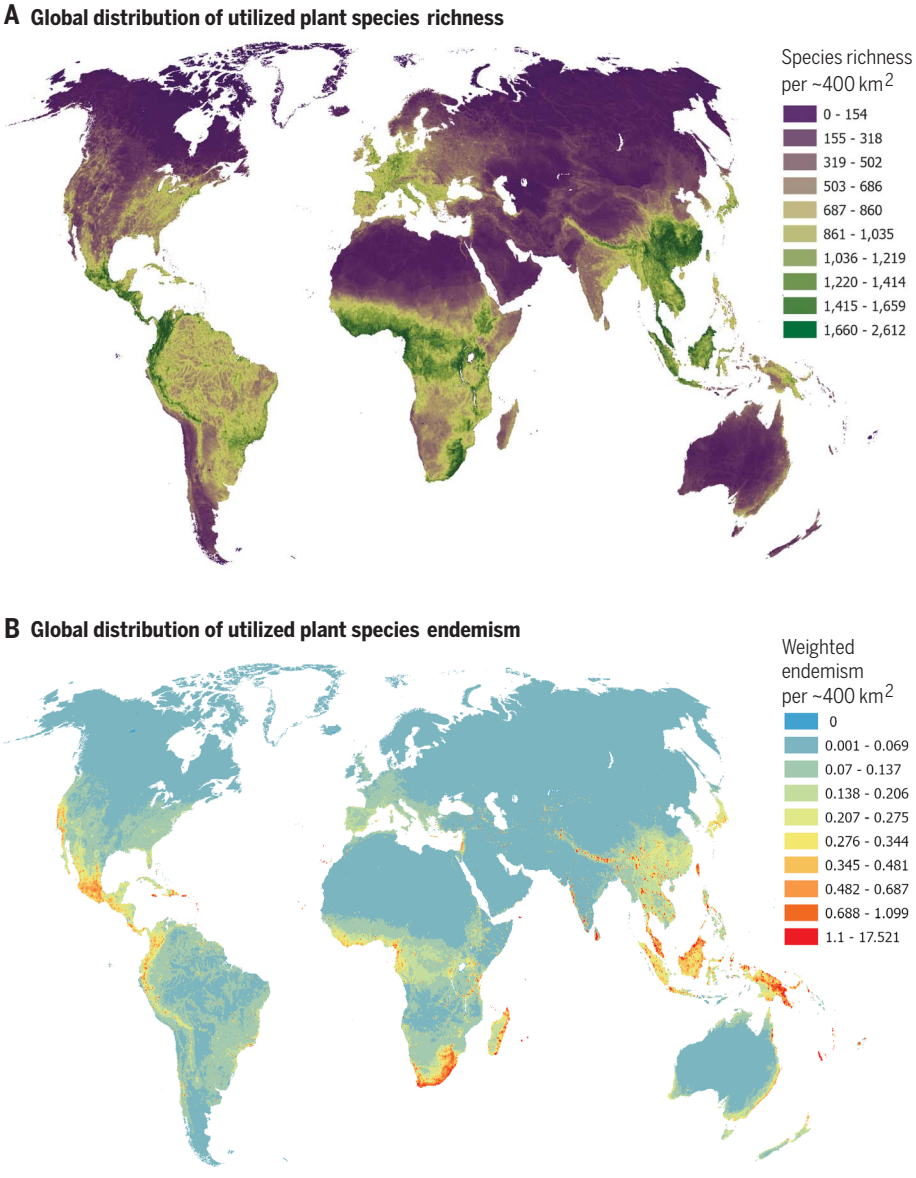


Fig. 1. Global species richness and endemism of plants with known uses by humans. (**A**) Utilized plant species richness corresponds to the sum of species occurrence probabilities predicted in each 10 arc min (~20 km) pixel found across their native and introduced ranges. (**B**) Utilized plant species endemism corresponds to the sum of species occurrence probabilities predicted in each pixel weighted by the inverse of their range size, calculated as the sum of the predicted probabilities within their study region (i.e., weighted endemism). High values are thus associated with areas containing high concentrations of species with small geographic ranges.

charcoal, alcohol), social uses (e.g., narcotics, ritual, religious uses), poisons (for both vertebrates and invertebrates), medicines (for both human and veterinary use), environmental uses (e.g., intercrops, windbreaks, ornamentals), and gene sources (e.g., crop wild relatives) (*19*, *21*). We find that latitudinal variation of utilized plant species richness is broadly consistent for all 10 use categories, with higher values in the tropics gradually declining toward high latitudes (Fig. 2 and figs. S9 and S10). Therefore, areas with high concentrations of utilized plant species also contain large numbers of species for each use type. Despite the overall similarity in latitudinal patterns among use categories, there are notable differences among temperate regions that are proportionally richer in plant species associated with vertebrate food, social uses, and poisons, compared to species-rich tropical environments that contain proportionally more species associated with the most essential uses for human subsistence (i.e., human food, material, and medicine). Concentrations of species used as gene sources are exceptionally high around the equator and thus diverge from domestication centers originally proposed by Vavilov (*27*). This is due to our consideration of a larger set of both domesticated species and wild relatives of potential interest for contemporary breeding programs (*15*). Utilized plant weighted endemism also follows a latitudinal gradient with larger relative concentrations of species at higher latitudes and consistent latitudinal variation among uses (figs. S11 to S13).

Spatial concordance between utilized plant species, total plant species, and human cultures

Although quantitative evidence is scarce, areas of high plant diversity are expected to contain more species that are beneficial to human populations (*5*). Our global analyses at (sub)country resolution indicate that utilized plant species richness is strongly positively associated with total plant species richness (t value = 20.703, $P < 0.001$; figs. S14 and S15 and table S2), and that this relationship holds for all categories of uses and for endemism (fig. S15 and table S3). It also highlights that large proportions of the flora of relatively low-diversity regions have documented uses (e.g., Scandinavia, Canada, Sahel), whereas smaller proportions of utilized species are reported across megadiverse regions (e.g., Madagascar, Brazil, tropical Andes; fig. S14). Future investigation will be required to identify whether this pattern is due to sampling gaps in our database or in the wider literature for these regions, or because the areas have reached a maximum capacity of utilized plant species richness. Overall, our findings substantiate the combined importance of preserving hotspots of plant diversity, which contain not only many distinctive species but also a considerable diversity of potential services for humanity (*1*). Nevertheless, although the spatial concordance between total plant diversity and utilized plant diversity is evident at a global scale, it is now crucial to assess whether this pattern holds at smaller scale, where political decisions are taken and management strategies implemented (*28*, *29*).

Biodiversity and cultural diversity have been shown to be highly intertwined spatially, giving rise to the notion of biocultural diversity (*7*, *30*). Our data suggest that cultural diversity not only correlates with total plant richness, but also covaries with utilized plant species richness (t value = 5.743, $P < 0.001$; figs. S14 and S15 and table S2) and inconsistently with endemism indices (fig. S15 and table S3). This finding supports previous hypotheses that geographic similarities between biodiversity and cultural diversity could be due to increased competition or reduced necessity for collaboration among human populations when biological resources (including plants) are widely available, ultimately causing social separation and

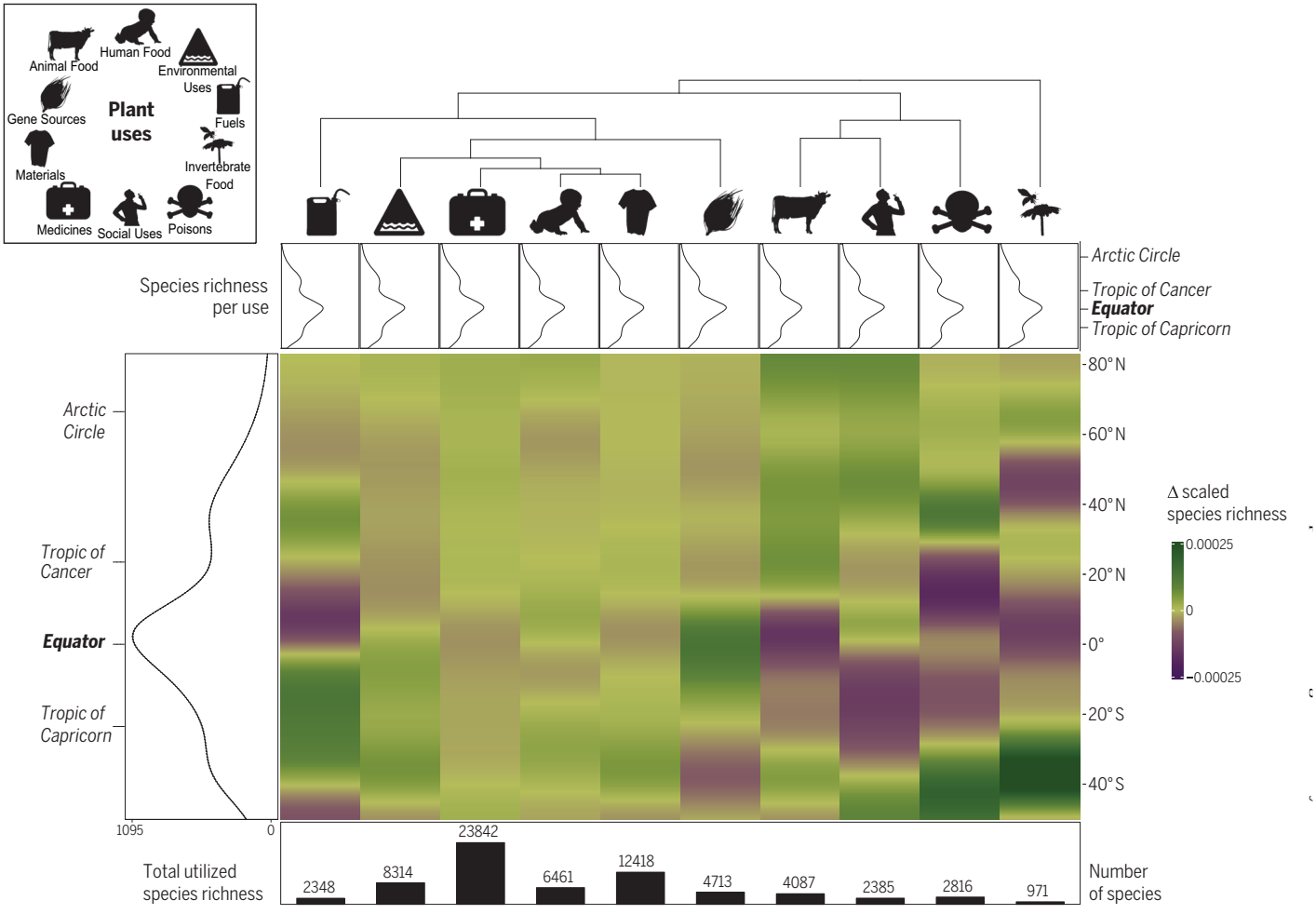


Fig. 2. Latitudinal distribution of utilized plant species richness across 10 categories of plant uses. The black curve on the left represents the latitudinal distribution of all utilized plant species richness. The dendrogram on the top orders the 10 use categories according to the (dis-)similarity of their species richness latitudinal profiles. Black curves underneath the dendrogram correspond to the species richness latitudinal profile for each use category. The heatmap describes the latitudinal variation in the deviation of utilized plant species richness for the 10 plant-use categories from total utilized plant species richness. Colors indicate higher (green) or lower (purple) proportions in utilized plant species richness of a given use relative to the total utilized plant species richness pattern. The bar chart underneath the heatmap shows the number of species considered in each use category.

generating greater linguistic diversity (*7*, *30*, *31*). However, other historical, evolutionary, and environmental factors may also be involved, and the identity and directionality of causal links for these correlations remain elusive and deserve future investigation (*7*, *30*).

Indigenous Peoples and protected areas: Preserving plant diversity and its contributions to people

Indigenous Peoples are particularly dependent on wild species for subsistence and well-being, in addition to being critical custodians of both plant diversity and traditional knowledge (*32*). Unexpectedly, at a large spatial scale, we find that the lands over which the estimated >370 million Indigenous Peoples of the world exert traditional rights do not contain higher concentrations of plant species with globally documented uses compared to neighboring non-Indigenous regions (Fig. 3 and fig. S17). This finding may reflect the fact that many Indigenous Peoples have been dispossessed of their lands throughout history (*33*), including biologically diverse areas, and that the largest remaining Indigenous territories are located in remote areas of low primary productivity (e.g., Greenland, Siberia, the Tibetan plateau, the Sahara, Sahel, Central Australia) (*34*). Exceptions include Indigenous lands located in multiple biocultural hotspots that harbor higher utilized plant species richness and endemism than do surrounding non-Indigenous regions: Central America, the Horn of Africa, South and Southeast Asia. Although Indigenous areas containing exceptionally high utilized plant diversity should be considered priorities for the joint conservation of nature and traditional knowledge (*34*, *35*), Indigenous lands containing fewer species should not be overlooked given that local populations may be particularly vulnerable to changing environmental conditions and species losses (*36*). Fostering the engagement of Indigenous, local, and scientific knowledge systems will be essential for enhancing ethics and actions toward protection at multiple scales (*37*).

Protected areas are at the forefront of global actions to preserve biodiversity and drive sustainable development (*38*). However, despite currently covering ~17% of Earth's terrestrial surface, the protected area network contributes to the conservation of a small fraction of plant diversity and ecosystem services (*16*). Spatial correlations between the proportion of land that is protected, and utilized plant species richness and endemism, indicate that regions with large protected area networks do not contain higher numbers or more distinctive utilized plant species than their nonprotected

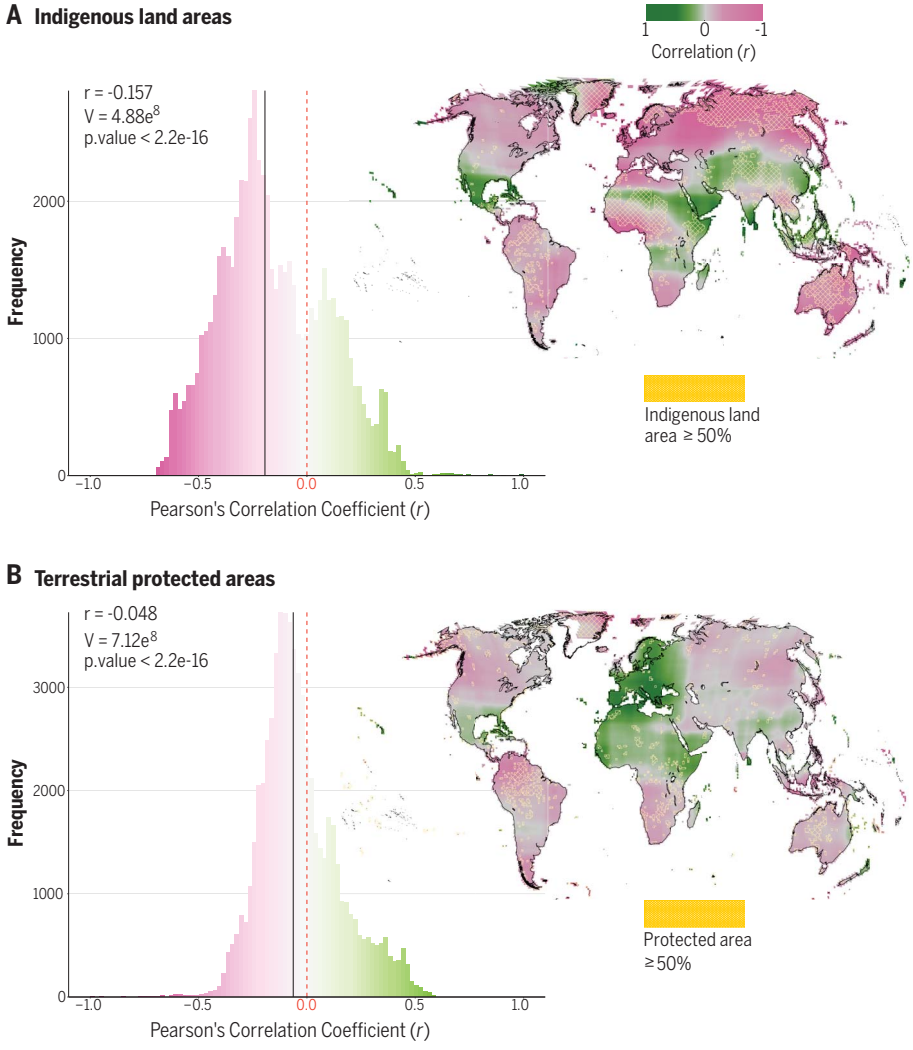


Fig. 3. Spatial correlations between utilized plant species richness and proportions of both Indigenous lands and terrestrial protected areas. (**A** and **B**) Pearson's correlation coefficients were computed across all values contained in 71 cells (~3550 km)–wide windows built around each pixel. Pixel color indicates regions where utilized plant species richness is positively (green) or negatively (pink) correlated with the proportion of Indigenous lands and terrestrial protected areas. Regions crossed in beige indicate pixels containing more than 50% of Indigenous lands and protected areas. All Indigenous lands and protected areas are thus not represented on the maps, although they are all accounted for in the analyses. Frequencies of Pearson's correlation coefficients found across the world are given in histograms. The median correlation across the world is indicated by the black vertical line, and zero correlation is indicated by the red dashed line. One-sample Wilcoxon signed rank tests were performed to assess whether median correlations are significantly different from zero.

counterparts (Fig. 3 and fig. S17). Indeed, although protected areas in Europe, the Mediterranean, West Africa, and the Horn of Africa contain more and more distinctive utilized plant species than nonprotected neighboring regions, several regions exhibit higher relative richness and endemism of utilized plant species outside of protected areas (e.g., Americas, Southern Africa, Southeast Asia, Australia). Our results point to the urgent need of considering plant diversity and its contributions to people in future area-based conservation planning (*10*, *29*, *39*), especially under the ambitious Target 4 of the GBF, which aims to conserve biodiversity across 30% of global land areas by 2030 (*40*). The latter also acknowledges the importance of "recognizing and respecting the rights of Indigenous Peoples and local communities" and "ensuring that any sustainable use…is fully consistent with conservation outcomes." In this context, it is essential to strike an appropriate balance between strictly protected areas that limit access to humans, and protected areas that accommodate the sustainable use of natural resources by local populations while preserving their well-being and cultural heritage (*41*, *42*).

Halting the overexploitation of species and ensuring their sustainable use have also been highlighted as a key priority by the GBF, notably in Target 5. The sustainable management of a few animal and plant species has proved to be an efficient tool for conservation (*43*, *44*). However, the sustainability of species use remains unknown across most plant diversity. Out of 2800 utilized plant species previously assessed by the International Union for Conservation of Nature (IUCN), more than 1 in 3 is considered to be at risk of global extinction (*43*). More than 1 in 10 plant species with a documented human food use in our study is also considered globally threatened (*45*). Although our findings show that utilized plant diversity remains largely underprotected in the wild, most species (and their genetic diversity) additionally lack representation in ex situ collections such as seed banks and botanical gardens (*46*). Documenting and understanding the diversity and distribution of plant species used by humans are thus crucial to implement conservation strategies and develop plant-based solutions to address global societal challenges such as hunger, diseases, and climate change (*47*–*49*). Our study aims to pave the way for efforts toward reconciling human needs and biodiversity protection for a more sustainable future.

REFERENCES AND NOTES

1. N. Myers, *A Wealth Of Wild Species: Storehouse For Human Welfare* (Routledge, 2019).
2. IPBES, Thematic Assessment Report on the Sustainable Use of Wild Species of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, J.-M. Fromentin *et al.*, Eds. (IPBES secretariat, 2022); https://doi.org/10.5281/zenodo.6448567.
3. S. Díaz *et al.*, *Science* **366**, eaax3100 (2019).
4. WWF, Living Planet Report 2022 - Building a nature-positive society (2022); https://www.wwf.org.uk/our-reports/living-planet-report-2022.
5. R. A. Mittermeier, W. R. Turner, F. W. Larsen, T. M. Brooks, C. Gascon, in *Biodiversity Hotspots*, F. E. Zachos, J. C. Habel, Eds. (Springer, 2011), pp. 3–22.
6. N. Myers, R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, J. Kent, *Nature* **403**, 853–858 (2000).
7. L. Maffi, *Annu. Rev. Anthropol.* **34**, 599–617 (2005).
8. D. Leclère *et al.*, *Nature* **585**, 551–556 (2020).
9. V. Reyes-García *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2217303120 (2023).
10. S. Díaz *et al.*, *Science* **370**, 411–413 (2020).
11. A. Antonelli *et al.*, "State of the World's Plants and Fungi 2020" (Royal Botanic Gardens, Kew, 2020); https://doi.org/10.34885/172.
12. L. Cai *et al.*, *New Phytol.* **237**, 1432–1445 (2023).
13. W. Barthlott, W. Lauer, A. Placke, *Erdkunde* **50**, 317–327 (1996).
14. J. T. Rieb *et al.*, *Bioscience* **67**, 820–833 (2017).
15. S. Pironon *et al.*, *Plants* **9**, 1128 (2020).
16. M. Jung, A. Arnell, X. de Lamo, S. García-Rangel, M. Lewis, *Nat. Ecol. Evol.* **5**, 1499–1509 (2021).
17. R. Chaplin-Kramer *et al.*, *Nat. Ecol. Evol.* **7**, 51–61 (2023).
18. K. M. A. Chan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 1462–1465 (2016).
19. Materials and methods are available as supplementary materials.
20. F. E. M. Cook, *Economic Botany Data Collection Standard* (Royal Botanic Gardens, Kew, 1995).
21. M. Diazgranados *et al.*, World Checklist of Useful Plant Species. Produced by Royal Botanic Gardens, Kew (2020); https://knb.ecoinformatics.org/view/doi:10.5063/F1CV4G34.
22. C. Meyer, P. Weigelt, H. Kreft, *Ecol. Lett.* **19**, 992–1006 (2016).
23. R. Govaerts, E. Nic Lughadha, N. Black, R. Turner, A. Paton, *Sci. Data* **8**, 215 (2021).

24. M. van Kleunen *et al*., *Nat. Commun.* **11**, 3201 (2020).
25. G. Kier *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9322–9327 (2009).
26. P. H. Raven *et al*., *Sci. Adv.* **6**, eabc6228 (2020).
27. N. I. Vavilov, M. I. Vavylov, N. Í. Vavílov, V. F. Dorofeev, *Origin and Geography of Cultivated Plants* (Cambridge Univ. Press, 1992).
28. N. Bystriakova *et al*., *PLOS ONE* **16**, e0256457 (2021).
29. S. J. E. Velazco *et al*., *Biol. Conserv.* **268**, 109525 (2022).
30. L. J. Gorenflo, S. Romaine, R. A. Mittermeier, K. Walker-Painemilla, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8032–8037 (2012).
31. R. Mace, M. Pagel, *Proc. R. Soc. London Ser. B* **261**, 117–121 (1997).
32. N. J. Turner, A. Cuerrier, L. Joseph, *People Nat.* **4**, 627–651 (2022).
33. K. S. Coates, *A Global History of Indigenous Peoples* (Palgrave Macmillan, 2004); https://link.springer.com/10.1057/9780230509078.
34. S. T. Garnett *et al*., *Nat. Sustain.* **1**, 369–374 (2018).
35. R. Cámara-Leret, M. A. Fortuna, J. Bascompte, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 9913–9918 (2019).
36. D. K. Bardsley, N. D. Wiseman, *Glob. Environ. Change* **22**, 713–723 (2012).
37. M. Tengö *et al*., *Curr. Opin. Environ. Sustain.* **26–27**, 17–25 (2017).
38. J. E. M. Watson, N. Dudley, D. B. Segan, M. Hockings, *Nature* **515**, 67–73 (2014).
39. G. G. Gurney *et al*., *Nature* **595**, 646–649 (2021).
40. E. Dinerstein *et al*., *Sci. Adv.* **5**, eaaw2869 (2019).
41. C. Sandbrook, J. A. Fisher, G. Holmes, R. Luque-Lora, A. Keane, *Nat. Sustain.* **2**, 316–323 (2019).
42. E. R. Cetas, M. Yasué, *Conserv. Biol.* **31**, 203–212 (2017).
43. S. M. E. Marsh *et al*., *Conserv. Biol.* **36**, e13844 (2022).
44. L. McRae *et al*., *One Earth* **5**, 422–433 (2022).
45. T. Ulian *et al*., *Plants People Planet* **2**, 421–445 (2020).
46. C. K. Khoury *et al*., *Ecol. Indic.* **98**, 420–429 (2019).
47. A. Di Sacco *et al*., *Glob. Chang. Biol.* **27**, 1328–1348 (2021).
48. A. Cantwell-Jones *et al*., *Nat. Plants* **8**, 225–232 (2022).
49. I. Raskin *et al*., *Trends Biotechnol.* **20**, 522–531 (2002).
50. S. Pironon *et al*., The global distribution of plants used by humans datasets: list of utilised species, occurrence data and model outputs at 10 arc-minutes spatial resolution [Data set], Zenodo (2023); https://doi.org/10.5281/zenodo.8176317.
51. I. Ondo, IanOndo/UsefulPlants: UsefulPlants (v1.0.0), Zenodo, (2023); https://doi.org/10.5281/zenodo.8180352.

ACKNOWLEDGMENTS

We thank nine anonymous reviewers for comments on a previous version of the manuscript. We are grateful to J. Eden for providing illustrations for the figures, N. Ash, J. Borrell, E. Hammond-Hunt, N. Kuhn, B. Kusumoto, R. Smith, M. Soto Gomez, C. Tovar, and P. Wilkin for comments on analyses and text, N. Black for assistance with data from RBG Kew's World Checklist of Vascular Plants, C. Phillips for help with cleaning occurrence records, and A. Barker for help in coordinating the project at RBG Kew. We thank G. Dauby and T. Couvreur for providing unpublished data from the Rainbio database, N. Castañeda-Álvarez and M. Obreza for extracting occurrence records from GENESYS, D. A. L. Canhos and S. De Souza for obtaining records from the SpeciesLink database, and S. Garnett, I. Leiper, O. McDermott-Long, and N. Burgess for help with Indigenous land data. **Funding:** This research was funded by Royal Botanic Gardens, Kew. S.P. and I.O. are supported by the Calleva Foundation. MGU and Wiet Pot Family Foundation have funded part of the data compilation work within the framework of the *Useful Plants Project*. A.A. is supported by the Swedish Research Council (2019-05191) and the Swedish Foundation for Strategic Environmental Research MISTRA (Project BioPath). R.C.-L. is supported by the Swiss National Science Foundation Starting Grant (INDIGENOMICS: TMSGI3_211659).

Author contributions: Original conception: S.P., I.O., M.D., T.U., K.J.W. Study design: S.P., I.O., M.D., T.U., A.A., K.J.W. Data acquisition: S.P., I.O., M.D., R.A., A.C.B., R.C.-L., C.C., Z.D.-C., R.G., S.H., A.J.H., R.L., W.M., M.N., K.P., G.S., R.M.T., T.R.A., T.U., K.J.W. Species name reconciliation: I.O., M.D., R.M.T. Analyses: I.O., S.P. with help from A.C.B. Data visualization: I.O. in close collaboration with S.P. Data interpretation: S.P., I.O., A.A., K.J.W. Writing – original draft: S.P., I.O. Writing – review and editing: S.P., I.O., M.D., R.A., A.C.B., R.C.-L., C.C., Z.D.-C., R.G., S.H., A.J.H., R.L., W.M., M.N., K.P., G.S., R.M.T., T.R.A., T.U., A.A., K.J.W. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Occurrence records, environmental layers, lists of names and uses of the 35,687 modeled utilized plant species, and outputs from species distribution models are available on Zenodo and Github (*50*, *51*) alongside the R package *UsefulPlants*, which gathers codes, functions, tutorials, and documentation allowing one to reproduce analyses of the paper.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adg8028
Materials and Methods
Figs. S1 to S17
Tables S1 to S3
References (*52*–*121*)
MDAR Reproducibility Checklist

Submitted 22 January 2023; accepted 18 December 2023
10.1126/science.adg8028

NATURAL HAZARDS

Surface deformations of the 6 February 2023 earthquake sequence, eastern Türkiye

Jiannan Meng[1,2]†, Timothy Kusky[1,2,3]*†, Walter D. Mooney[4]*, Erdin Bozkurt[3], Mehmet Nuri Bodur[5], Lu Wang[1]

Two powerful earthquakes struck Türkiye on 6 February 2023. The initial rupture was on the Dead Sea fault zone, yet maximum displacements and energy release [moment magnitude (M_w) 7.8] occurred 24 seconds later when rupture transferred to the East Anatolian fault zone (EAFZ). More than 7 hours later, a M_w 4.5 aftershock at the junction of the EAFZ with the east-west striking Çardak-Sürgü fault was followed 86 minutes later by the second large (M_w 7.5) earthquake, suggesting a causal relationship. We provide quantitative ground and aerial documentation of surface offsets and kinematics from the slipped faults, providing important data on surface deformation during large continental strike-slip earthquakes, rupture propagation mechanisms, and how slip may be transferred between complex fault systems. We also provide insight into how slip along linked fault systems accommodates global plate motions.

Large earthquakes in populated regions are responsible for some of the greatest disasters in human history. In epicentral regions on land, obtaining valuable early field observations is often prevented by infrastructure destruction, rescue and recovery operations, hazardous strong aftershocks, inaccessibility, and rapidly eroding fault-scarp features. Despite these challenges, a limited number of studies have been able to correlate point observations of surface ruptures with overall fault displacements (*1*–*4*). The pair of destructive earthquakes of 6 February 2023 in Türkiye (Fig. 1) were densely measured by numerous seismic stations locally and globally, providing a wealth of seismic information. We were able to access the epicentral regions beginning 1 day after these earthquakes; we completed detailed surface mapping of fault ruptures, slip lineations, and offsets along the new fault traces and obtained an uncomm… detailed view of surface deformation … tures before postseismic deformation and erosion.

The first event (Pazarcık earthquake) occurred at 4:17 a.m. (local time) with a moment magnitude (M_w) of 7.8, making it the strongest earthquake in Türkiye in the past century (*5*). The results were catastrophic (figs. S1 to S3), with the devastation of entire cities and officially more than 59,000 deaths (*6*, *7*). The M_w 7.8 fault rupture extended about 350 km (*8*) along the northeast-southwest–oriented East Anatolian fault zone (EAFZ) (Fig. 1). The second (Elbistan) earthquake (M_w 7.5) occurred 9 hours later, at 13:24 (local time), rupturing an estimated 170 km of the approximately east-west–oriented Çardak-Sürgü fault zone (Ç-SF) (*6*, *9*). Previous geologic

[1]State Key Laboratory of Geological Processes and Mineral Resources, Center for Global Tectonics, School of Earth Sciences, China University of Geosciences, Wuhan 430074, China. [2]Badong National Observatory and Research Station for Geohazards, China University of Geosciences, Wuhan, Wuhan 430074, China. [3]Department of Geological Engineering, Middle East Technical University, Ankara 06800, Türkiye. [4]US Geological Survey, Menlo Park, CA 94025, USA. [5]Faculty of Civil Engineering, Hakkari University, Hakkari 30000, Türkiye.
***Corresponding author. Email: tkusky@gmail.com (T.K.); mooney@usgs.gov (W.D.M.)**
†These authors contributed equally to this work.

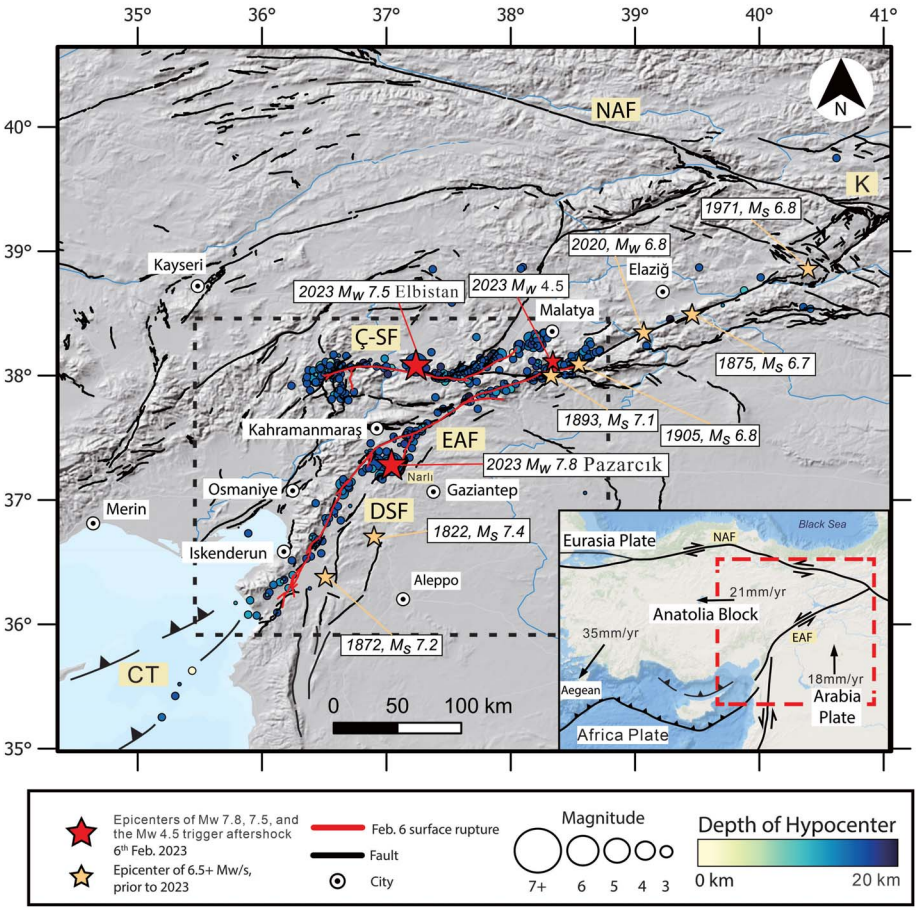


Fig. 1. Location of the M_w 7.8 Pazarcık and M_w 7.5 Elbistan 6 February 2023 earthquakes of the East Anatolian fault system. Aftershocks (blue circles) during the first week after the 6 February 2023 earthquakes and historical earthquakes along the East Anatolian fault are shown. NAF, North Anatolia fault zone; EAF, East Anatolian fault zone; DSF, Dead Sea fault zone; K, Karlıova; Ç-SF, Çardak-Sürgü fault zone; CT, Cyprus Trench. The East Anatolia and North Anatolia faults converge at the point where three plates (Anatolia, Arabia, and Eurasia) meet near Karlıova, known as the Karlıova triple junction (K). The East Anatolian fault, the Dead Sea fault, and the Cyprus Trench meet the Kahramanmaraş triple junction near the first, M_w 7.8 earthquake, at an imprecisely defined location, where the Arabian, African, and Anatolian plates meet. [Sources: faults (*10*); digital topography (*58*); recent and historical seismicity (*5*)]

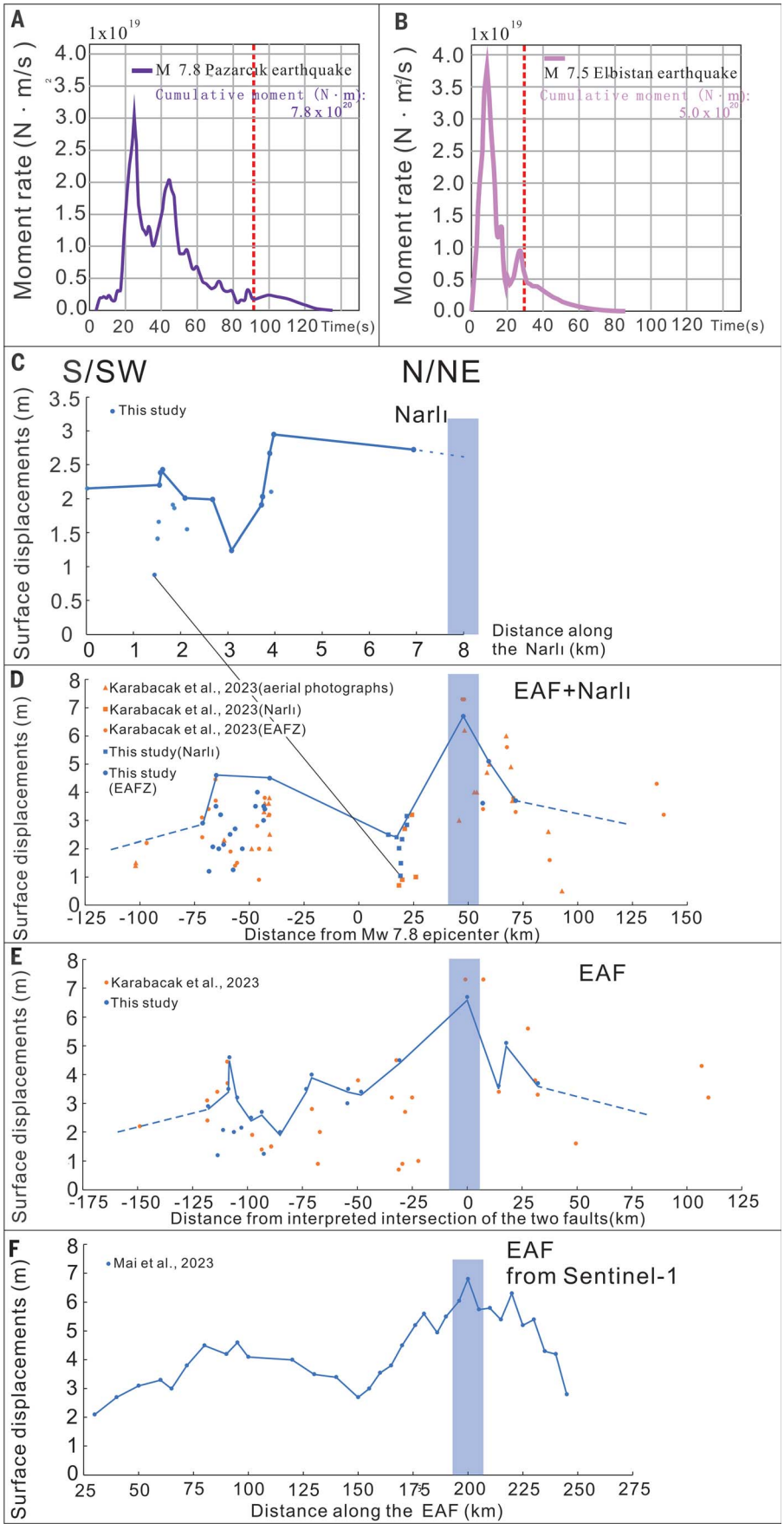


Fig. 2. Energy release spectra compared with displacements measured by field mapping and estimated displacement profiles from field and Sentinel-1 satellite data. (**A** and **B**) Seismic moment release versus time for the (A) M_w 7.8 Pazarcık earthquake and (B) M_w 7.5 Elbistan earthquake of 6 February 2023. In the first M_w 7.8 event, the seismic moment remained very low for the first 16 s and then accelerated rapidly, reaching a peak 24 s after initiation. The rupture became bilateral, and a second peak in seismic moment release occurred at 44 s. The red dashed line indicates the time when the rupture ended, with later moment release being scattered energy. The seismic moment of the second M_w 7.5 event reached its peak in less than 10 s. Data are from the US Geological Survey reported in (*29*). (**C** to **F**) Displacements along the ruptured fault segments, with the thin blue line indicating the maximum measured displacement profile estimated from the field observations reported here. The vertical blue column indicates the intersection region of the Narlı segment of the Dead Sea fault (DSF) and the EAFZ. (C) Measured displacements along the Narlı segment, and (D) the Narlı plus EAFZ, plotted against distance from the epicenter (0 km). The thin line connecting (C) and (D) points to the same blue data point. (E) Displacements along the EAFZ (excluding Narlı) plotted versus distance (+ is to northeast, − is to southwest) from fault intersection zone (table S1) (*59*). (F) Displacements from Sentinel-1 satellite data plotted along the length of the EAFZ (*9*). The maximum displacements occur near the intersection of the Narlı segment of the DSF and the EAFZ, at a distance of 47.5 km from the epicenter. Assuming that the maximum displacements occurred when and where the maximum energy was released at 24 s after the initial rupture, we estimate that the rupture propagated at 2.03 km/s.

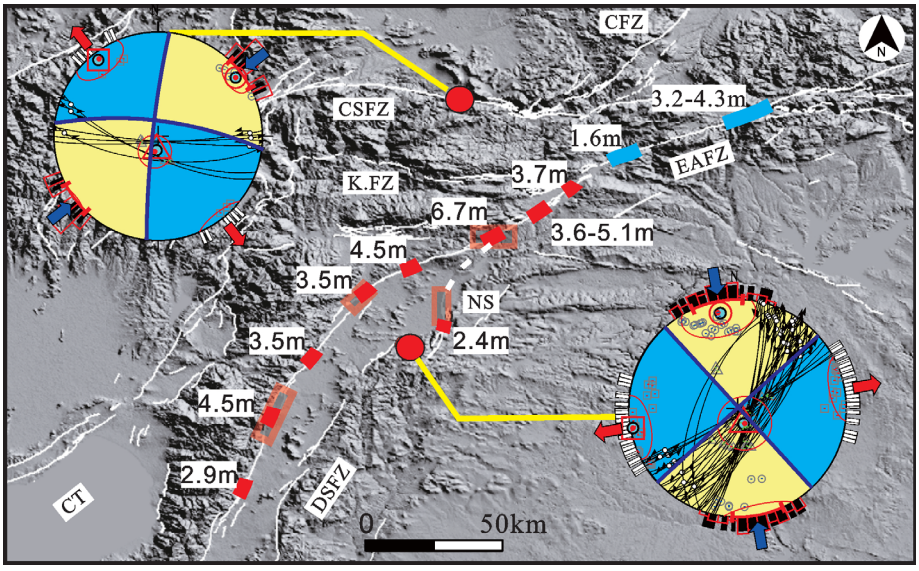


Fig. 3. Field observation of ground offsets and kinematic features around the earthquake area. Shown is the area identified with the black dashed rectangle in Fig. 1. The map superimposes the regional active fault map (*10*) over a digital elevation model (*24*), with measured offsets from our field survey. Offsets labeled with red boxes are from our field observations (table S1), whereas those labeled with blue boxes are compiled from other publications (table S1). The orange rectangles indicate the location of the drone maps [(*11*), data S1 and S2]. The first, M_w 7.8 event occurred on the Narlı segment (NS) of the Dead Sea fault zone (DSFZ), ~47.5 km southwest of the triple junction region between the DSFZ, EAFZ, and Cyprus Trench (CT), where the African, Anatolian, and Arabian plates meet (*7*). (Insets) Plots of the earthquake mechanism solutions on lower-hemisphere Schmidt projections (blue, compressional; yellow, extensional) together with our fault kinematic measurements [blue arrows indicate direction of σ_1, and red arrows indicate σ_3, calculated by using the methods of (*60*)]. The blue dots with circles, triangles, and rectangles indicate the directions of σ_1, σ_2, and σ_3, respectively, and the thin red lines indicate the uncertainties (mean cone angle, 1σ SD). The black and white lines on the outside of the plot indicate the maximum and minimum horizontal stress axes, respectively. The thin great circles are the fault planes, and the arrows on it indicate the slip motions, showing a general consistency but a slight discrepancy related to the surface faults being en echelon splays that likely merge with single faults at depth [(*32*), doc. S2]. NS, Narlı segment of the northern branch of the Dead Sea transform fault; CSFZ, Çardak-Sürgü fault zone; CFZ, Ciglik fault zone; KFZ, Kahramanmaraş fault zone. The white dashed line indicates our interpretation from published active fault maps (*10*). The maximum measured displacement of 6.7 m occurs at the intersection of the Narlı segment of the DSFZ and the EAFZ, 47.5 km from the epicenter. Comparison with Fig. 2 shows that the maximum displacement occurred when the smaller slip on the Narlı segment of the DSFZ transferred to the EAFZ at the moment of peak energy release 24 s after the initial rupture, activating the entire EAFZ. More fault kinematic data on specific locations are provided in fig. S4.

mapping (*10*) indicated that the eastern terminus of the Ç-SF merges into the EAFZ, and the second main earthquake occurred shortly (86 min) after an aftershock (M_w 4.5) of the first event struck the junction of the two faults [(*11*), imagery S1].

These dual (*9*) earthquakes and aftershocks, known as the Kahramanmaraş earthquake sequence (*6*), are associated with various complex strands of the left-lateral Dead Sea fault system, East Anatolian fault system, and their related faults. These faults form the complex and evolving boundaries of the Kahramanmaraş triple junction between the Arabian, African, and Anatolian tectonic plates (*6*, *7*, *12*, *13*). Along these boundaries, the Arabian plate is sliding northward at 18 mm/year relative to the African plate on the Dead Sea transform fault and colliding with the Eurasian plate, whereas the African plate is subducting northward beneath Anatolia in the Cyprus Trench, and the Anatolian plate is extruding westward at 21 mm/year along the East Anatolian and North Anatolian transform faults (Fig. 1) (*14*–*24*). The earthquake sequence can be described simply in a plate tectonic framework as the Arabian plate taking a step to the north, and the Anatolian plate responding with a jump to the west (*7*). The data and observations we report here provide a snapshot of the surface deformation patterns related to the lateral movement of tectonic plates. When linked with recordings and interpretations of the same phenomena through geophysical instrumentation, these data and observations provide valuable multiscale and multidisciplinary perspectives. They also may provide insight into interpreting plate tectonic movements recorded in the geological and seismic record, along with rupture mechanisms along intracontinental plate boundaries and how slip is transferred from one complex plate boundary fault system to another. Further, our observations could ultimately help to protect populations that live in other potentially vulnerable plate boundary regions by informing building codes and preparedness exercises.

Seismological background and earthquake characteristics

The history of strong earthquakes on the central and southern East Anatolian fault is well documented (*25*, *26*). The previous largest event [surface wave magnitude (M_s) 6.8] in the 20th century occurred in 1905 (Fig. 1). More recently, an earthquake of M_w 6.7 occurred on 24 January 2020 on the northeastern portion of the East Anatolian fault system (Fig. 1). Geodetic data indicate that the strain rate on the central portion of the East Anatolian fault is about 10 mm/year (*26*, *27*). Assuming this strain rate, the ~6 m of slip that occurred during the M_w 7.8 earthquake would require 600 years to accumulate. However, strain rates may have varied, and several other earthquakes released some of the strain over this time interval. Paleoseismic data yield an average recurrence interval of 190 years over the past 3800 years (*28*).

Seismic focal mechanisms indicate that the M_w 7.8 and M_w 7.5 events occurred on nearly vertical strike-slip faults (*29*). Aftershocks are largely restricted to the upper 20 km (Fig. 1) of the crust, below which crustal displacements are accommodated by ductile deformation. The initial focal depth for the M_w 7.8 earthquake was 10 km (*6*), and the energy release spectra suggest that it took 10 to 20 s for the rupture to affect the entire upper brittle crust, as the initial rupture propagated north to the EAFZ and south along the Narlı segment of the Dead Sea fault zone (*29*). The depth extent of aftershocks reveals the rupture width (depth) multiplied by the rupture length (along strike) that defines the total surface area of the rupture. The amount of energy released during the earthquake is proportional to the seismic moment, defined as

$$M_o = \mu DA \quad (1)$$

Where μ is the average shear modulus along the fault, D is the average displacement on the rupture plane, and A is the surface area of the rupture plane.

Estimates of the lengths of the fault rupture for the two 6 February 2023 earthquakes from the distribution of aftershocks and satellite-based optical remote sensing images of displacement (Fig. 1) yield fault rupture lengths of ~350 ± 30 km, depth of 20 km and ~170 ± 20 km, depth of 20 km for the M_w 7.8 and M_w 7.5 earthquakes, respectively (*8*, *9*, *29*, *30*). However, the rupture lengths vary depending on which remotely sensed data or geophysical

model is used (*31*). Using our field data reported here and Eq. 1, we estimated an average slip on the EAFZ of 3.2 m, and on the basis of Sentinel-1 data, the average displacement for the M_w 7.5 rupture of the Ç-SF is 4.2 m (*8*, *29*). Using Eq. 1 and a shear modulus of 3.5×10^{10} N/m^2, the seismic moment (M_o) of the M_w 7.8 and M_w 7.5 earthquakes is 7.8×10^{20} and 5.0×10^{20} N-m, respectively.

The rate of seismic moment release during an earthquake helps to determine the duration and amount of energy released. Seismic data reveal the timing of the seismic moment release from the fault rupture plane (*29*). The timing of seismic moment release indicates that the M_w 7.8 event grew over a period of 24 s (Fig. 2A), when the energy release peaked. We interpret this time to be when the rupture had propagated from the Narlı segment of the Dead Sea transform to the East Anatolian transform. From there, the rupture became bilateral along the EAFZ, propagating 125 km to the northeast and 175 km to the southwest (*29*). The M_w 7.5 event peaked much faster (Fig. 2B), with the rupture propagating bilaterally roughly 55 km both east and west (*29*) within the first 9 s. The rapid peak in the seismic moment rate suggests that the Ç-SF was nearly at the rupture stage when it was activated by stress changes from the M_w 7.8 event and associated aftershocks.

Ground and uncrewed aerial vehicle (drone) mapping of fault ruptures

We obtained detailed observations of the surface deformation features, including fault displacement and ground ruptures, and were able to compare our ground-based results with satellite and seismic data as well as with geophysical models for the two large earthquakes. After initial reconnaissance, logistical arrangements, and ground-based observations starting the day after the earthquakes, our quantitative drone-based field mapping started 10 days after the earthquakes [(*32*), doc. S1]. This allowed us to extract kinematic and displacement data from fault planes in surface soils, rocks, and human constructions such as roads, fences, farm field borders, railroads, and buildings. Many of these features were quickly eroded by rains or destroyed by reconstruction after the earthquakes. We investigated most of the length of the rupture along the EAFZ, extending more than 300 km (Fig. 3). In addition to our surface observations, we also report detailed georeferenced maps of surface ruptures of four key places, imaged with the Real Time Kinematic Drone (DJI Phantom 4 RT). These data complement our ground observations of the deformation features recorded in the days immediately after the earthquakes (figs. S4 to S6 and tables S1 and S2) [(*11*), drone imagery S2 to S7 and data S1 and S2].

The structural and tectonic setting is complex near the epicentral regions. Three fault systems are involved (*10*), including the main (transform) EAFZ, the Narlı fault segment of the Dead Sea transform fault (*33*, *34*), and the east-west Ç-SF (Fig. 3), which merges into the EAFZ along a fault branch known as the Sürgü fault (*10*). The M_w 7.8 earthquake initiated on the Narlı fault (*6*). The main Ç-SF slipped 9 hours after the first event, resulting in the second, M_w 7.5 event of the sequence (*6*), but the Sürgü fault became inactive, and aftershocks next migrated northeast along a new rupture (Ciglik fault) parallel with the main EAFZ (Figs. 1 and 3) [(*11*), imagery S1].

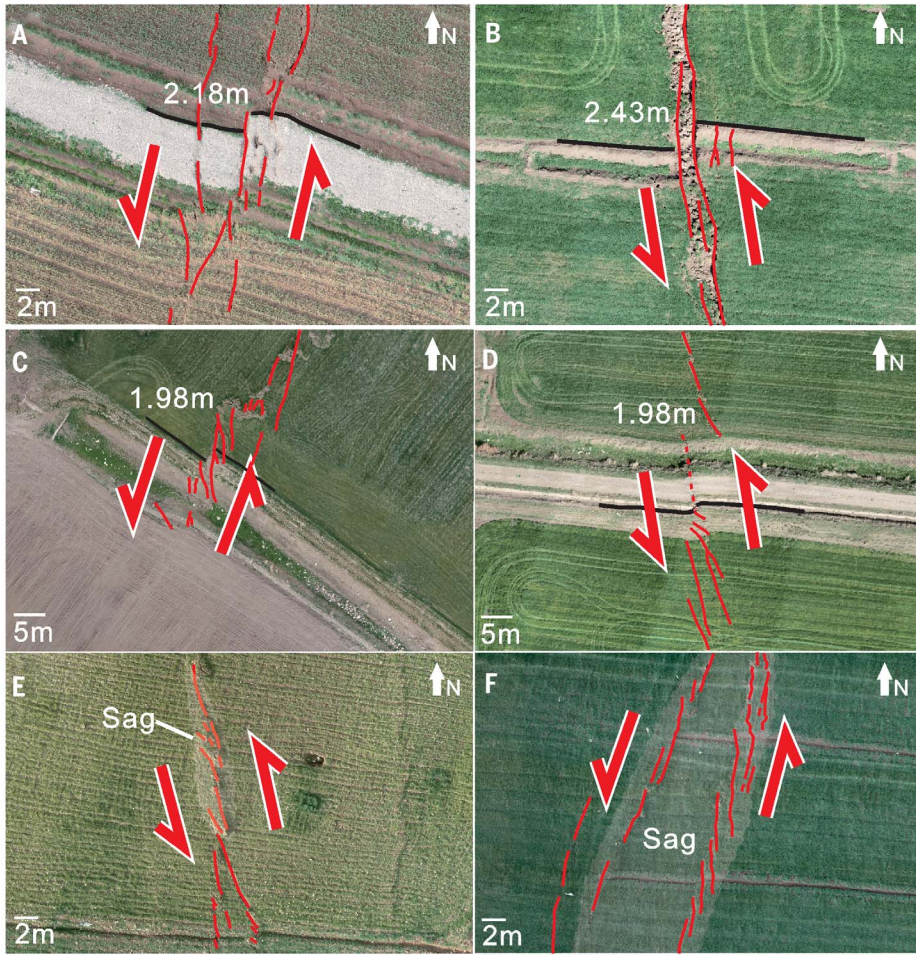


Fig. 4. Horizontal ground offsets and ruptures of the Narlı segment of the Dead Sea transform fault. Images are from aerial drone mapping. Exact locations and other parameters of all images are available in (*11*), data S1. The locations of the drone mapping areas are shown in Fig. 3. (**A**) A road with a 2.18-m offset. Black lines are used as a reference. Thin red lines indicate the deformation pattern, and the arrows indicate the motion direction. (**B**) The field edge with a 2.43-m offset. Black lines are used as a reference. Thin red lines indicate the deformation pattern, and the arrows indicate the motion direction. (**C**) A road with a 1.98-m offset. Black lines are used as a reference. The field in the southwest part of the image is already plowed over, removing all traces of the fault trace. Thin red lines indicate the deformation pattern, and the arrow indicates the motion direction. Synthetic en echelon faults formed at a low angle to the main ruptures, antithetic faults are at high angles to the main fault, and through-going connecting faults link the different segments. These typically are best developed in a several-meter-wide damage zone, marking the main fault trace. There are some ground distortions and rotations between the ground ruptures. (**D**) The road with a 1.98-m offset. Black lines are used as a reference. Thin red lines indicate the deformation pattern, and the arrow indicates the motion direction. The north-northeast "lines" are artifacts of the image merging. The location of (D) is close to the edge of the fault segment; the deformation is not as strong as in (A) to (C). (**E** and **F**) Distributed en echelon ruptures and also small sag basins related to the strike-slip fault. Scale bars provided at bottom left, (A) to (F).

Measured ground offsets

Ground offsets (horizontal) along all three faults are sinistral, varying from <2 to 6.7 m in different locations (Fig. 3 and fig. S4) [(*11*), data S1 and S2]. According to our field observations, the closest surface rupture to the epicenter is located around 5 km to the east, along the Narlı segment of the Dead Sea (transform)

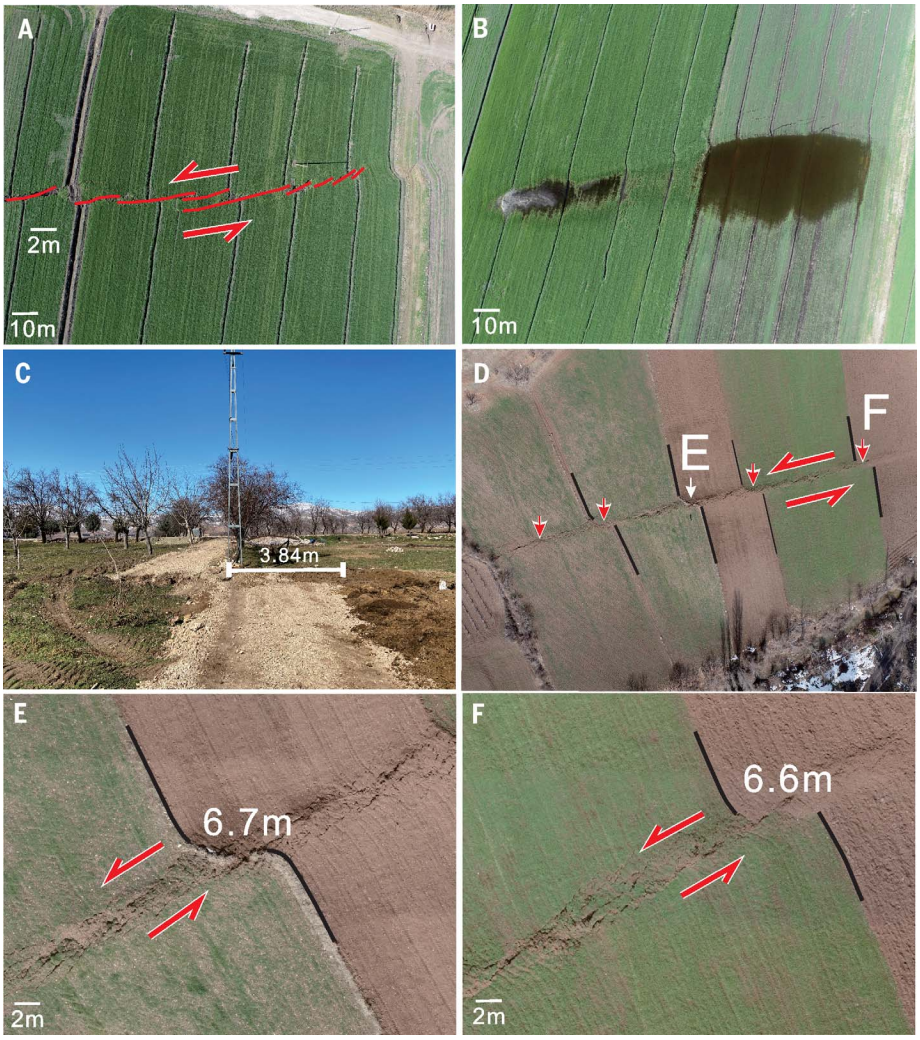


Fig. 5. Surface expression of maximum displacement area on the EAFZ, located at the triple junction where the Dead Sea transform intersects the East Anatolian transform. (**A**) A farm field is offset by synthetic en echelon pattern faults. (**B**) A newly formed sag basin (indicated by recent pond) formed through subsidence along an extensional jog in the fault system. (**C**) A road is offset by 3.84 m. All observations indicate sinistral motion. (**D**) The fault cuts through the field and exhibits the largest ground offset that we observed. (**E** and **F**) Close-up views showing that the displacement is accommodated across a 3- to 4-m-wide zone of distributed deformation, or damage zone (*61*). Locations of images (E) and (F) are labeled in (D).

fault zone, extending around 7.2 km from north to south (Fig. 3). The geometry of this fault strand is generally north-south striking (ranging between 346° and 030°), on nearly vertically dipping (78° to 89°) planes, and the surface lateral offsets are around 2 to 2.5 m. (Figs. 3 and 4) [(*11*), data S1 and S2], with some unusual oblique reverse-slip motions at the tips of the rupture areas and along isolated pop-up pressure ridge structures, as described along other strike-slip fault systems (*27*, *28*).

The Narlı fault strand of the Dead Sea transform fault merges with the main EAFZ about 47.5 km northeast of the epicentral region that marks the triple junction between the Arabian, African, and Anatolian plates (Figs. 1 and 3). The rupture propagated from the Narlı segment to the EAFZ during the initial 10 to 25 s of the M_w 7.8 event (Fig. 2), depending on which data are used for the seismic inversion (*8*, *29*). The EAFZ generally has larger horizontal offsets than those of the Narlı fault segment (Fig. 2, C to F), where the first M_w 7.8 earthquake initiated, and the energy release spectra (Fig. 2A) were noticeably less on the Narlı segment. The strike of the EAFZ in our field observations ranges from 347° to 110° along 300 km, and the horizontal ground offsets vary from 3.4 to 6.7 m (Figs. 3 and 5, D, E, and F) [(*11*), imagery S2 to S5 and data S2].

The area of maximum offset from this earthquake event is located 47.5 km to the northeast of the epicenter (37.5830042°N, 37.316717°E), along the Pazarcık segment of the EAFZ at the triple junction between the three plates (Fig. 3). The maximum displacement is in a farm field in which individual previously linear roads and crop rows are offset by the 051° striking fault from 6.5 to 6.7 m along a 200-m-long strand of the fault. The images we documented show the undisturbed offsets soon after the earthquakes, but because they are in farm soils, they were subsequently obscured by plowing of the soil for planting the next harvest of crops. Neither clear fault planes nor clear vertical offsets were observable at this location (Fig. 5D) [(*11*), data S2], with the displacement accommodated across a 3- to 4-m-wide deformation zone (Fig. 5, E and F) [(*11*), data S2].

Our observations indicate that the greatest ground deformation and area of maximum slip is not located in the epicenter region. Instead, as the fault rupture propagated over 90 s, it took 24 s for the main rupture with the highest energy release (Fig. 2A) to occur on the Pazarcık segment of the EAFZ (Fig. 2, C to F). The rupture can be interpreted to have propagated at 2.03 km/s from where it initiated on the Narlı fault to the location where greatest energy was released at the intersection of the Narlı and East Anatolian faults, also yielding the largest displacements. The maximum displacement and energy release occurred almost precisely when the rupture propagated to the main EAFZ. Our observations of the distance that the rupture traveled to the maximum offset area when the maximum energy was released are slightly lower than that from geophysical numerical rupture models that use joint kinematic inversion of high-rate global navigation satellite system (HR-GNSS) and strong motion data, which suggest supershear rupture propagation velocities of up to 3.2 km/s for the M_w 7.8 event (*8*), but are consistent with a slow moment release on the Narlı fault during the first 12 s (*29*). Twenty min after this event, new earthquakes started propagating northeast and southwest along the EAFZ [(*11*), imagery S1].

Seven hours and 40 min after the main M_w 7.8 event, a M_w 4.5 aftershock struck the junction of the east-west Ç-SF with the EAFZ on the Sürgü fault (Fig. 1) (*10*) and activated seismicity on the entire Ç-SF system [(*11*), imagery S6]. One hour 20 min after the M_w 4.5 earthquake, the M_w 7.5 Elbistan event occurred in the center of the Ç-SF, possibly induced by stress changes caused by the earlier earthquakes (*2*). During our observations of the Ç-SF, the mountains were covered with thick snow, followed by springtime torrential rains and disastrous flooding, partially erasing many of the delicate fault surface features and hampering further observations of ground offsets and fault planes. Thus, no convincing offset was measurable on the Ç-SF during our fieldwork, although the rupture

surfaces and slip lineations were visible and measured (fig. S5, A to F). However, reports exist indicating that the Ç-SF has around 7 to 8 m of local horizontal offsets, but these have not been confirmed with quantified field observations (*30*, *35*).

Vertical offsets on the faults, ranging from 0.5 to 1.7 m on different fault segments, are not as large as the major horizontal displacements. Vertical offsets occur either as ground ruptures or as oblique thrust-strike slip offsets on pressure ridges along strike-slip faults (figs. S5, A, B, and E) or as oblique-normal displacements along extensional jogs and sag basins (Fig. 5B), which is consistent with the remote sensing observations derived from ESA Sentinel-1 radar image pairs acquired on 28 January 2023 and 10 February 2023 (*29*, *30*).

Other surface deformation features

En echelon fracture patterns are the most common surface deformation style along the fault zone (Figs. 4 and 5A) [(*11*), data S1 and S2]. The surface deformation, especially in clays and unconsolidated alluvium, is characterized in many places by a series of discontinuous synthetic en echelon faults that theoretically should lie at an acute angle to the main fault at depth, and the en echelon faults on the surface may merge with a major fault plane near or above the brittle-ductile transition (*5*–*7*, *36*–*39*). This pattern is consistent with the nearly parallel but slightly oblique relationships between the measured surface faults and the recorded seismic moments, indicating sinistral motions (fig. S4, A to I). Ground deformation in this pattern gradually decreases at the end of the individual deformation belts and is picked up by the next synthetic en echelon fault.

The surface ruptures and newly formed faults do not always follow topographic changes or preexisting weak surfaces. Some of the ruptures cut through the centers of basins, and others cut across the topographic highs directly, instead of following known fault traces, which is inconsistent with expectations that the new ruptures follow known active faults or topographic changes. Old fault-related sag basins and pressure ridges were activated by these events, and there are also some newly formed sag and small pull-apart basins and pressure ridges (*39*) that emerged along the faults, often bordered by splays (Figs. 4, E and F, and 5B).

Progressive development of through-going fault surfaces with increasing displacement in unconsolidated soils

The mechanisms by which strike-slip fractures propagate and coalescence into major through-going faults in rocks and other materials have been examined experimentally, theoretically, and with a limited number of field observations (*36*, *40*–*43*). Many of the field examples show that large strike-slip faults often form by the linking of many en echelon faults or fractures [for example, (*9*, *39*, *43*, *44*–*49*)], a hypothesis that our observations confirm. Previous studies have generally suggested that the propagation and growth of fractures into faults is variable and depends on the material properties of the affected area, the stress and strain conditions, and the presence or absence of preexisting fractures. However, relatively few studies have examined fault propagation in soils or unconsolidated sediments, with some exceptions noted above (*1*–*4*, *49*–*52*) and some theoretical studies [for example, (*50*, *53*)]. The detailed imagery and maps presented here [(*11*), data S1 to S3] offer an unprecedented view of the propagation and development of surface fault features in soils and bedrock associated with two major earthquakes.

Snapshots of our drone images and maps of the EAFZ ruptures in farm field soils [(*11*), data

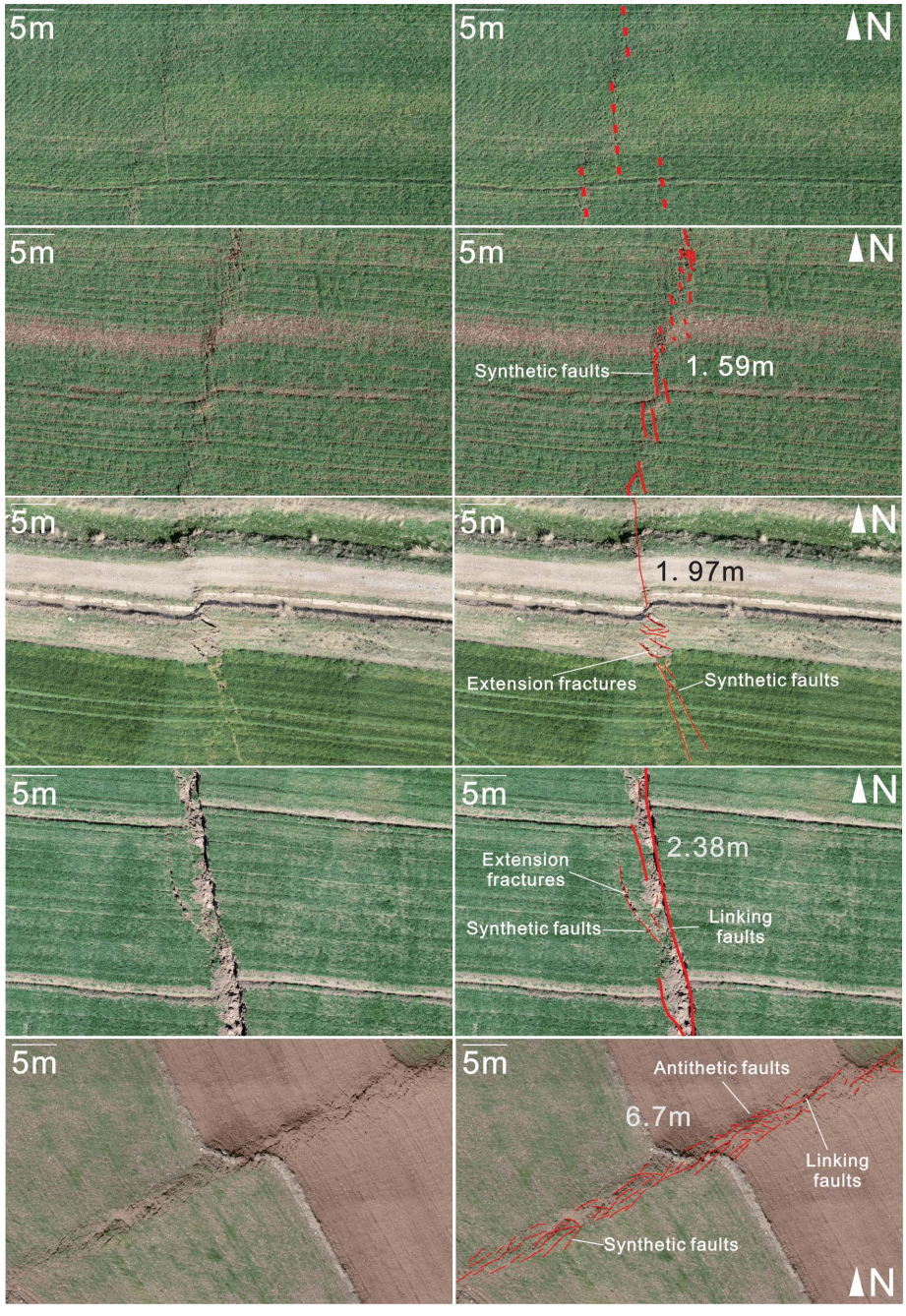


Fig. 6. Progressive development of through-going fault planes from initial series of en echelon faults along the EAFZ. (Left) Uninterpreted images. (Right) The same images including our interpretations of the new fault planes, with increasing displacements from top to bottom. Locations of images are given in (*11*), data S1.

S1] are shown in Fig. 6 in a progression from low displacements at top to the 6.7-m maximum displacements at bottom [(*11*), data S2]. There is a progression from small, isolated, nearly parallel fractures at the smallest displacements to synthetic en echelon faults at about 1.5-m displacements, with through-going linking faults parallel to the main fault beginning to cut both synthetic and antithetic en echelon faults at displacements greater than 2.3 m. The largest offset areas are characterized by through-going faults and continued development of en echelon faults, defining a wider damage zone for the main fault. The progressive structural development of the fault features in soils is very similar to that in rocks with different material properties (*48*–*53*). This observation suggests that the material properties may be less important than other factors (stress and strain) in determining the structural development of large-scale, high-displacement faults in unconsolidated sediments.

Discussion and implications

Our mapping of the surface deformation features of the earthquakes of 6 February 2023 provides new information on the distribution of slip along two major continental transform faults and provides insight into rupture propagation and slip transfer on complex fault systems. Our rapid deployment to the field made it possible to obtain high-quality uncrewed aerial vehicle (drone) and ground observations of ephemeral surface deformation features. These data, together with the corresponding seismic data, illustrate the process through which tectonic plates interact within two large-scale continental transform systems. The dual M_w 7.8 and M_w 7.5 earthquakes demonstrate how the sudden motion of one plate (Arabia, moving to the north) can lead to the response to that motion by another plate (Anatolia, moving to the west). Thus, the observed surface displacements correlated with the seismic moment release provide insight into how continental plate tectonic processes occur with the coupled rupturing of two plate boundaries, and how the surface expression of these plate movements is formed in the geological and geomorphic record.

Our observations of surface displacements (Fig. 2, C to F) add to the existing knowledge base of the patterns of surface rupture for major earthquakes. For example, a broadly similar pattern of surface displacement was documented from the 2002 Denali (Alaska) earthquake, with branching of fault strands on the Denali and Totschunda faults and maximum fault displacement far from the epicenter (*3*). However, in general maximum surface offsets far from the epicenter, as we have documented, are not common (*4*). We found that the largest ground offset is not near the epicenter within the Narlı segment of the Dead Sea transform fault. Rather, the largest offset is 47.5 km to the northeast at the intersection of the Dead Sea and East Anatolian fault systems. This situation is unusual because observations from other earthquake events typically show that fault junctions foster rupture arrest, not increases in slip (*54*–*56*). We suggest that the observed spike in the relative seismic moment rate (energy release in newton-meters per second), which occurred 24 s after the origin time, relates to the time when the Narlı rupture intersected with the EAFZ, which then "exploded" with 6.7 m of slip at the intersection (at the triple junction plate boundary). The rupture then became bilateral, activating the entire central and southern EAFZ transform plate boundary. This burst of released energy, with maximum fault slip, may be due to this portion of the EAFZ having accumulated several hundred years of stress build-up (*25*–*28*). On the basis of seismic waveform modeling, the rupture propagation on the EAFZ was at supershear velocities (>~3 km/s), which is perhaps 1.55 times the shear wave velocity, creating a "Mach cone" that is analogous to a sonic boom in the air (*8*, *57*). However, our observations suggest that the initial rupture propagation on the Dead Sea transform (Narlı fault) was at a subshear velocity and accelerated only after transfer to the EAFZ, where the rupture became bilateral (*29*, *57*). Our data show that the maximum slip on the EAFZ (6.7 m) was very near the intersection of the two fault systems, about 5 to 10 km southwest of the maximum slip location predicted by the US Geological Survey finite fault model [(*32*), doc. S3] (*29*).

Our data, together with the analysis of the seismic data and stress modeling reported elsewhere, relate to the important question of how the slip migrated from the EAFZ to the Ç-SF. More than 7 hours after the rupture propagated north along the EAFZ, a M_w 4.5 earthquake occurred at the junction with the easternmost portion of the Ç-SF. The east-west Çardak-Sürgü system then became seismically active, followed some 86 min later by the M_w 7.5 earthquake [(*11*), imagery S1]. The timing suggests that the M_w 7.5 event on the Ç-SF may have been triggered by a stress change or other fault parameters, such as a pore pressure increase, that were altered because of the M_w 7.8 event. The M_w 7.5 earthquake produced an abundant aftershock sequence on the Ç-SF. However, the aftershocks show that slip did not reconnect with the EAFZ but unexpectedly turned northeast along a new fault rupture plane on the Ciglik fault (*10*), which is parallel to the EAFZ (Figs. 1B and *3*).

The large distance (47.5 km) between the epicenter location and the area of maximum fault displacement calls into question the interpretation of historical earthquake catalogs that probably record the location of the most damaged population center rather than the epicenter where historical earthquakes were initiated. These earthquake catalog locations can be interpreted as the locations of maximum energy release.

Our data address the question of where the largest surface deformation may be expected for a large earthquake. Our field investigation shows that surface fault deformation is typically distributed on a series of en echelon faults within a several-meter-wide damage zone, whereas at depth, the maximum displacement may be on a discrete fault plane where the individual en echelon faults merge (*1*–*3*). The shape of the displacement field may be complex and intersect the surface obliquely from the hypocentral location for a dipping fault.

The M_w 7.8 and M_w 7.5 earthquakes of February 2023 provide a wealth of information on fault rupture propagation, how slip is transferred between fault systems, how this behavior is related to large-scale linked tectonic plate motions, and how seismic and other geophysical data correlate with observable ground deformation features. Our ground-truth observations provide quantitative measurements of surface offsets and associated deformation. Continued comparison of ground deformation, dynamic rupture models, Coulomb stress models, finite fault models, and other geophysical measurements may improve understanding of the physics of large earthquakes and how individual fault systems accommodate global plate motions, and this knowledge may contribute to protecting society from the consequences of major earthquakes.

REFERENCES AND NOTES

1. G. P. Biasi, R. J. Weldon, *Bull. Seismol. Soc. Am.* **96**, 1612–1623 (2006).
2. G. P. Biasi, S. G. Wesnousky, *Bull. Seismol. Soc. Am.* **106**, 1110–1124 (2016).
3. P. J. Haeussler *et al*., *Earthq. Spectra* **20**, 565–578 (2004).
4. M. A. Hemphill-Haley, R. J. WeldonII, *Bull. Seismol. Soc. Am.* **89**, 1264–1279 (1999).
5. US Geological Survey, Earthquake Lists, Maps, and Statistic (accessed March 18, 2023); https://www.usgs.gov/natural-hazards/earthquake-hazards/lists-maps-and-statistics.
6. R. Schmitt, M. Herman, W. Barnhart, K. Furlong, H. Benz, The 2023 Kahramanmaraş, Turkey, Earthquake Sequence, (USGS Geologic Hazards Center, March 28 2023); https://storymaps.arcgis.com/stories/355bfc8b3c5941e683d4f258e8fb2dfa.
7. T. M. Kusky, E. Bozkurt, J. Meng, L. Wang, *J. Earth Sci.* **34**, 291–296 (2023).
8. D. Melgar *et al*., *Seismica* **2**, 3 (2023).
9. P. M. Mai *et al*., *The Seismic Record* **3**, 105–115 (2023).
10. Ö. Emre *et al*., *Bull. Earthquake Eng.* **16**, 3229–3275 (2018).
11. J. Meng *et al*., Other supplementary materials for Surface deformations of the February 6, 2023 earthquake sequence, eastern Türkiye. figshare (2023); https://doi.org/10.6084/m9.figshare.24872043.
12. D. Whitney *et al*., *Geology* **51**, 673–677 (2023).
13. A. C. Şengör, C. Zabcı, B. A. Natal'in, *Continental Transform Faults in Encyclopedia of Geology* (Elsevier, 2021), pp. 913–929.
14. A. C. Şengör, N. Canitez, *Geodyn. Ser.* **7**, 205–216 (1982).
15. J. Dewey, A. C. L. Şengör, *Geol. Soc. Am. Bull.* **90**, 84–92 (1979).
16. D. McKenzie, *Earth Planet. Sci. Lett.* **29**, 189–193 (1976).
17. E. Bozkurt, *Geodin. Acta* **14**, 3–30 (2001).
18. X. Le Pichon, C. Kreemer, *Annu. Rev. Earth Planet. Sci.* **38**, 323–351 (2010).
19. A. Şengör *et al*., *Annu. Rev. Earth Planet. Sci.* **33**, 37–112 (2005).
20. F. A. Capitanio, *Earth Planet. Sci. Lett.* **448**, 122–132 (2016).
21. C. Faccenna *et al*., *Rev. Geophys.* **52**, 283–332 (2014).
22. L. Jolivet *et al*., *Tectonophysics* **597-598**, 1–33 (2013).

23. A. Şengör, M. Yazıcı, *Mediterranean Geoscience Reviews* **2**, 327–339 (2020).
24. J. Meng *et al*., *Sci. Rep.* **11**, 6486 (2021).
25. N. N. Ambraseys, *Geophys. J. Int.* **96**, 311–331 (1989).
26. B. Aktug *et al*., *J. Geodyn.* **94-95**, 1–12 (2016).
27. R. Reilinger *et al*., *J. Geophys. Res. Solid Earth* **111**, B5 (2006).
28. A. Hubert-Ferrari *et al*., *Earth Planet. Sci. Lett.* **538**, 116152 (2020).
29. D. E. Goldberg *et al*., *The Seismic Record* **3**, 156–167 (2023).
30. N. G. Reitman *et al*., (2023). Fault rupture mapping of the 6 February 2023 Kahramanmaraş, Türkiye, earthquake sequence from satellite data. US Geological Survey Data Release (2023); https://doi.org/10.5066/P985I7U2.
31. S. Barbot *et al*., *Seismica* **2**, 2 (2023).
32. Materials and methods are available as supplementary materials.
33. D. Perinçek, I. Çemen, *Tectonophysics* **172**, 331–340 (1990).
34. Y.-S. Kim, J. R. Andrews, D. J. Sanderson, *Geosci. J.* **4**, 53–72 (2000).
35. L. Dal Zilio, J.-P. Ampuero, *Commun. Earth Environ.* **4**, 71 (2023).
36. L. N. Y. Wong, Y. Zhang, *J. Geophys. Res. Solid Earth* **124**, 10916–10939 (2019).
37. A. Sylvester, *Geol. Soc. Am. Bull.* **100**, 1666–1703 (1988).
38. Y. Cheng, L. N. Y. Wong, C. Zou, *Eng. Geol.* **195**, 312–326 (2015).
39. P. Mann, M. R. Hempton, D. C. Bradley, K. Burke, *J. Geol.* **91**, 529–554 (1983).
40. M. Doblas, *Tectonophysics* **295**, 187–197 (1998).
41. A. G. McGrath, I. Davison, *J. Struct. Geol.* **17**, 1011–1024 (1995).
42. P. A. Cowie, Z. K. Shipton, *J. Struct. Geol.* **20**, 983–997 (1998).
43. Y.-S. Kim, D. C. P. Peacock, D. J. Sanderson, *J. Struct. Geol.* **25**, 793–812 (2003).
44. D. A. Rodgers, in *Sedimentation in Oblique-Slip Mobile Zones*, P. F. Ballance, H. G. Reading, Eds. (Wiley, 1980), pp. 27–41.
45. T. Kusky, W. Kidd, D. Bradley, *J. Geodyn.* **7**, 105–133 (1987).
46. T. Kusky, C. Robinson, F. El-Baz, *J. Geol. Soc. London* **162**, 871–888 (2005).
47. B. Tikoff, C. Teyssier, *Geology* **20**, 927–930 (1992).
48. D. Faulkner, A. Lewis, E. Rutter, *Tectonophysics* **367**, 235–251 (2003).
49. E. J. Willemse, D. C. Peacock, A. Aydin, *J. Struct. Geol.* **19**, 1461–1477 (1997).
50. D. Loukidis, G. D. Bouckovalas, A. G. Papadimitriou, *Soil. Dyn. Earthquake Eng.* **29**, 1389–1404 (2009).
51. T. M. Kusky *et al*., *J. Earth Sci.* **21**, 910–922 (2010).
52. B. A. Brooks *et al*., *Sci. Adv.* **3**, e1700525 (2017).
53. J. Shi, L. Guan, D. Zhuang, X. Chen, D. Ling, *J. Rock Mech. Geotech. Eng.* **15**, 2973–2993 (2023).
54. Y. Cheng, Y. Ben-Zion, *Geophys. Res. Lett.* **47**, e2020GL089650 (2020).
55. R. D. Koehler *et al*., *Seismol. Res. Lett.* **92** (2A), 823–839 (2021).
56. J. Liu-Zeng *et al*., *Earthquake Research Advances* **2**, 100140 (2022).
57. B. Delouis, M. van den Ende, J. P. Ampuero, *Bull. Seismol. Soc. Am.* **113**, 1–15 (2023).
58. Advanced Spaceborne Thermal Emission and Reflection Radiometer, (ASTER) Global Digital Elevation Model, version 3 (Jet Propulsion Laboratory, 2019); https://asterweb.jpl.nasa.gov/gdem.asp.
59. V. Karabacak *et al*., *J. Geol. Soc. London* **180**, jgs2023–jgs2023-020 (2023).
60. D. Delvaux, B. Sperner, in *New Insights into Structural Interpretation and Modelling*, D. A. Nieuwland, Ed., Geographical Society Special Publications vol. 212 (Geographical Society of London, 2003), pp. 75–100.
61. E. Hussain, S. Kalaycıoğlu, C. W. D. Milliner, Z. Çakir, *Nat. Rev. Earth Environ.* **4**, 287–289 (2023).

ACKNOWLEDGMENTS

We are grateful to the people of Türkiye for their assistance during the field work. We thank D. E. Goldberg, J. Nevitt, B. Sherrod, E. Van Dyke, L. Hanson, C. Barrera-Lopez, R. Chidlow, J. Hardebeck, S. Hecker, and three anonymous readers for their review comments. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the China University of Geosciences, Middle East Technical University, Hakkari University, or the US government. **Funding:** This work was supported by the National Natural Science Foundation of China, 41888101 (T.K. and L.W.) and 91755213 (T.K., L.W., and E.B.); the State Key Lab of Geological Processes and Mineral Resources, China University of Geosciences (Wuhan), MSFGPMR2022-7 (L.W.); and the US Geological Survey Earthquake Hazards Program (W.D.M.) **Author contributions:** Conceptualization: T.K. and W.D.M. Methodology: J.M. and T.K. Investigation: J.M., T.K., W.D.M., E.B., M.N.B., and L.W. Visualization: J.M., W.D.M., and T.K. Funding acquisition: T.K., L.W., and E.B. Project administration: T.K. and L.W. Supervision: T.K., W.D.M., and E.B. Writing – original draft: J.M., T.K., and W.D.M. Writing – review and editing: J.M., T.K., W.D.M., E.B., and L.W. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text, the supplementary materials, and at the locations listed at (*11*), the latter directing to imagery S1 to S7 and data S1 to S3.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adj3770
Materials and Methods
Supplementary Text
Figs. S1 to S7
Tables S1 to S3
References (*62*–*69*)

Submitted 23 June 2023; resubmitted 18 September 2023
Accepted 8 December 2023
10.1126/science.adj3770

NATURAL HAZARDS

Supershear triggering and cascading fault ruptures of the 2023 Kahramanmaraş, Türkiye, earthquake doublet

Chunmei Ren[1]†, Zexin Wang[1]†, Tuncay Taymaz[2], Nan Hu[3], Heng Luo[1], Zeyan Zhao[1], Han Yue[1]*, Xiaodong Song[1,4]*, Zhengkang Shen[5,1], Haoyu Xu[1], Jianghui Geng[6], Wei Zhang[3], Teng Wang[1], Zengxi Ge[1], T. Serkan Irmak[7], Ceyhun Erman[2], Yijian Zhou[8], Zhen Li[1], Hang Xu[1], Bonan Cao[1], Hongyang Ding[1]

On 6 February 2023, two large earthquakes (moment magnitude 7.8 and 7.6) shocked a vast area of southeastern Türkiye and northern Syria, leading to heavy casualties and economic loss. To investigate the rupture process over multiple fault segments, we performed a comprehensive analysis of local seismic and geodetic data and determined supershear ruptures on the initial branch and the Pazarcık and Erkenek segments and subshear ruptures on the Amanos segment of event 1. The bilateral rupture of event 2 also presents distinct sub- and supershear velocities. The dynamic stress of the branch fault rupture triggered the Pazarcık segment initial rupture at a point 9 kilometers west of the junction of these two faults, boosting the supershear rupture of the Pazarcık segment of the main fault. The geometry and prestress level of multiple segments controlled the rupture behaviors and influenced the ground shaking intensity.

On 6 February 2023, a devastating earthquake of moment magnitude (M_w) 7.8 ("event 1") near the city of Kahramanmaraş shocked southeastern Türkiye and northern Syria. The event was initiated at 04:17:34 a.m. local time [or 01:17:34 UTC, per the United States Geological Survey (USGS)], and it lasted for ~90 s and produced ~380-km-long surface ruptures. Event 1 was followed by another major event of M_w 7.6 ("event 2") about 9 hours later (at 10:24:48 UTC, per the USGS), which produced ~200-km surface ruptures along a separate fault located ~100 km north of the event 1 epicenter (Fig. 1A). This 2023 Kahramanmaraş earthquake doublet resulted in catastrophic destruction with substantial devastation to a populated area covering several major cities in southeastern Türkiye and northwestern Syria (e.g., Kahramanmaraş, Adıyaman, Şanlıurfa, Antakya, Gaziantep, Malatya, İskenderun, and Adana, among others) and caused more than 50,000 deaths and 100,000 injuries across both countries. The total economic loss is estimated to be >100 billion USD. Event 1 is one of the largest continental strike-slip events recorded by modern seismological instruments, comparable to the 2001 M_w 7.8 Kokoxili earthquake (*1*), the 2002 M_w 7.8 Denali earthquake (*2*), and the 1906 M_w 7.7–7.8 San Francisco earthquake (*3*). The doublet is the largest continental earthquake doublet ever recorded by modern seismological instruments (*4*).

The 2023 Kahramanmaraş earthquake doublet (Fig. 1) occurred in the East Anatolian Fault Zone (EAFZ), which is a major fault zone created by the oblique collision between the Anatolian and Arabian plates. The regional tectonic setting is controlled by the interaction between the Anatolian plate and the surrounding Eurasian, African, and Arabian plates (*5*–*7*). Aside from the continental collision, the subduction of the African plate also leads to the opening of the Aegean Sea (*6*, *8*), and the net force of collision and dilatation causes a westward movement of the Anatolian plate, resulting in the formation of two major fault systems along its northern and eastern boundaries: the North Anatolian Fault Zone (NAFZ) and the EAFZ (*7*, *9*). The NAFZ slips at a high rate of between 18 and 22 mm/year (*10*) and has produced high seismic activity along the fault during the 20th century (*11*, *12*) (Fig. 1A). The EAFZ slips at a relatively lower rate: ~10 mm/year on its central to northeastern (from Nurdağı-Pazarcık to Pütürge) segments and ~4.5 mm/year on its southwestern (Amanos) segment (Fig. 1). The EAFZ was ruptured by several historical events before the 20th century, including the 1513 M 7.4, 1822 M 7.5, 1866 M 7.2, 1874 M 7.1, and 1893 M 7.1 events (*13*) (Fig. 1B), but it was relatively inactive in the 20th century and the first two decades of the 21st century, until an M_w 6.8 event ruptured the Pütürge segment of the fault in 2020 (*14*).

Segmented rupture appears to be the m… characteristic of both fault zones, which i… cates that junctions of segments—for example, stepover, bending, and intersections—serve as barriers to stop dynamic ruptures (*13*). However, a large earthquake may break through these barriers and sequentially rupture several segments, as occurred during the 1668 Anatolian earthquake (*15*) (Fig. 1A). Event 1 of the 2023 doublet is also such an event; it ruptured three segments (Amanos, Pazarcık, and Erkenek) hosting three historical events (1822, 1513, and 1893 events). This behavior of multiple segment ruptures poses a challenge for forecasting the earthquake potential of the region (*16*). Meanwhile, supershear rupture is commonly found in large strike-slip events (*17*), in which the rupture propagation velocity exceeds the shear wave velocity, causing intensive ground shaking in the surrounding region. Of particular interest in the 2023 Kahramanmaraş doublet is the role that supershear rupture may have played in triggering multisegment ruptures. The near-fault seismic and geodetic observations available for this doublet provide us with an unprecedented opportunity to determine the multisegment triggering mechanism and ground shaking intensity from a perspective of rupture kinematics.

Faulting geometry and slip model

As continental earthquakes, the surface rupture and ground deformation of the 2023 Kahramanmaraş doublet were well monitored by seismic and geodetic observations. From various synthetic aperture radar (SAR) interferograms and optical images, we constructed three-dimensional (3D) ground deformation of coseismic slips from azimuth, range, and pixel offsets [Fig. 2A; see section S1.4 in (*18*)]. Surface ruptures are well mapped by these images. Event 1 rupture covers a lateral extent of ~380 km, spanning the Amanos ("Seg A"), Pazarcık ("Seg P"), and Erkenek ("Seg E") segments, with surface offsets (averaged for fault traces with > 0.1 m coseismic offsets) of 1.5, 2.8, and 2.1 m resolved on each segment, respectively. A subfault parallel to Seg A is consistent with a branch of aftershock lineation and hypocenter location, which appears to be the initial branch (Nurdağı fault) that hosts the hypocenter and triggered the cascading rupture on the other segments (Fig. 2, A and B). The surface rupture of event 2 is also imaged by the SAR data, which extends ~200 km in the east-west direction spanning the Sürgü, Çardak, and Savrun segments. The epicenter is located at the center of surface rupture, representing a bilateral rupture.

To determine the geometry of fault orientation at depth, we performed a multi-point-source (MPS) inversion (*19*) using teleseismic body waves [section S1.6 in (*18*)]. This method decomposes the kinematic rupture process into sequential ruptures of a cluster of point sources

[1]School of Earth and Space Sciences, Peking University, Beijing 100871, China. [2]Department of Geophysical Engineering, The Faculty of Mines, Istanbul Technical University, 34467 Sarıyer, Istanbul, Türkiye. [3]Department of Earth and Space Sciences, Southern University of Science and Technology, Shenzhen, Guangdong 518055, China. [4]SinoProbe Laboratory, School of Earth and Space Sciences, Peking University, Beijing 100871, China. [5]Department of Earth, Planetary, and Spaces Sciences, University of California, Los Angeles, CA 90095, USA. [6]GNSS Research Center, Wuhan University, Wuhan 430079, China. [7]Department of Geophysical Engineering, Faculty of Engineering, Kocaeli University, 41380 Umuttepe, Kocaeli, Türkiye. [8]Department of Earth and Planetary Sciences, University of California, Riverside, CA 92521, USA.
***Corresponding author. Email: yue.han@pku.edu.cn (H.Y.); xiao.d.song@gmail.com (X.S.)**
†These authors contributed equally to this work.

whose locations and focal mechanisms reflect the fault planes and slips of different segments. The MPS results are consistent with the branch initiation and bilateral rupture of event 1. The high-dip angles of all the MPS solutions for the doublet suggest that the hosting faults are mostly subvertical (Fig. 2 and tables S8 and S9). This feature is also consistent with the relocated aftershocks at depth (Fig. 2B and figs. S8 and S9). Considering the above evidence and focal mechanisms of background seismicity (*20*), we parameterized the faults to be vertical above 10 km and dipping slightly to the northwest direction (between 80° and 85° for different segments) at deeper depths.

The 2023 Kahramanmaraş doublet was well recorded by numerous near-field monitoring stations of various types (Fig. 2A), allowing for a joint inversion of the rupture process using regional strong motion (SM) (table S2), high-rate GPS, static GPS (table S3), and SAR images (table S4). The joint inversion technique exploits the complementary resolution of different datasets (*21*), yielding finite-fault models (FFMs) of both events with robust spatial-temporal resolutions. We also obtained the slip models of the two events simultaneously (Fig. 2C) [section S2.1 in (*18*)], to take advantage of the high spatial resolution of the SAR images (*22*, *23*). We performed a hypocenter relocation for event 1 and set the hypocenter at 37.003°N, 37.235°E, 12.6 km depth, at 1:17:33 UTC, for the FFM inversion. Note that the hypocentral location and timing have a substantial impact on the kinematic rupture models, especially for the initial branch [section S1.5 in (*18*)]. Hereafter, we use timing relative to the hypocenter to describe the rupture process.

For event 1, the rupture started at the southwest of the initial branch, propagated unilaterally to the northeast, triggered Seg P rupture at about 11 s, and was followed by the subsequent bilateral rupture propagation. In the northeast direction, the rupture broke through the Seg P and E junction and continued along Seg E for about 60 km. The major asperity (a locked area of high prestress that is released during an earthquake) was located near the junction between the initial branch and Seg P with a peak slip of ~12 m. Another asperity with a peak slip of ~10 m was identified at the center of Seg E. Although the slips on Seg P and E seem continuous in all ruptured portions, the slip amounts vary substantially. To the southwest direction, the rupture broke the whole of Seg A for ~150 km, on which we observe two main asperities, located 70 and 120 km from the junction, with peak slips of ~6 and ~9 m, respectively. The initial rupture of event 2 was located at the center of the Çardak fault and propagated bilaterally. Its rupture terminated at the western end of the Savrun fault and the eastern end of the Sürgü fault, which bent to the southwest and northeast directions, respectively. The event 2 rupture is distributed more on the Savrun fault than on the Sürgü fault, with a peak slip of ~12 m near the hypocenter (Fig. 2C).

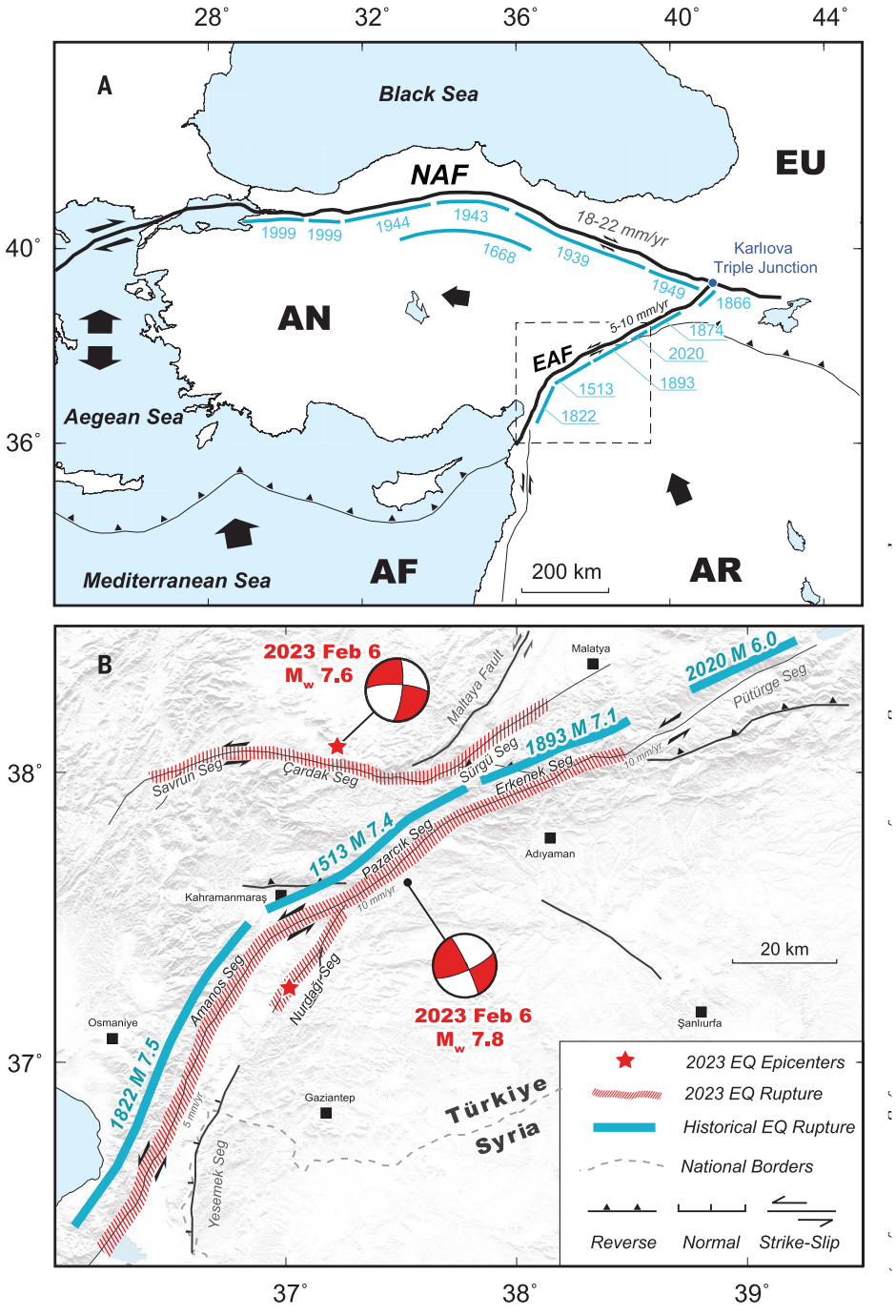


Fig. 1. Maps of faults and earthquakes in the region. (**A**) Regional tectonic setting. Eurasian (EU), Anatolian (AN), African (AF), and Arabian (AR) plates are labeled, with their relative moving directions marked by arrows. The North and East Anatolian Faults are labeled (NAF and EAF, respectively), and the ruptures of large historical earthquakes are marked with blue bars. (**B**) Magnification of the region marked with a dashed box in (A). Seismogenic faults of the 2023 southeastern Türkiye doublet. Ruptured segments of the doublet are marked as red-hatched segments. Moment focal mechanisms of major events during the 2023 sequence are plotted as red-and-white beach balls.

A clear complementary pattern between the distributions of aftershocks (*24*) and coseismic slips was found for both events, but details vary by segment. Aftershocks on Seg P and E were mostly located near the bottom of coseismic rupture, whereas aftershocks on Seg A were more concentrated between asperities,

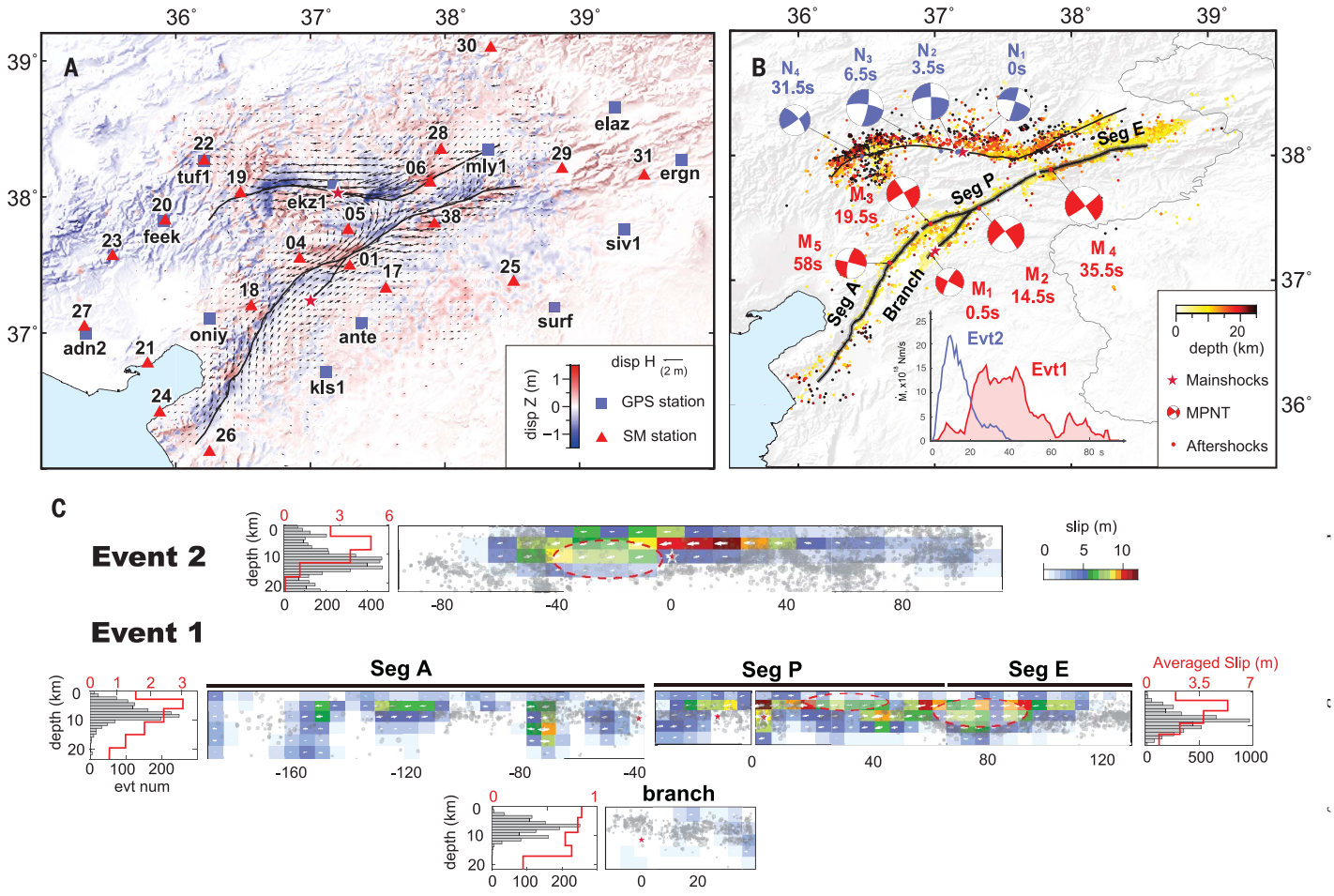


Fig. 2. Coseismic deformation and slip models. (**A**) Coseismic vertical ground uplift (colors) and horizontal displacements (arrows) at the surface. Fault traces are identified from the surface offsets (black curves). Other symbols represent the hypocenters of the doublet (red stars), GPS stations; blue squares), and SM stations (red triangles; assigned names listed in table S2) used in FFM inversion. (**B**) Earthquake and fault rupture data: rupture segments (lines), MPS inversion focal mechanisms (beach balls), and relocated aftershocks (colored dots). The inset shows the source time function of the doublet. The fault segments for event 1 are labeled for the initial branch ("Branch"), Pazarcık ("Seg P"), Amanos ("Seg A"), and Erkenek ("Seg E"). The MPS results are marked as M_1 to M_5 for event 1 and N_1 to N_4 for event 2. (**C**) Slip distributions (colors) of the doublet on the fault segments, along with initiation points of each segment (red stars) and the aftershock distributions (gray dots). The histograms on the side show the depth distribution of the aftershocks (gray filled histogram) and the averaged slip distribution (red curve) along the depth profile in each segment. Voids of aftershocks are marked as dashed red ellipses.

causing different depth concentrations (shallower for Seg A and deeper for Seg P and E) (Fig. 2C). The on-fault aftershocks were related to the stress concentration on the lateral and bottom edges of the asperities, whereas the asperities on Seg A were more scattered, with ruptures being large at shallow depths. We observed two voids of aftershocks on the central Seg P of event 1 and the western portion of event 2 (Fig. 2C), which may be related to supershear rupture behavior that will be discussed in the next section.

Detailed rupture kinematics

The 2023 Kahramanmaraş doublet displays distinctive features in its rupture kinematics (Fig. 3), which illustrate the spatial-temporal evolution of the rupture front over multiple fault segments. The rupture kinematics of event 1 are well constrained by the analysis of the SM data and the joint inversion results (Fig. 3, fig. S13, and table S13).

For event 1, several SM stations are well distributed along Seg A for recording and deciphering fault ruptures, but SM stations located around the initial branch and Seg P and Seg E are not as suitably located, thus making it necessary to use different techniques to analyze the data. For the initial branch, we analyzed the waveforms of two near-fault stations, 01 and 02, which exhibit supershear rupture [section S3.1 in (*18*)]. A pre-hypocentral *S*-wave arrival is identified at station 01 (Fig. 3E), which requires a supershear rupture of the initial branch. Three main asperities are identified from our slip model of the initial branch (Fig. 2 and fig. S19). By carefully examining waveforms at station 01, we identify *P* and *S* phases that are associated with the start and termination of asperity 1 near the hypocenter (fig. S18). Using the relative *S* and *P* times and assuming a straight path of along-fault rupture propagation (1D path), we determined a rupture velocity (V_r) of about 4.0 km/s. A more realistic assumption is a rupture propagation along asperities distributed on the fault plane (2D path). In this case, the along-path V_r is estimated even higher (4.9 to 5.4 km/s; figs. S19 and S20) from asperity 1 to asperity 3 (close to the junction). Given that the upper-crust shear velocity of this region is ~3.6 km/s (*25*), the averaged V_r is well located in the supershear domain no matter which propagation path is assumed [section S3.1 in (*18*)]. Another path is also identified with an upward propagation from asperity 1 to asperity 2 (near the center of the initial branch), which has a distinct subshear V_r (2.6 to 3.0 km/s).

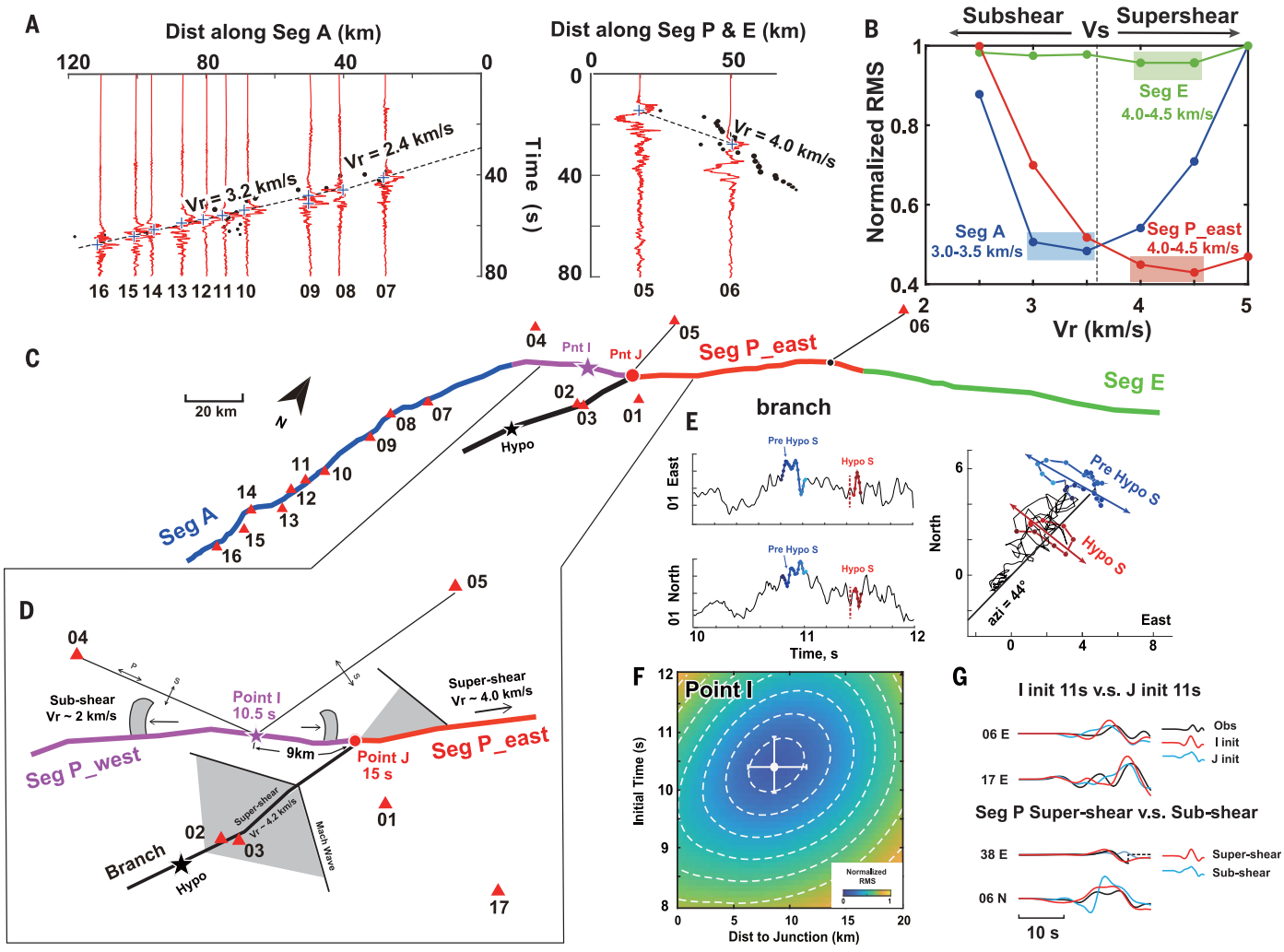


Fig. 3. Analyses of event 1 rupture kinematics. (**A**) SM stations near the fault segments (Seg A, P, and E) whose data are used to estimate the rupture velocities (V_r). The stations are labeled and shown in (C). Rupture arrival times are estimated from peaks of the fault-parallel velocity (blue crosses). The dashed lines indicate linearly fitted V_r from the peak arrivals. (**B**) Estimates of V_r from trade-off curves of residual root mean square (RMS) with V_r in the joint inversion. The curves for Seg A, P, and E are plotted using identical color coding as in (C). The estimated ranges of the optimal V_r are shaded and labeled for each segment. (**C**) Fault segments (colors) of event 1, along with other information, including nearby SM stations (red triangles), hypocenter (black star), the initial point I of Seg P (purple star), and junction point J (red star). (**D**) Magnified view of the initial branch and Seg P for V_r analyses. Areas swept by sub- and supershear wavefronts are illustrated by gray curved shading (upper left) and triangular shading (upper right and bottom), together with estimated V_r values. (**E**) Horizontal ground velocity waveforms (east and north) (left) and particle motions (right) of station 01. Two *S*-wave arrivals are marked, from the hypocenter (red) and a pre-hypocentral arrival (blue). (**F**) Data post-fit residual distribution of station 01 for the initial time and location of point I. The optimized solution and uncertainties are marked with a white cross. (**G**) Waveform comparison for different kinematic parameters. Examples of waveforms show the observed (black) and synthetic waveforms for different initial points (I and J in red and cyan, respectively) and V_r values (super- and subshear in red and cyan, respectively) on the top and bottom panels, respectively.

The discrepancy between the two paths is expected for the respective rupture modes whose paths are parallel and normal to the slip direction (*26*).

Intuitively, the initial branch rupture should have reached the junction (point J) first and triggered subsequent rupture to propagate bilaterally on Seg P. However, the rupture sequence on Seg P is quite complicated. The sequence and triggering mechanism of the rupture from the branch to Seg P are a key part of the rupture process of event 1, which we determined as follows. First, the rupture on Seg P was initiated at another location, about 9 km west of the junction at about 10.5 s after the start of the event [point I, Fig. 3F; section S3.2 in (*18*)]. By analyzing waveform polarity and amplitude of two near-field stations (stations 04 and 05; Fig. 3D and fig. S21), we obtained the particle motion and arrival times of *P* and *S* phases of the Seg P initial rupture, and by backtracking along the ray path and fitting the arrival times of both phases, we could pin down the Seg P initial rupture (point I). This result is also consistent with a 2D grid search for the initial point of Seg P in FFM inversions (fig. S25). Our theoretical travel time calculation using the FFM suggests that the supershear rupture of the initial branch sent out a Mach wave subparallel to the southwest portion of Seg P (fig. S25), which arrived at points I and J in sequence within a short time (between 9 and 10 s). The initiation of point I rupture is likely associated with the Mach wave arrival. Incorporating a numerical simulation experiment (*27*, *28*), we calculated the dynamic stress field caused by the initial branch rupture and found a peak coulomb stress impulse at point I (~1.4 MPa) at ~9.5 s, which is consistent with

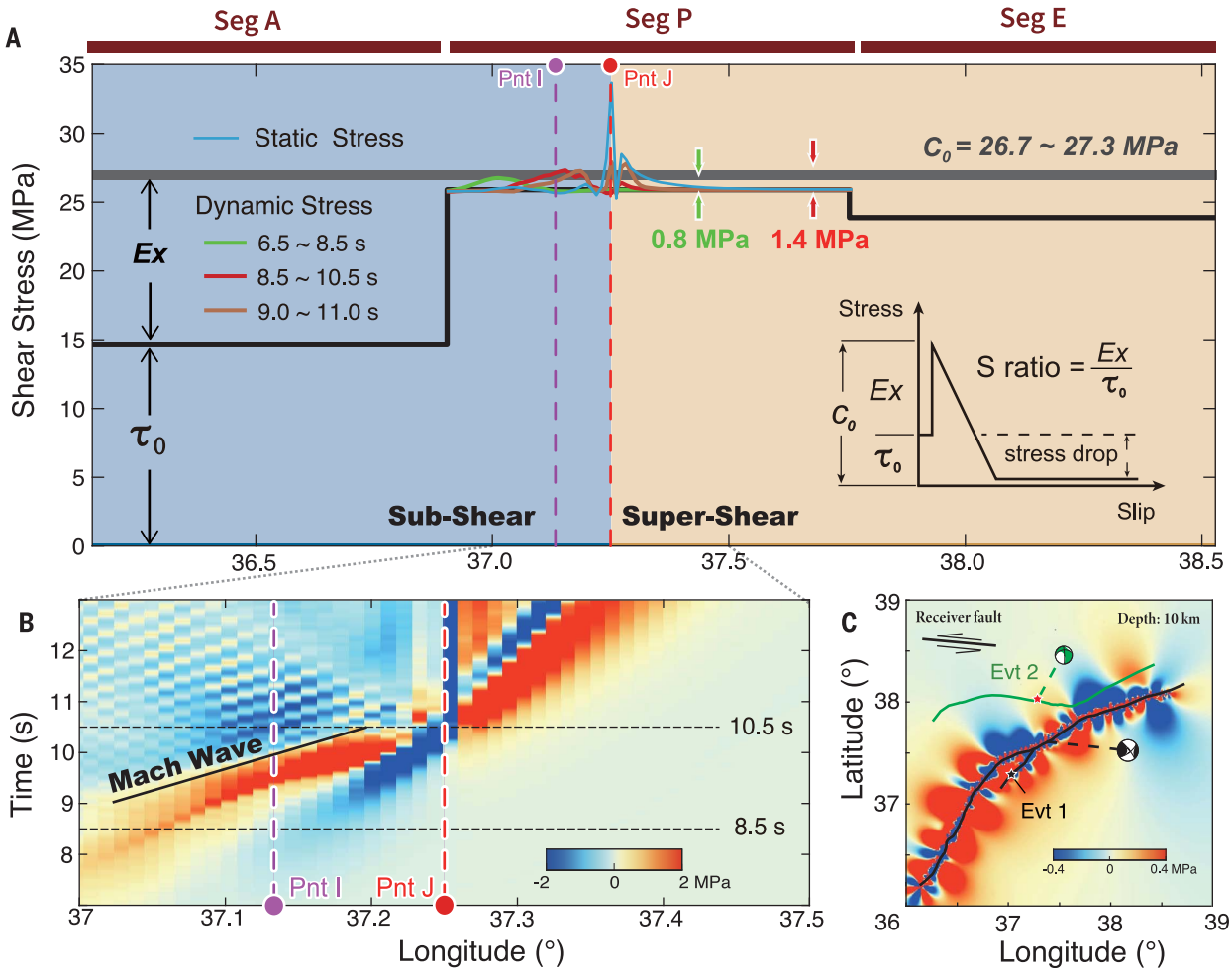


Fig. 4. Stresses on the fault segments. (**A**) Along-strike distribution of prestress and dynamic and static stresses. The prestress is estimated from the stress drop (fig. S25) on Seg A, P, and E and is denoted as black lines for each segment (marked on top). The peak dynamic and static Coulomb stresses due to the initial branch rupture are plotted as colored curves. The purple and red dashed lines mark point I and point J, and blue and orange patches denote the subshear and supershear portions, respectively. (Inset) A graph illustrating the definition of stress parameters. (**B**) Spatial-temporal variation of dynamic Coulomb stress distribution due to initial branch rupture. The Mach wavefront is denoted as a black line. The time range used to extract triggering shear stress (8.5 to 10.5s) is marked with black dashed lines. (**C**) Map view of event 1 coseismic Coulomb stress change at 10 km depth. Other markers are fault traces of event 1 ("Evt 1," black) and event 2 ("Evt 2," green) and dislocation mechanism on receiver fault used to calculate the coseismic Coulomb stress change (top left).

the peculiar triggering sequence (Fig. 4 and figs. S34 and S35). Second, the rupture propagated from point I bilaterally at a subshear V_r (between 2.0 and 2.5 km/s). Its eastward rupture reached the junction point (point J) at ~15 s and then triggered rupture (or rerupture) of the eastern portion of Seg P at a supershear velocity (4.0 km/s). The optimized initial times and V_r values are determined from the waveforms of near-field stations (stations 05, 06, and 01; figs. S26 and S27) as well as a grid search [section S3.3 in (*18*)] to determine the optimal parameters of the FFM. The supershear rupture of Seg P continues to Seg E with a velocity of 4.0 to 4.5 km/s, as determined from FFM inversion [section S2.4 in (*18*)].

The V_r of Seg A is well constrained by the SM stations distributed along the fault, which can be used directly to measure the rupture front (Fig. 3, A and C). By making a correction of the travel times from the fault plane to the stations, we mapped the waveforms to the projection point on the fault and measured the V_r by linearly fitting the times of *S*-wave peak velocities. Waveforms in the distance-time domain indicate that rupture on Seg A initiated at about 27 s after the start of event 1 and propagated southwestward with two-stage velocities: 2.4 km/s for the initial 40 km and 3.2 km/s for the remaining 80 km (Fig. 3A and fig. S28). The averaged V_r of Seg A (3 km/s) is in the subshear regime, which is lower than the shear velocity (~3.6 km/s) during the stable second stage. We classify Seg A as a subshear rupture, although supershear may occur locally (Fig. 3A).

The bilateral V_r of event 2 was determined from the FFM inversion [section S2.4 in (*18*)], which is supershear (4.0 to 4.5 km/s) to the west direction; while its eastward V_r is 3.0 to 3.5 km/s (fig. S11F), close to the shear wave velocity. Kinematic rupture models were also obtained by several studies with various reports of rupture behaviors (*29*–*31*). We believe that our kinematic model makes considerable improvements over the previous ones by aligning the results of joint inversion of multiple datasets with direct waveform analyses [section S2.6 in (*18*)].

Triggering mechanisms and fault interactions

Our analysis shows that the kinematic rupture process varied substantially on the initial branch and the three major segments of event 1, being subshear for Seg A and supershear for the initial branch, Seg P, and Seg E. A theoretical rupture phase diagram (*32*) indicates that higher prestress on a fault tends to host supershear ruptures. We inferred that for event 1, the prestress level might have been the major factor controlling the variation in V_r, as different

elapsed times since the last major events should have introduced various prestress levels at different segments.

We assumed a complete stress drop for all fault segments after the event and set the stress drop as prestress (τ_0) (Fig. 4). We estimated the averaged stress drop of each segment that is weighted by the mean coseismic slip [section S4.2 in (*18*)] and obtained a prestress level of 25.9, 23.9, and 14.6 MPa for Seg P, E, and A, respectively (fig. S31). By comparing the stress drop and stress loading rate [(*14*), section S4.4 in (*18*)], we also found that the ruptured portion of Seg E might not have been ruptured during the 1893 *M* 7.1 event, otherwise the elapsed time would have yielded insufficient prestress to be released by the 2023 event (fig. S36). To simulate and analyze the dynamic triggering process, we adopted elastic wave propagation and slip-weakening friction law and computed the dynamic stress field through a finite difference simulation (*27*, *28*) with the interpolated kinematic slip model [section S4.1 in (*18*)]. The stress excess ($c_0 - \tau_0$), which is defined by the difference between failing criterion (c_0) and prestress (τ_0), is regarded as the minimum stress increase required to initiate a rupture. Thus, it can be estimated from dynamic stresses for Seg P using the following derivation. Considering that the Seg P rupture was initiated at point I at about 10.5 s and that a nucleation time may be required for the rupture to grow from initial triggering to a detectable amplitude (*33*, *34*), the associated peak dynamic stress (1.4 MPa) preceding the initial rupture (8.5 to 10.5 s) is responsible for the triggering, thus it should serve as an upper bound for the local stress excess (Fig. 4A). *S* waves from the initial branch arrived at the southwest portion of Seg P at earlier times, between 8.5 and 9.5 s, yet did not trigger its rupture (Fig. 4A). Thus, peak dynamic stress between 6.5 and 8.5 s on Seg P (0.8 MPa) serves as the lower bound of the stress excess. Adding the Seg P prestress, we determine its c_0 to be between 26.7 and 27.3 MPa (Fig. 4A). Such dynamic triggering analysis is not available for the other fault segments, yet we consider the EAFZ to be a well-developed plate boundary with similar tectonic environments among segments, thus we made a first-order approximation by assuming a uniform c_0 for all segments and determined the stress excess of other segments (Fig. 4A). The associated *S* ratio (*35*), which equals $(c_0 - \tau_0)/\tau_0$, can be computed and mapped for all segments of event 1 (fig. S36). Theoretically, the dynamic V_r is related to the *S* ratio, and it is clear that the supershear segments have a substantially lower *S* ratio relative to the subshear segments, which is strong evidence that the V_r of event 1 was controlled by the prestress level—in other words, high prestress level tends to produce supershear rupture. For event 2, its stress excess is not as well constrained as in event 1, and the estimated *S* ratio also implies a prestress control of V_r during event 2 because its western-part *S* ratio is smaller than that of the eastern part (fig. S37).

We also infer that the initial-branch supershear rupture may have facilitated the supershear rupture excitation on Seg P. Numerical experiments show that the development of supershear generally needs rupture to propagate over some distance, reach the surface, or meet some stress heterogeneities to realize a transition from subshear to supershear, to realize a transition from subshear to supershear (*32*, *36*). In our rupture kinematics, the branch Mach waves arrived at points I and J within a rather short time window (~1 s; Fig. 4), and point J was loaded with concentrated stress near the tip of the initial branch (Fig. 4A), the nucleation or initial rupture of point J might have started around 11 s. Then point J was further loaded by the point I rupture at ~15 s, which might have boosted point J rupture to a supershear velocity [section S3.3 in (*18*)]. Thus, the point I rupture might have facilitated the nucleation or rerupture of point J with extra stress loading [section S4.3 in (*18*)]. Numerical tests suggest that such a rerupturing is capable of initiating a supershear rupture (*32*). The dynamic Coulomb stress change along Seg P increased at point I between 8.8 and 10 s and decreased after the Mach wave arrival (Fig. 4B). Note that the Coulomb stress change before and after the Mach wave arrival was controlled by bulk (*P* wave) and shear (*S* wave) stress changes, respectively, thus, when considering the "direct" effect of the Mach wave (*S* wave), its shear stress inhibited ruptures on Seg P, while the dilatational stress (*37*) preceding the Mach wave was more related to the point I triggering (*33*).

We also performed a scenario earthquake simulation by assuming a subshear V_r rupture of the initial branch, leading to a much-delayed causal triggering (after 12 s; figs. S34 and S35), thus the numerical simulation performed for the kinematic process serves as a cross-validation of the point I initiation on Seg P. The numerical analysis performed in this study is based on kinematic rupture analysis adopting stress failing laws, while our conclusion of low *S* ratio and triggering sequences is also supported by a full-dynamic simulation validation (*33*, *38*). The triggering between events 1 and 2 appears to be associated with static stress instead of dynamic stress, because the events were temporally separated by 9 hours. The coseismic Coulomb stress increase produced by event 1 is about 0.1 MPa near the event 2 hypocenter, which is close to the threshold of coseismic stress triggering (Fig. 4C).

In this study, we resolved rupture behaviors of different fault segments during the 2023 Kahramanmaraş earthquake doublet. We find that initial branch triggering, multisegment rupture, and supershear rupture are three important factors related to the magnitude and damage of this event. The supershear rupture on the initial branch plays an important role in triggering and facilitating the supershear rupture of Seg P. A similar branch initiation occurred for several great continental strike-slip events (e.g., the 2001 Kokoxili, 2002 Denali, 2010 El Mayor-Cucapah, and 2016 Kaikoura earthquakes) (*2*, *16*, *39*, *40*). Such a branch initiation works as a "trigger" of a domino effect, in which a small energy release cascades into a big earthquake. Without the branch rupture, Seg P may have needed another ~80 years to accumulate enough stress to initiate its own rupture from its current loading rate (fig. S38), assuming a uniform and stable stress loading from the creeping faults at depth [section S4.4 in (*18*)]. Specifically, the initial rupture on the main fault developed near the junction between Seg A and Seg P, which was a geometrical barrier owing to fault bifurcation that might have accumulated high prestress to be released only during a large earthquake (*41*). The rupture on the initial branch sent out strong dynamic waves to the subparallel section of the main fault loaded with high prestress and triggered its subsequent rupture. Such splay branch faults are commonly associated with continental strike-slip faults, and because of their close proximities, monitoring and understanding the slip behaviors of these branches is crucial for assessing earthquake hazards of major faults.

The 2023 Kahramanmaraş doublet is also an extraordinary example of interactions between fault segments, in which a physical barrier may be breached during a large event. Previous studies suggest that supershear ruptures are more effective in jumping across fault stepovers and triggering nearby faults (*42*). Besides the cascade rupture sequence, the ground shaking intensity produced by sub- and supershear ruptures also varies substantially. For example, in a ground motion simulation of the 2023 Kahramanmaraş earthquake sequence we find that the supershear rupture increases strong shaking area (intensity above VI) by ~23% (~35,090 km^2) in comparison with a subshear scenario (fig. S39). Thus, the supershear rupture behavior influenced the damage caused by event 1 in terms of both the scope of rupture and the level of ground shaking intensity. The mainshock rupture behavior is influenced by the prestress level, which is worth particular investigation for a better understanding of faulting behaviors of continental fault zones. Earthquake scenarios that incorporate realistic prestress, multisegment rupture, and variation of V_r values are crucial for quantitatively assessing earthquake ground shaking hazards.

REFERENCES AND NOTES

1. C. Lasserre *et al.*, *J. Geophys. Res. Solid Earth* **110**, B12408 (2005).
2. D. Eberhart-Phillips *et al.*, *Science* **300**, 1113–1118 (2003).

3. D. J. Wald, H. Kanamori, D. V. Helmberger, T. H. Heaton, *Bull. Seismol. Soc. Am.* **83**, 981–1019 (1993).
4. X. Jiang, X. Song, T. Li, K. Wu, *Earthq. Sci.* **36**, 169–174 (2023).
5. N. N. Ambraseys, *Geophys. J. Int.* **96**, 311–331 (1989).
6. T. Taymaz, J. Jackson, D. McKenzie, *Geophys. J. Int.* **106**, 433–490 (1991).
7. T. Taymaz, H. Eyidoğan, J. Jackson, *Geophys. J. Int.* **106**, 537–550 (1991).
8. T. Taymaz, J. Jackson, R. Westaway, *Geophys. J. Int.* **102**, 695–731 (1990).
9. N. Turkelli *et al*., *Geophys. Res. Lett.* **30**, 2003GL018023 (2003).
10. A. Hubert-Ferrari, R. Armijo, G. King, B. Meyer, A. Barka, *J. Geophys. Res.* **107**, ETG 9-1–ETG 9-33 (2002).
11. R. S. Stein, A. A. Barka, J. H. Dieterich, *Geophys. J. Int.* **128**, 594–604 (1997).
12. Z. Peng, Y. Ben-Zion, *Geophys. J. Int.* **160**, 1027–1043 (2005).
13. T. Y. Duman, Ö. Emre, *Spec. Publ. Geol. Soc. Lond.* **372**, 495–529 (2013).
14. L. Pousse-Beltran *et al*., *Geophys. Res. Lett.* **47**, e2020GL088136 (2020).
15. C. Zabci *et al*., *Turk. J. Earth Sci.* **20**, 411–427 (2011).
16. T. Wang *et al*., *Earth Planet. Sci. Lett.* **482**, 44–51 (2018).
17. H. Bao *et al*., *Nat. Geosci.* **15**, 942–949 (2022).
18. See supplementary materials.
19. H. Yue, T. Lay, *J. Geophys. Res. Solid Earth* **125**, e2019JB018601 (2020).
20. S. E. Güvercin, H. Karabulut, A. Ö. Konca, U. Doğan, S. Ergintav, *Geophys. J. Int.* **230**, 50–69 (2022).
21. H. Yue, Y. Zhang, Z. Ge, T. Wang, L. Zhao, *Sci. China Earth Sci.* **63**, 492–511 (2019).
22. H. Yue *et al*., *Earth Planet. Sci. Lett.* **570**, 117066 (2021).
23. H. Yue *et al*., *J. Geophys. Res. Solid Earth* **122**, 9166–9183 (2017).
24. H. Ding *et al*., *Earthq. Sci.* **36**, 417–432 (2023).
25. T. Taymaz *et al*., *Tectonophysics* **804**, 228745 (2021).
26. L. B. Freund, *Dynamic Fracture Mechanics* (Cambridge Univ. Press, 1990).
27. W. Zhang, Z. Zhang, X. Chen, *Geophys. J. Int.* **190**, 358–378 (2012).
28. W. Zhang, X. Chen, *Geophys. J. Int.* **167**, 337–353 (2006).
29. Z. Jia *et al*., *Science* **381**, 985–990 (2023).
30. C. Liu *et al*., *Nat. Commun.* **14**, 5564 (2023).
31. L. Xu *et al*., *Commun. Earth Environ.* **4**, 379 (2023).
32. J. Xu, H. Zhang, X. Chen, *Geophys. J. Int.* **202**, 2194–2206 (2015).
33. X. Ding *et al*., *Seismica* **2**, 10.26443/seismica.v2i3.1083 (2023).
34. P. Dong *et al*., *Seismol. Res. Lett.* **94**, 913–924 (2023).
35. S. Xu *et al*., *Nat. Geosci.* **16**, 94–100 (2023).
36. E. M. Dunham, P. Favreau, J. M. Carlson, *Science* **299**, 1557–1559 (2003).
37. M. Mello, H. S. Bhat, A. J. Rosakis, *J. Mech. Phys. Solids* **93**, 153–181 (2016).
38. Z. Wang *et al*., *Geophys. Res. Lett*. **50**, e2023GL104787 (2023).
39. Y. Klinger *et al*., *Bull. Seismol. Soc. Am.* **95**, 1970–1987 (2005).
40. M. Wei, D. Sandwell, Y. Fialko, R. Bilham, *Geophys. Res. Lett.* **38**, L01308 (2011).
41. Z. K. Shen *et al*., *Nat. Geosci.* **2**, 718–724 (2009).
42. R. A. Harris, S. M. Day, *J. Geophys. Res.* **98**, 4461–4472 (1993).

ACKNOWLEDGMENTS

We thank three anonymous reviewers for their constructive comments, which helped improve the article. **Funding:** This project was supported by the National Key R&D Program of China (grant 2022YFF0800602) and the National Science Foundation of China (grants 42025401, U1901602, and 42174059). T.T. acknowledges the Istanbul Technical University-Research Fund (ITU-BAP) and the Alexander von Humboldt Foundation Research Fellowship Award for providing computing facilities and other relevant resources through Humboldt-Stiftung Follow-Up Program. The Peking University research group acknowledges the China National Petroleum Corporation–Peking University Strategic Cooperation Project of Fundamental Research. **Author contributions:** C.R. constructed the fault model and performed MPS and FFM inversion. Z.W. conducted waveform-based analysis of the kinematics. N.H. and W.Z. performed wavefield simulation. Z.Z. conducted the stress analysis. All authors contributed to the initial draft of the manuscript. H.Y. and X.S. conceived of and administrated the project, guided the scientific discussion, wrote most of the manuscript, and acquired the funding. Z.S. contributed to the scientific discussion and manuscript editing. H.L., T.W., Z.L., and Hang Xu processed the SAR and optical images and contributed to manuscript writing on SAR data processing. T.T., T.S.I., and C.E. collected continuous local and regional broadband seismograms, strong ground motion waveforms, and GPS and high-rate GPS data and contributed to manuscript editing and interpretation of the active tectonics of Eastern Anatolia, Türkiye. J.G. processed the original GPS data. Haoyu Xu, Z.G., Y.Z., H.D., and B.C. contributed to the hypocenter relocation of the mainshocks and the aftershock catalog. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The teleseismic data were download from IRIS (https://www.iris.edu/hq/). The strong motion waveforms were retrieved from AFAD (https://tdvms.afad.gov.tr/event_spec_data). The GPS data were provided by the CORS-TR (TUSAGA-Aktif: https://www.tusaga-aktif.gov.tr) GNSS network administrated by the General Directorate of Land Registry and Cadastre (TKGM) and General Directorate of Mapping (HGM-Ankara, Türkiye). The Sentinel-1 SAR data and Sentinel-2 optical data were provided by the European Space Agency (https://scihub.copernicus.eu), and the Sentinel-1 data were distributed by the Alaska Satellite Facility (https://asf.alaska.edu/how-to/data-tools). The Advanced Land Observation Satellite-2 (ALOS-2) SAR data used in this work are under the copyright of the Japan Aerospace Exploration Agency (JAXA) and were openly accessed at https://www.eorc.jaxa.jp/ALOS/en/dataset/alos_open_and_free_e.htm.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adi1519
Supplementary Text
Figs. S1 to S39
Tables S1 to S13
References (*43*–*73*)

Submitted 15 April 2023; accepted 5 December 2023
10.1126/science.adi1519

DNA-COLLOID CRYSTALS

Space-tiled colloidal crystals from DNA-forced shape-complementary polyhedra pairing

Wenjie Zhou[1,2]†, Yuanwei Li[1,3]†, Kwanghwi Je[4], Thi Vo[4], Haixin Lin[1,2], Benjamin E. Partridge[1,2], Ziyin Huang[1,5], Sharon C. Glotzer[4,6]*, Chad A. Mirkin[1,2,3,5]*

Generating space-filling arrangements of most discrete polyhedra nanostructures of the same shape is not possible. However, if the appropriate individual building blocks are selected (e.g., cubes), or multiple shapes of the appropriate dimensions are matched (e.g., octahedra and tetrahedra) and their pairing interactions are subsequently forced, space-filled architectures may be possible. With flexible molecular ligands (polyethylene glycol–modified DNA), the shape of a polyhedral nanoparticle can be deliberately altered and used to realize geometries that favor space tessellation. In this work, 10 new colloidal crystals were synthesized from DNA-modified nanocrystal building blocks that differed in shapes and sizes, designed to form space-filling architectures with micron-scale dimensions. The insights and capabilities provided by this new strategy substantially expand the scope of colloidal crystals possible and provide an expanded tool kit for researchers interested in designing metamaterials.

Polyhedral nanocrystals (NCs) can be assembled into a wide variety of complex colloidal crystals (*1*, *2*), most of which have structural counterparts on the atomic scale. Other structures have only been observed with colloidal crystals, and access to such structural diversity makes possible new materials with unprecedented properties (*3*–*5*). DNA ligands chemically immobilized on polyhedral NCs can program interactions between individual building blocks; this approach can direct the thermodynamic assembly of highly ordered colloidal crystals with almost 100 different crystal symmetries (*6*–*14*). However, the long DNA ligands conventionally used in colloidal crystal engineering precluded the assembly of space-filling polyhedral particle shapes into their three-dimensional (3D) tilings because the ligands can lead to loss of the NC anisotropy, which impacts their assembly (Fig. 1A). For example, DNA ligation of a nanocube can produce a particle with an altered shape that will pack into crystal structures more complex than the face-to-face, simple cubic packing of the bare nanocube (*15*, *16*). This modification of particle shape prevents the use of simple geometric principles to predict the appropriate polyhedral NCs for assembly into space-filling structures, and many of the previously observed assemblies are often counterintuitive (*15*–*17*). Decoupling the interplay between particle-particle interactions and core geometry would be a major advance in the design of space-filling NC assemblies.

We explored decoupling of the contributions between the DNA ligand shell and core shape in driving polyhedral NC assembly using shorter DNA ligands (~20 base pairs) to achieve space-filling arrangements that reflected the NC shape. However, such particles have thus far not yielded highly ordered crystals (*16*, *18*, *19*). The observed frustrations of NC assembly arise because of two reasons: (i) limited flexibility of the short, rigid, double-stranded DNA (dsDNA), which cannot compensate for the intrinsic NC polydispersity and (ii) high charge densities associated with the dsDNA that disfavor facet alignment. A key challenge in the use of DNA bonding agents for assembling colloidal crystals into space-filling structures is to design ligands that mediate facet-to-facet alignment of polyhedral NCs that reflect NC shape.

In this work, we present a DNA design that involves a short, flexible, charge-neutral region: an oligoethylene glycol (OEG) spacer (Fig. 1, B and C) adjacent to the sticky end that minimally alters the native geometries of NCs and facilitates the formation of highly ordered assemblies from shape-complementary polyhedral NCs (Fig. 1D). Placing a flexible spacer between the rigid dsDNA anchor strand and the single-stranded DNA sticky end decoupled the binding of sticky ends from the NC shape.

This decoupling enabled two new design strategies possible, which we demonstrated experimentally and confirmed with molecular simulations and free energy calculations. First, decoupling allowed imperfect polyhedral building blocks (ones that varied slightly in size and shape) to assemble into highly ordered space-filling structures commensurate with the 3D tilings associated with the perfect shape. We then used geometric tilings consisting of particles that differ in shape to define Voronoi polyhedra—larger particles whose shapes reflected that of the Voronoi cells of a target crystal [illegible] order to design assemblies of highly ordered crystals, even for NCs that were not perfectly space filling. Second, for non–space-filling polyhedral NCs, the additional degrees of freedom provided by the flexible spacer gave access to complex crystals with symmetries not previously possible with DNA.

Assembly of space-filling NCs

OEG is a highly flexible, charge-neutral spacer molecule that can be incorporated readily into synthetic oligonucleotides and used as a flexor region (Fig. 1, B and C). DNA strands featuring *n*-hexaethyleneglycol units (sp*n* for n = 0, 1, 2, 3, 5) were designed with self-complementary sticky ends that could hybridize to the sticky ends on any other particle in the system (DNA designs and sequences are illustrated in fig. S1).

To assess whether DNA with an OEG flexor region would permit the assembly of polyhedral NCs able to tile 3D space into space-filling colloidal crystals, we investigated the assembly of three space-filling NC shapes. Specifically, cubes, truncated octahedra (TO), and rhombic dodecahedra (RD) with high yield and size polydispersity (CV) <5% (detailed statistics are provided in table S2) were synthesized and functionalized with anchor DNA strands complementary to DNA linker strands containing the flexor region through salt-aging. NCs were then assembled by using the linker strands through slow-cooling and annealing (supplementary materials). We identified lattice symmetry and particle orientation by combining both electron microscopy (EM) (figs. S30–S32) and synchrotron small-angle x-ray scattering (SAXS) analyses (fig. S29) of the superlattices.

From geometry, cubes, TO, and RD are the shapes of the Voronoi cells of simple cubic (sc), body-centered cubic (bcc), and face-centered cubic (fcc) structures, respectively. Based on the complementary contact model (*7*), which maximizes DNA hybridization by facet alignment between neighboring shapes, these Voronoi polyhedra should assemble into their space-filling crystal structures with a high degree of ordering if the OEG spacer strategy successfully decoupled linker architecture from the core particle shape (Fig. 2, A to C) (*20*). Nanocubes (edge length E = 110 nm) with sp0 linkers (the rigid DNA linker that contains no OEG spacers) produced malformed colloidal crystals with few discernible domains with sizes >1 μm (fig. S2A). Increasing the length of the OEG moiety led to single sc crystals with dramatically larger crystal domains (Fig. 2D): cubes with sp1, sp2, sp3, and sp5 linkers assembled into large sc crystals with cubic habits and average crystal domain sizes of 6, 10, 20, and 40 μm, respectively (Fig. 2G and fig. S2A). The largest domain size (56 μm) was achieved with sp5 linkers. Similarly, TO (E = 40 nm)

[1]International Institute for Nanotechnology, Northwestern University, Evanston, IL 60208, USA. [2]Department of Chemistry, Northwestern University, Evanston, IL 60208, USA. [3]Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL 60208, USA. [4]Department of Chemical Engineering, University of Michigan, Ann Arbor, MI 48109, USA. [5]Department of Materials Science and Engineering, Northwestern University, Evanston, IL 60208, USA. [6]Biointerfaces Institute, University of Michigan, Ann Arbor, MI 48109, USA.
*Corresponding author. Email: chadnano@northwestern.edu (C.A.M.); sglotzer@umich.edu (S.C.G.)
†These authors contributed equally to this work

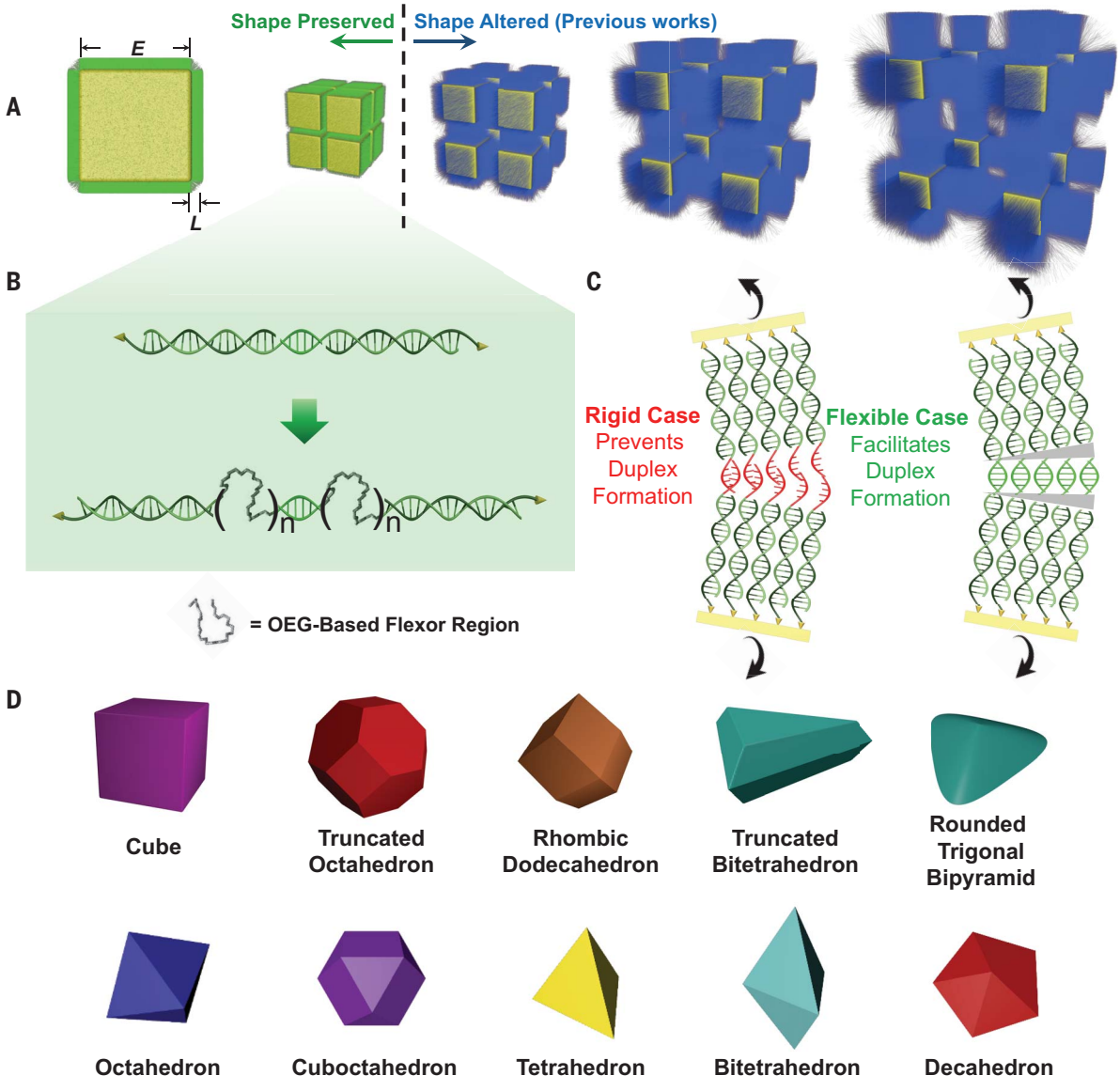


Fig. 1. DNA-crystallized polyhedral nanocrystals. (**A**) Schematic illustration of colloidal crystals assembled from cube NCs functionalized with different lengths of DNA (*L*). For a given size of NC *E*, long DNA ligands (conventionally used in colloidal crystal engineering with DNA) can alter the shape of underlying polyhedron core. (**B**) Illustrations of previously used rigid DNA design and newly developed flexible DNA design. (**C**) Depiction of rigid and flexible DNA attached to two flat surfaces upon angular strain. The gray triangles represent the OEG flexor regions. (**D**) 3D models of all shapes of NCs used in this work. The statistical analysis for each polyhedral NC sample can be found in table S2.

assembled into bcc colloidal crystals with average domain sizes up to 37 μm (Fig. 2, B, E, and H, and fig. S2B), and the largest observed domain size was 51 μm. RD (*E* = 52 nm) assembled into fcc colloidal crystals with truncated tetrahedral habits and average domain sizes up to 40 μm (Fig. 2, C, F, and I, and fig. S2C), and the largest observed domain size was 57 μm.

Assembly of nearly space-filling NCs

The OEG modification lengthened the ligands by only a small amount (0 to 1.6 nm per sp unit depending on whether the flexors coil or fully stretch), and thus, the most they could contribute to the interparticle spacing was ~8 nm for sp5, which was orders of magnitude less than the resulting average crystal domain sizes. However, the small flexor regions provide sufficient flexibility to compensate for the size polydispersity that suppressed long-range ordering in their absence. We hypothesized that this compensating characteristic might be sufficiently strong to enable the assembly of not-quite-space-filling polyhedral NCs into the crystal structures associated with the Voronoi polyhedron most similar in shape to the NC. Among the NC shapes that are synthetically accessible, very few can fill space perfectly, and their geometric packings contain structural voids and nonparallel gaps. We hypothesized that the flexor regions in the DNA linkers would eliminate these "tiling defects" that would otherwise hinder successful crystallization by leaving many DNA sticky ends unlinked while keeping the original shape unaltered. If successful, this approach would dramatically expand the number of space-filling colloidal crystals possible with DNA even without having access to space-filling NC shapes.

To test this hypothesis, we first considered the truncated bitetrahedron (TBT), a polyhedral NC whose shape closely approximates the Voronoi cell shape of the interlocked honeycomb (ih) lattice (Fig. 3A and fig. S3A). When TBT NCs (*E* = 100 nm) were assembled with sp1, sp2, and sp3 linkers, ih colloidal crystals were observed with hexagonal prismatic habits and an average domain size approaching 30 μm (Fig. 3B).

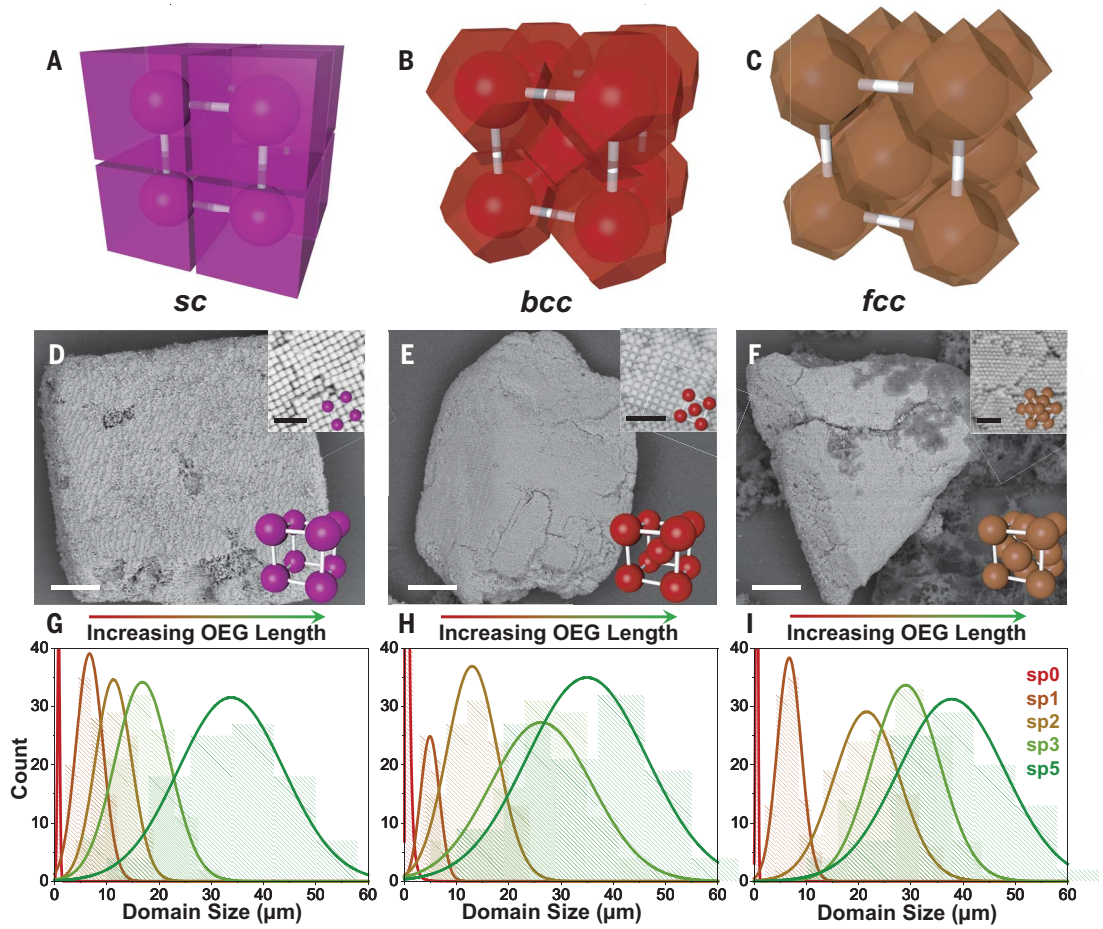


Fig. 2. Flexible packing of NCs with space-filling shapes. (**A** to **C**), Schematic illustration of sc (A), bcc (B), and fcc (C) lattices composed of cubes, TOs, and RDs, respectively. (**D** to **F**) Representative scanning electron microscopy (SEM) images of colloidal crystals formed from DNA-engineered cubic- [(D), *E* = 110 nm], TO- [(E), *E* = 40 nm], and RD- shaped [(F), *E* = 52 nm] NCs. Scale bars, 5 μm; inset scale bars, 500 nm. (**G** to **I**) Statistical analyses of single-crystalline domain sizes crystallized from cubic- (G), TO- (H), and (I) RD-shaped NCs by means of sp0, sp1, sp2, sp3, and sp5 oligonucleotide linkers.

We next considered "squashed" triangular bipyramid (TBP) NCs (*E* = 200 nm), which were previously reported to assemble into three types of clathrate colloidal crystals (types I, II, and IV), the most complex colloidal crystal structures obtained to date (*18*). The same type of clathrate colloidal crystals could be produced with high quality (fig. S12) by using the OEG flexor approach. Although the squashed TBP resembles the ih Voronoi polyhedron, there are substantial differences, such as the number and spatial arrangements of their lateral edges and midsection vertices, and evidently, the OEG flexor approach did not compensate for such differences. However, when TBP NCs (*E* = 150 nm) with rounded midsection vertices (rTBPs), analogous to the ih Voronoi polyhedron, were assembled by using the flexor DNA design (sp2, sp3, and sp5), ih crystals were formed with elongated hexagonal bipyramidal crystal habits (Fig. 3, C and D, and fig. S3C).

Assembly of non–space-filling NCs

Octahedra are substantially different from known Voronoi cells and do not fill 3D space well, and octahedral NCs assemble into crystals with many different symmetries (*21*). Moreover, multiple assembled structures often coexist for a given set of experimental growth conditions (*22*, *23*), which hinders assembly guided by simple geometric principles. Using flexible DNA linkers, we discovered four distinct assemblies that depend on the size of octahedral NCs (*E* between 30 and 120 nm, linker = sp1, sp2, and sp3). When large (*E* > 80 nm) octahedra were used, crystals with simple hexagonal (sh) arrangements and hexagonal prismatic habits were observed in high yield (Fig. 3, E and I). When smaller octahedra (*E* between 70 and 80 nm) were assembled, we observed base-centered monoclinic (bcm) structures, Minkowski structures (*24*), or mixtures of the two (Fig. 3, F and J, fig. S15, and supplementary materials). When octahedral NCs with *E* between 55 and 75 nm were used, I-43d lattices (cubic lattices containing 16 octahedra within each unit cell, all with different orientations) (figs. S30 to S32) were obtained (Fig. 3G and fig. S4C). This I-43d crystal, also referred to as the high-pressure lithium structure (*21*), exhibits a triakis tetrahedral crystal habit (Fig. 3K). The smallest set of octahedra (*E* < 40 nm) assembled into bcc lattices with RD crystal habits (Fig. 3H, L), consistent with a previous report (*19*).

We hypothesized that the observed assembly differences were correlated with the ratio of DNA length (*L*) to octahedron edge length (*E*). To test this hypothesis, instead of varying the octahedron size while keeping the DNA lengths the same (as above), we used octahedral NCs with a single edge length (*E* = 85 nm) and assembled them with DNA linkers of different lengths (sp1, no additional double-stranded extension region and one OEG unit; d20, a 20-bp extension region; and d40, a 40-bp extension region) (table S1 and fig. S1C). We observed the same trend as when octahedron size was varied and DNA length was held constant: as *E*/*L* decreased, we observed the formation of sh crystals with the shortest d0 linkers, I-43d crystals with the d20 linkers, and bcc crystals with the longest d40 linkers (Fig. 3M, fig. S5, and table S2).

The various observed colloidal crystals of octahedra can be explained by considering the DNA spatial distribution about the core octahedron (fig. S20) (*15*, *25*). Given the same ligand length, we expect increased DNA partitioning toward the corners and edges of the octahedron with decreasing core size. Prior work has predicted and demonstrated this partitioning tendency for various types of ligands arising from the entropic penalty suffered when long, flexible ligands or bulky ligands are confined to a facet (*15*, *26*). The OEG flexor adds flexibility to the DNA ligands, making partitioning likely, especially at the highly acute polar tips of the

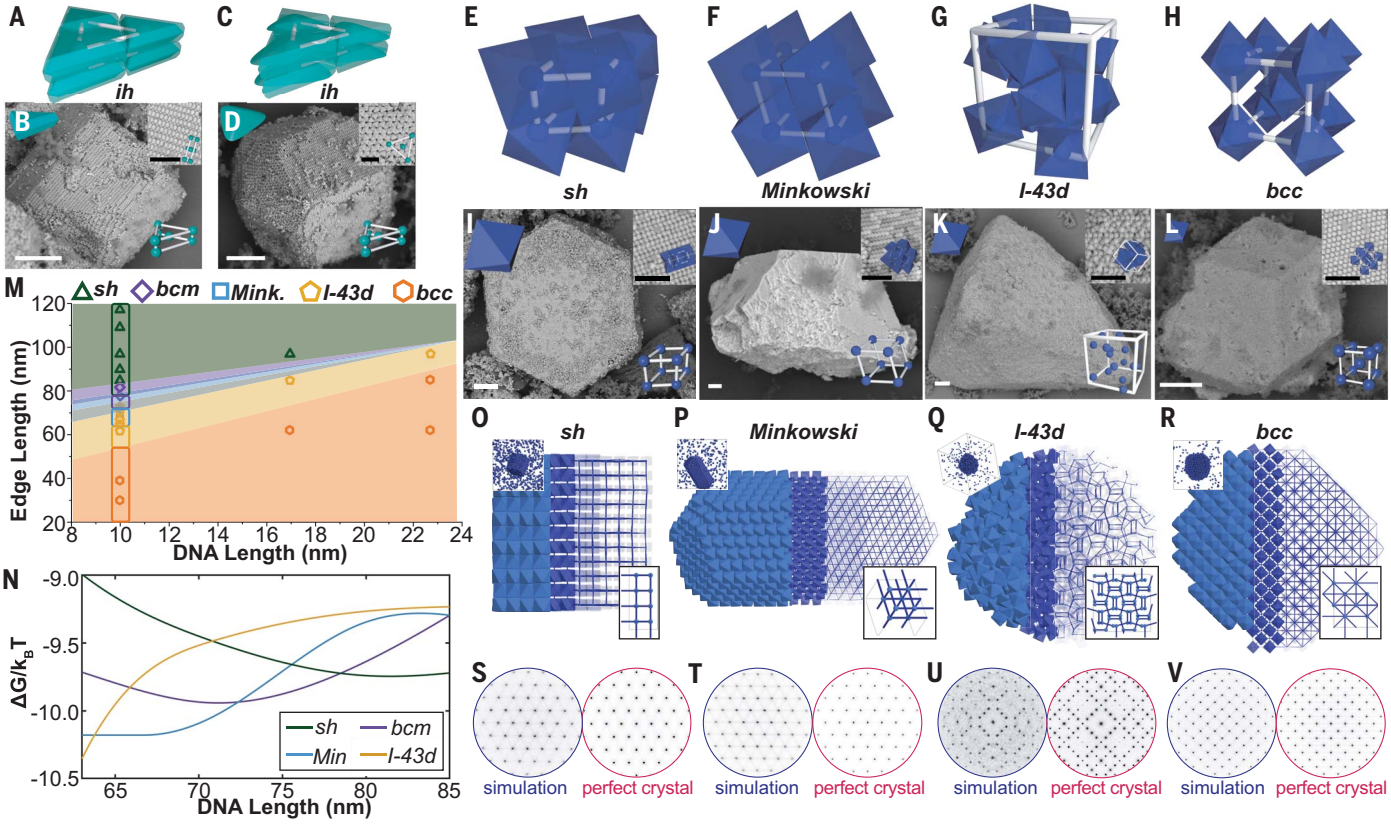


Fig. 3. Flexible packing of NCs with non-space-filling shapes. (**A**) Schematic illustration of ih lattice composed of TBT NCs. (**B**) Representative SEM images of a hexagonal prism–habited colloidal crystal with ih symmetry formed from TBT-shaped NCs (E = 100 nm). (**C**) Schematic illustration of ih lattice composed of rTBP NCs. (**D**) Representative SEM images of elongated hexagonal prism–habited colloidal crystals with ih symmetry formed from rTBP-shaped NCs (E = 150 nm). (**E** to **H**) Schematic illustrations of sh (E), Minkowski (F), I-43d (G), and bcc (H) lattices packed from octahedra. (**I** to **L**) Representative SEM images of colloidal crystals with sh (I), Minkowski (J), I-43d (K), and bcc (L) symmetries formed from octahedral NCs (E = 95, 71, 65, and 39 nm, respectively). Identifiable crystal habits are hexagonal prisms (I), triakis tetrahedra (K), and RDs (L). Scale bars, 2 μm; inset scale bars, 500 nm. (**M**) Experimental and theoretical phase diagrams for octahedron-assembled lattices as a function of octahedron edge length and DNA length. At DNA length L = 10 nm, the experimental phase diagram is overlaid by color-coded boxes representing theoretical phase diagram. (**N**) Lattice energies of competing structures sh, bcm, Minkowski, and I-43d near the regime (E/L) ~ 8. (**O** to **R**) Superlattices of octahedra assembled in MD simulations. In each figure, a snapshot of simulation box (top left), solid clusters extracted from the snapshot (center), and 2×2×2 unit cells of crystals (bottom right) are shown. With decreasing E/L, octahedra assemble into sh (O), Minkowski (P), I-43d (Q), and bcc (R) phases. For solid clusters, the transparency of octahedron is increased from left to right to show local coordination. (**S** to **V**) Diffraction patterns obtained from snapshots in (O) to (R) (left) and from perfect crystals (right), showing the same symmetry and peak positions.

octahedron (in contrast to the other shapes studied). For a given linker length, larger octahedra should have higher DNA ligand densities on their faces than smaller octahedra, favoring close-packed face-face alignment and thus, the sh phase. Smaller octahedra conversely should have a higher ligand concentration at the corners and edges, favoring corner-edge alignments that result in open lattices such as that of bcc.

To test this hypothesis, we used thermodynamic perturbation theory (TPT) (*27*) to compute the lattice free energy of formation of competing structures (supplementary materials). Lattice energies were calculated across a range of octahedron edge lengths E from 20 to 120 nm, with the length L of the DNA ligand kept constant at 10 nm. The predicted phase diagram is shown in Fig. 3M (color-coded boxes). Over a wide range of edge lengths, the theoretical results agree well with experiments (Fig. 3M). Near the $(E/L) \approx 8$ regime, the energy differences between bcm, Minkowski, I-43d, and sh phases were very small (less than ~1.5× the thermal energy k_BT, where k_B is Boltzmann's constant and T is temperature) (Fig. 3N), which can result in the formation of multiple phases from the same octahedron, as observed in our experiments.

We also performed molecular dynamics (MD) simulations of octahedral NCs with DNA ligands using the HOOMD-Blue simulation toolkit (*28*–*30*). We modeled each octahedral NC as a rigid polyhedron core surrounded by a flexible DNA ligand shell (supplementary materials and fig. S21) using a simplified effective potential (supplementary materials). To account for the hypothesized spatial distribution of DNA on the NC surface, the interaction strength between DNA ligands on the NC facets was increased as the ratio E/L increased (supplementary materials). We selected four E/L ratios—12, 7, 6, and 3—and simulated the assembly (both seeded and nonseeded) of octahedral NCs with the hypothesized spatial distribution of DNA. We characterized the resulting colloidal crystals visually and by their diffraction patterns (Fig. 3, S to V). For (E/L) = 12, 7, 6, and 3, we observed the formation of sh (Fig. 3O and fig. S24), Minkowski (Fig. 3P and fig. S25), I-43d (Fig. 3Q and fig. S23), and bcc (Fig. 3R and fig. S22), which were in agreement with our experiments and free energy calculations. The TPT calculations and MD simulations confirmed that the spatial distribution of DNA ligands was critical to explaining the octahedral NC assembly results.

Coassembling tilings of shape-complementary polyhedral NCs

Although many convex polyhedra cannot tile space on their own, they can tessellate with other polyhedra to fill space (*31*, *32*). We hypothesized

that, by considering the shape complementarity of a pair of polyhedra with matching edge lengths, we could use the flexor DNA design to engineer space-filling, multicomponent colloidal crystals. We first chose one of the simplest 3D space-tessellation pairs, cuboctahedra and octahedra, for which there is only one possible configuration to tessellate space, a rectified cubic honeycomb (rch) packing (*32*) (Fig. 4, A and B). Cuboctahedral and octahedral NCs synthesized with identical edge lengths (both E = 61 or 95 nm) were functionalized with the same surface-bound DNA and mixed at a ratio of 1:1 before slow-cooling along with the self-complementary DNA linker (sp1 for E = 61 nm, sp5 for E = 95 nm). Coassembled CsCl crystals with rch packing symmetry formed as the main product (Fig. 4C). The self-complementarity of the DNA on both types of particles could have enabled phase separation of the two shapes into non–space-filling, single-component colloidal crystals, each of which could result in many unhybridized DNA linkers.

The combination of tetrahedra and octahedra presented a more challenging space-tessellation pair because these shapes can tile 3D space with two different configurations, both with a ratio of one octahedron to two tetrahedra. Three octahedra and four tetrahedra together comprise a local space-filling arrangement, forming a triangular honeycomb structure (fig. S6A), which can then be expanded in two ways: the tetra-octa honeycomb (toh) (Fig. 4, D and E, and fig. S6B) and the gyrated tetra-octa honeycomb (gtoh) (fig. S6C) (*33*). Although both tessellation arrangements are space filling, the symmetries and geometric configurations are very different. Notably, when tetrahedral and octahedral NCs were coassembled in the same manner as the cuboctahedra and octahedra above, the toh structure with an octahedral crystal habit was formed, but the gtoh structure was not observed (Fig. 4F). TPT calculations of the lattice free energies of the two phases (fig. S26) show that toh is lower in free energy than gtoh and thus, more thermodynamically stable, supporting our experimental observations.

Notably, in toh, tetrahedra are surrounded only by octahedral neighbors, whereas in gtoh, each tetrahedron contacts one other tetrahedron. In fact, all of the tetrahedral building blocks in the gtoh structure can be grouped into bitetrahedral pairs to form a tessellation of octahedra and bitetrahedra (a Johnson solid, J12). Based on this tessellation, we hypothesized that replacing two tetrahedral building blocks with one equivalently sized bitetrahedral building block would favor formation of the gtoh structure (Fig. 4, G and H). Indeed, such geometric replacement could only be experimentally realized when the DNA ligands minimally affected the native geometries of the NCs (fig. S33), a crucial feature enabled by the flexible OEG linkers. To test this hypothesis, DNA-functionalized octahedral and bitetrahedral NCs with identical edge lengths (E = 78 nm) were synthesized and coassembled as described above. As predicted, a gtoh structure with a well-defined hexagonal prismatic crystal habit was observed as the main product (Fig. 4I, fig. S7, and supplementary materials). The coassembly of octahedral and bitetrahedral NCs demonstrated that sophisticated colloidal crystal structures can be obtained by compounding shape-complementary building blocks.

Using MD simulations, we assembled each of the three colloidal crystals formed by the space-filling polyhedron pairs. We observed the assembly of the rch phase from a fluid mixture of DNA ligand–coated octahedra and cuboctahedra with 1:1 stoichiometric composition, the toh phase from a mixture of octahedra and tetrahedra with 1:2 stoichiometric composition, and the gtoh phase from a mixture of octahedra and bitetrahedra with 1:1 stoichiometric composition (Fig. 5, A to C). The diffraction patterns of the assembled superlattices show well-defined Bragg peaks that correspond to each of these phases. These crystal phases could be assembled in simulations from binary mixtures of hard polyhedra (fig. S27, A to C), demonstrating that shape complementarity of the space-filling polyhedron pair was the essential feature driving the assembly of these superlattices, and that our flexible DNA-graft minimally altered this shape complementarity.

Next, we applied this shape complementarity strategy to a quasi–space-filling pair: decahedra (a Johnson solid, J13) and octahedra. A

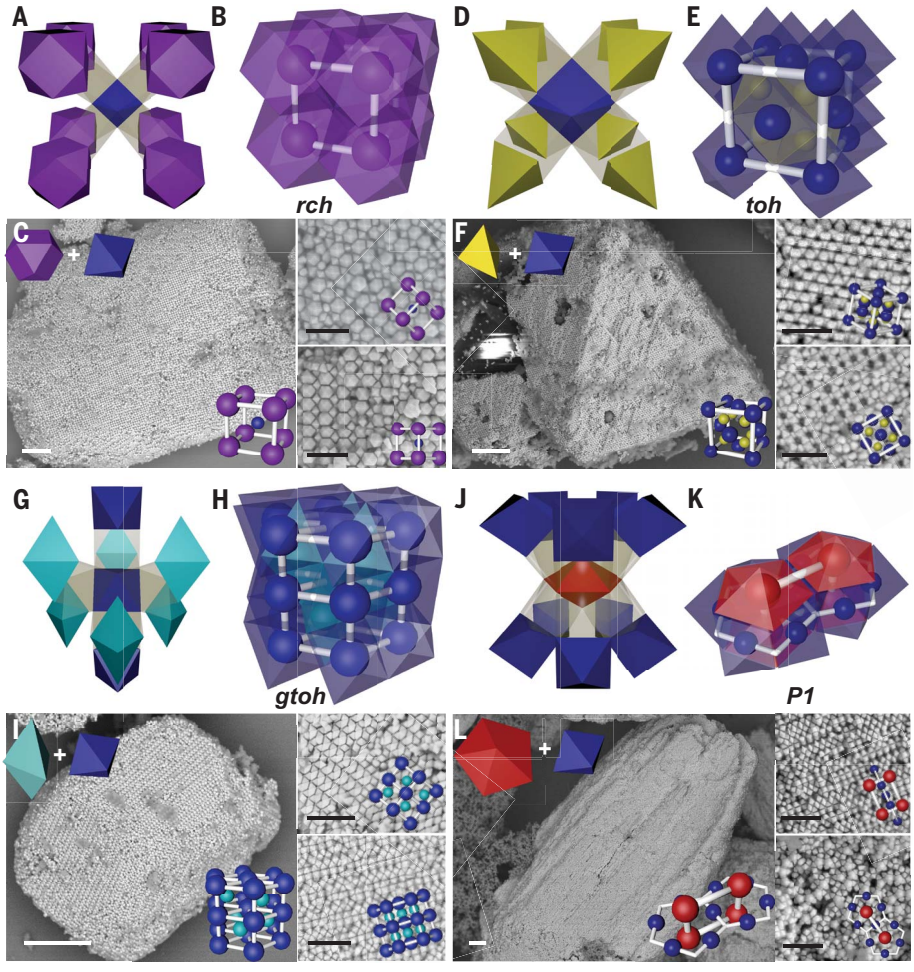


Fig. 4. Flexible copacking of NCs with complementary shapes. (**A** and **B**) Schematic illustration of rch space-filling configuration. (**C**) Representative SEM images of cubic-habited colloidal crystals with rch symmetry formed from cuboctahedral and octahedral NCs (both E = 95 nm). (**D** and **E**) Schematic illustration of the toh space-filling configuration. (**F**) Representative SEM images of octahedral-habited colloidal crystals with toh symmetry formed from tetrahedral and octahedral NCs (both E = 82 nm). (**G** and **H**) Schematic illustration of the gtoh space-filling configuration. (**I**) Representative SEM images of hexagonal prism–habited colloidal crystals with gtoh symmetry formed from bitetrahedral and octahedral NCs (both E = 78 nm). (**J** and **K**) Schematic illustration of P1 quasi–space-filling configuration. (**L**) Representative SEM images of colloidal crystals with P1 symmetry formed from decahedral and octahedral NCs (both E = 85 nm). Scale bars, 2 μm; inset scale bars, 500 nm.

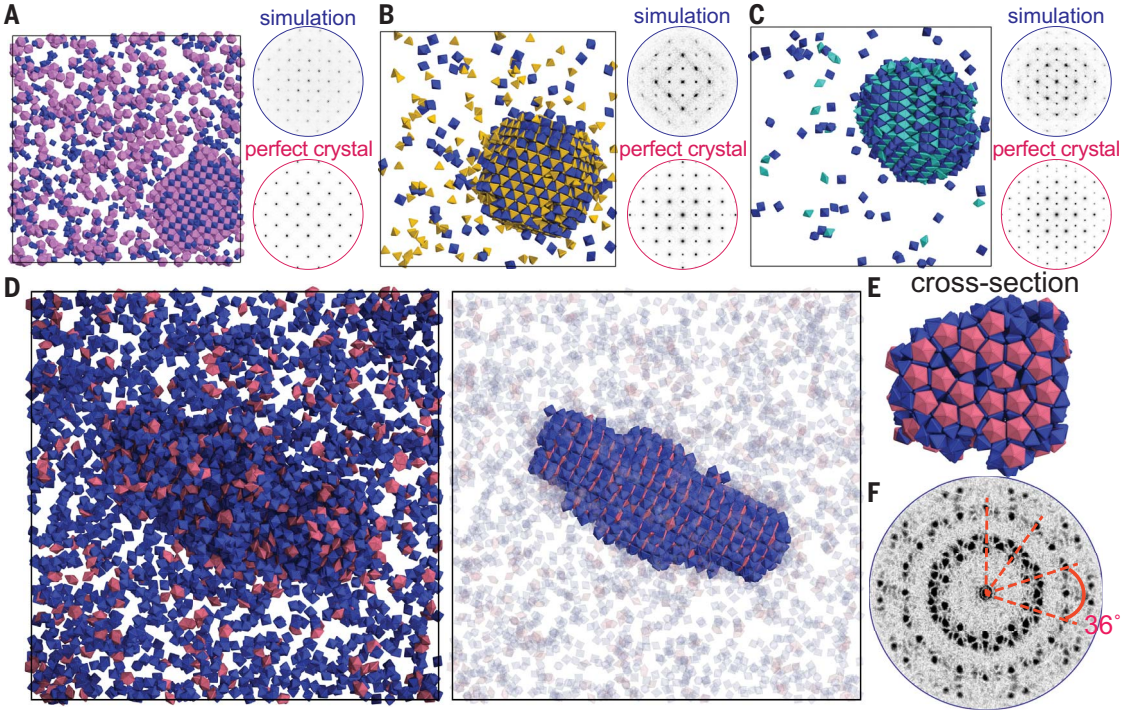


Fig. 5. Molecular dynamics simulations of cocrystallization. (**A** to **C**) Assembly of rch phase from DNA-coated cuboctahedra (purple) and octahedra (blue) (A), assembly of toh phase from DNA-coated tetrahedra (yellow) and octahedra (blue) (B), and assembly of gtoh phase from DNA-coated bitetrahedra (cyan) and octahedra (blue) (C). (Left) MD simulation snapshots of the particle assemblies. (Right) diffraction patterns obtained from the snapshots (upper right) and from perfect crystals (lower right). Diffraction patterns show the same symmetry and peak positions. (**D**) Snapshots obtained from MD simulations of binary mixture of DNA-coated decahedra (red) and octahedra (blue). The same snapshot is shown twice: (left) the particle assembly and (right) to visualize the P1 superlattice in the assembly. Particles not belonging to the superlattice are colored translucently (supplementary materials). The superlattice is a periodic stacking of layers. (**E**) Cross section of the P1 superlattice viewed along the stacking direction. (**F**) Diffraction patterns of the P1 superlattice, indicating decagonal symmetry.

decahedron can be approximated as a compound of five face-sharing tetrahedra, albeit with five small gaps of 1.47° each (fig. S8A) (*34*). The exposed faces of the tetrahedra still have the same complementarity to octahedra; each face can bind with a neighboring octahedron. When grouped into a decahedron, this shape-complementary arrangement can be viewed as the local packing of 10 octahedra around one decahedron (Fig. 4J). Such local complementarity can only periodically extend itself along the polar direction, forming an infinitely long pillar with aligning decahedral polar tips (Fig. 6A). Two pillars can be strongly connected equatorially by one row of octahedra that are shared between the decahedra from both pillars (Fig. 4K). By extending the connection between pillars, the pillars can be arranged into a superlattice (Fig. 6A). This proposed superlattice has a similar tessellation rule as the Penrose P1 tiling and can therefore be considered an infinitely repeated 2D pattern (figs. S8, D and E, S34, and S35) (*35*). We refer to this superlattice as the P1 superlattice.

To test our hypothesized geometric design, we synthesized decahedral and octahedral NCs with identical edge lengths (E between 60 and 100 nm) and assembled them using sp0, sp1, sp2, sp3, and sp5 DNA linkers. Because decahedral and octahedral NCs form quasi–space-filling pairs (fig. S8), there was a small oblique gap between neighboring octahedra within pillars. Without sufficient linker flexibility, this gap would cause a structural strain inside the 3D superlattices; vertically, the oblique gaps between decahedra and octahedra would be perfectly parallel (Fig. 6D and fig S35), but horizontally, these gaps (1.47°) would hinder the 2D tessellation of pentagons (Fig. 6E and fig S34). Indeed, when NCs with E = 85 nm were assembled with sp0 linkers, icosahedral clusters in which local strains were radially distributed around all 12 decahedra were observed exclusively (Fig. 6, B and C) instead of the proposed P1 superlattice.

This strain was relieved by increasing the length of the OEG flexor region: sp1 linkers enabled very small domains (Fig. 6F), and sp2 linkers enabled needle-shaped superlattices as long as 20 μm and less than 5 μm wide (Fig. 6G) but with a substantial number of defects that limited continued growth. NCs with even longer flexor regions (sp3 and sp5) assembled into smaller aspect ratio superlattices with much wider cross sections and relatively few structural defects (Fig. 6, H and I). Notably, we obtained superlattices as long as 100 μm, with cross sections as wide as 50 μm, using sp5 linkers (Fig. 6, J and K). The nature of the 2D tessellation in these samples was characterized by sectioning (Fig. 6, L and M). The superlattices appeared in cross sections as a nonperiodic tessellation of pentagons, with well-resolved multiple layers of 10-fold fast Fourier transform (FFT) rings [Fig. 6L (inset) and fig. S34].

The DNA-mediated assembly of the P1 superlattice was further confirmed by MD simulations. We simulated binary mixtures of decahedral and octahedral NCs with 1:4 stoichiometric composition and observed the assembly of a P1 superlattice with a periodic stacking of layers (Fig. 5D). A cross section of the P1 superlattice viewed along the stacking direction (Fig. 5E) showed local order that agreed with the geometric design (Fig. 6A). From a large single domain of a P1 superlattice assembled in simulation from a seed (supplementary materials), we found that the P1 superlattice could be mapped to a nonperiodic tessellation of pentagons (fig. S28E).

The diffraction patterns of the P1 superlattices viewed along the stacking direction showed a decagonal symmetry, in agreement with the experimental FFTs (Fig. 5F and fig. S28, D to F). Binary mixtures of hard decahedra and hard octahedra did not assemble into P1 superlattices but instead assembled into periodically ordered crystals with twofold and sixfold rotational symmetries (fig. S27, D to F). This result demonstrated that the flexible DNA graft could alter the assembly of this particular quasi–space-filling pair, unlike other space-tilling polyhedron pairs. Taken together, the assembly of octahedra and coassembly of decahedron-octahedron mixtures showed that we could use flexible segments in DNA ligands as an important additional design handle for tuning the arrangements of mixtures of quasi–space-filling polyhedra.

Conclusions

The ability to engineer large, space-filling colloidal crystals from both space-filling and non–space-

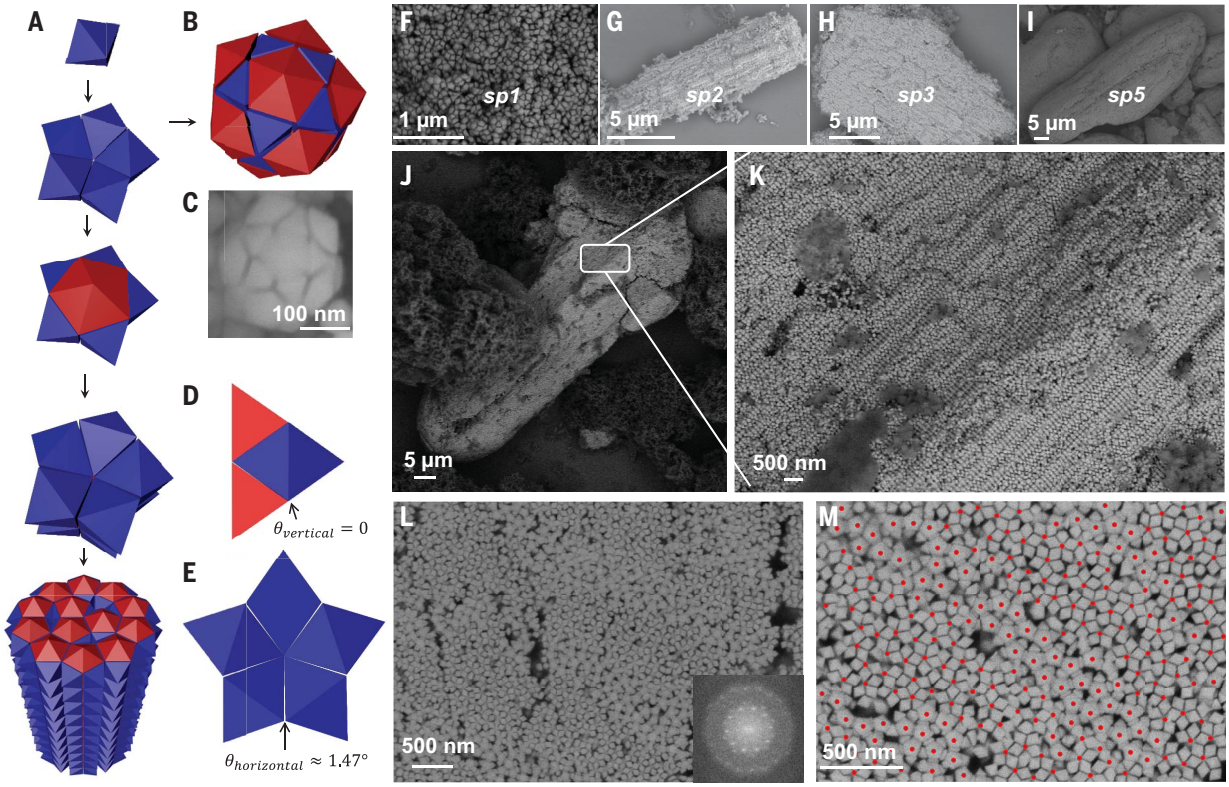


Fig. 6. Construction of P1 superlattices. (**A**) Schematic illustration of P1 lattice formed from decahedral and octahedral NCs. (**B** and **C**) Schematic illustration (B) and SEM image (C) of icosahedral clusters formed from decahedral and octahedral NCs. (**D** and **E**) Schematic illustration of gaps between decahedra and octahedra (D), as well as between octahedra (E). (**F** to **M**) Representative SEM images of colloidal crystals formed from decahedral and octahedral NCs (both E = 85 nm), assembled by sp1 (F), sp2 (G), sp3 (H), and sp5 [(I) to (M)] linkers, where [(G) to (K)] shows the sideview, and [(L) and (M)] shows the cross section of the P1 superlattice. The arrangements of decahedral NCs, whose centers are represented by red dots in (M), can be mapped to nonperiodic tessellation of pentagons.

filling NCs with simple geometric considerations unleashes the opportunity to explore an enormous library of structures with DNA. The 10 assemblies presented in this work represent a small fraction of the total potential design space. Considering the recent interest in dense, highly ordered plasmonic superlattices in the field of optical metamaterials (*36*, *37*), this polyhedron construction strategy will likely lead to crystalline architectures with unnatural chemical and physical properties. More broadly, this strategy for realizing geometry-dictated organization into space-filing structures is easily translatable to other self-assembling systems.

REFERENCES AND NOTES

1. D. Samanta, W. Zhou, S. B. Ebrahimi, S. H. Petrosko, C. A. Mirkin, *Adv. Mater.* **34**, e2107875 (2022).
2. K. Deng, Z. Luo, L. Tan, Z. Quan, *Chem. Soc. Rev.* **49**, 6002–6038 (2020).
3. Y. Li *et al.*, *Nature* **611**, 695–701 (2022).
4. S. Zhou *et al.*, *Nature* **612**, 259–265 (2022).
5. Y. Li *et al.*, *Small* **18**, e2201171 (2022).
6. C. A. Mirkin, R. L. Letsinger, R. C. Mucic, J. J. Storhoff, *Nature* **382**, 607–609 (1996).
7. R. J. Macfarlane *et al.*, *Science* **334**, 204–208 (2011).
8. M. R. Jones *et al.*, *Nat. Mater.* **9**, 913–917 (2010).
9. M. N. O'Brien, M. R. Jones, B. Lee, C. A. Mirkin, *Nat. Mater.* **14**, 833–839 (2015).
10. C. R. Laramy *et al.*, *ACS Nano* **13**, 1412–1420 (2019).
11. W. Liu *et al.*, *Science* **351**, 582–586 (2016).
12. M. He *et al.*, *Nature* **585**, 524–529 (2020).
13. W. B. Rogers, V. N. Manoharan, *Science* **347**, 639–642 (2015).
14. F. Lu, K. G. Yager, Y. Zhang, H. Xin, O. Gang, *Nat. Commun.* **6**, 6912 (2015).
15. F. Lu *et al.*, *Sci. Adv.* **5**, eaaw2399 (2019).
16. M. N. O'Brien *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10485–10490 (2016).
17. Y. Zhang, F. Lu, D. van der Lelie, O. Gang, *Phys. Rev. Lett.* **107**, 135701 (2011).
18. H. Lin *et al.*, *Science* **355**, 931–935 (2017).
19. M. N. O'Brien, H. X. Lin, M. Girard, M. Olvera de la Cruz, C. A. Mirkin, *J. Am. Chem. Soc.* **138**, 14562–14565 (2016).
20. P. Engel, *Geometric Crystallography: An Axiomatic Introduction to Crystallography* (Springer Netherlands, vol. 23, 1986), doi:10.1007/978-94-009-4760-3.
21. P. F. Damasceno, M. Engel, S. C. Glotzer, *ACS Nano* **6**, 609–614 (2012).
22. J. Henzie, M. Grünwald, A. Widmer-Cooper, P. L. Geissler, P. Yang, *Nat. Mater.* **11**, 131–137 (2011).
23. J. Gong *et al.*, *Nat. Commun.* **8**, 14038 (2017).
24. S. Chaladus, *Math. Comput.* **58**, 341–345 (1992).
25. K. C. Elbert *et al.*, *ACS Nano* **13**, 14241–14251 (2019).
26. A. Santos, J. A. Millan, S. C. Glotzer, *Nanoscale* **4**, 2640–2650 (2012).
27. M. S. Wertheim, *J. Stat. Phys.* **42**, 459–476 (1986).
28. V. Ramasubramani, T. Vo, J. A. Anderson, S. C. Glotzer, *J. Chem. Phys.* **153**, 084106 (2020).
29. J. Towns *et al.*, *Comput. Sci. Eng.* **16**, 62–74 (2014).
30. J. A. Anderson, J. Glaser, S. C. Glotzer, *Comput. Mater. Sci.* **173**, 109363 (2020).
31. F. A. Escobedo, *J. Chem. Phys.* **145**, 211903 (2016).
32. M. R. Khadilkar, F. A. Escobedo, *J. Chem. Phys.* **137**, 194907 (2012).
33. A. T. Cadotte, J. Dshemuchadse, P. F. Damasceno, R. S. Newman, S. C. Glotzer, *Soft Matter* **12**, 7073–7078 (2016).
34. J. S. Du, W. Zhou, S. M. Rupich, C. A. Mirkin, *Angew. Chem. Int. Ed.* **60**, 6858–6863 (2021).
35. R. Penrose, *Bull. Inst. Math. Appl.* **10**, 266–271 (1974).
36. S. Kim *et al.*, *Nano Lett.* **20**, 8096–8101 (2020).
37. J. H. Huh, J. Lee, S. Lee, *Nano Lett.* **20**, 4768–4774 (2020).

ACKNOWLEDGMENTS

The authors thank J. S. Du and S. Wang [Northwestern University (NU)] for discussions, E. W. Roth (NU) for help with ultramicrotomy, and S. H. Petrosko (NU) for editing the manuscript. **Funding:** This material is based upon work supported by the Air Force Office of Scientific Research award FA9550-22-1-0300 (nanoparticle synthesis and assembly) and the Center for Bio-Inspired Energy Science, an Energy Frontier Research Center funded by the US Department of Energy, Office of Science, Basic Energy Sciences award DE-SC0000989 (EM characterization, theory, modeling, and simulation). Z.H. acknowledges support by the NU Graduate School Cluster in Biotechnology, Systems, and Synthetic Biology, which is affiliated with the Biotechnology Training Program funded by NIGMS grant T32 GM008449. This work made use of the EPIC facility of the NU*ANCE* Center at NU, which has received support from the Soft and Hybrid Nanotechnology Experimental (SHyNE) Resource (NSF NNCI-1542205), the International Institute for Nanotechnology (IIN), the Keck Foundation, and the State of Illinois, through the IIN. T.V. was supported by a grant from the Simons Foundation (256297, SCG). K.J. acknowledges support from the University of Michigan Rackham Predoctoral Fellowship Program. Simulations made use of the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant no. ACI-1548562; XSEDE award DMR 140129. Computational resources and services were also provided by Advanced Research Computing, a division of Information and Technology Services at the University of Michigan, Ann Arbor. **Author contributions:** W.Z. and Y.L. synthesized the particles and designed and assembled the colloidal crystals. Z.H.

and Y.L. performed the synthesis of the oligonucleotides. W.Z., Y.L., and H.L. performed electron microscopy studies. K.J. performed molecular dynamics simulations, and T.V. performed lattice free energy calculations. C.A.M. and S.C.G. supervised the project. All authors contributed to the data analysis and writing of the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. **License information:**

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adj1021
Materials and Methods
Supplementary Text
Figs. S1 to S36
Tables S1 to S15
References (*38*–*46*)
Movie S1

Submitted 6 June 2023; resubmitted 25 September 2023
Accepted 8 December 2023
10.1126/science.adj1021

IMMUNOLOGY

Impact of HLA class I functional divergence on HIV control

Mathias Viard[1]†, Colm O'hUigin[1]†, Yuko Yuki[1], Arman A. Bashirova[1], David R. Collins[2], Jonathan M. Urbach[2], Steven Wolinsky[3], Susan Buchbinder[4,5], Gregory D. Kirk[6], James J. Goedert[7], Nelson L. Michael[8], David W. Haas[9], Steven G. Deeks[10], Bruce D. Walker[2,11], Xu Yu[2], Mary Carrington[1,2]*

Heterozygosity of *Human leukocyte antigen* (*HLA*) class I genes is linked to beneficial outcomes after HIV infection, presumably through greater breadth of HIV epitope presentation and cytotoxic T cell response. Distinct allotype pairs, however, differ in the extent to which they bind shared sets of peptides. We developed a functional divergence metric that measures pairwise complementarity of allotype-associated peptide binding profiles. Greater functional divergence for pairs of HLA-A and/or HLA-B allotypes was associated with slower AIDS progression and independently with enhanced viral load control. The metric predicts immune breadth at the peptide level rather than gene level and redefines *HLA* heterozygosity as a continuum differentially affecting disease outcome. Functional divergence may affect response to additional infections, vaccination, immunotherapy, and other diseases where *HLA* heterozygote advantage occurs.

It has been over two decades since a robust demonstration of *Human leukocyte antigen* (*HLA*) class I heterozygous advantage was documented for people living with HIV (PLWH) (*1*), in which progression to AIDS after seroconversion was lengthened in proportion to the number of heterozygous class I gene loci. The blunt assignment of heterozygosity, however, does not account for the degree of functional similarity between pairs of allotypes encoded by a given *HLA* genotype. A larger repertoire of HIV peptides presented by two allotypes that bind very distinct types of peptides may improve immune surveillance relative to a smaller repertoire contributed by functionally similar allotypes. In support of this model, *HLA* sequence divergence caused by radical substitutions has been shown to be selectively favored (*2*), and indeed, has been proposed as a source of fitness differences among heterozygous individuals (*3*), potentially accounting for maintenance of allelic lineages with up to 10% sequence difference. Metrics estimating genetic distance between *HLA* allelic pairs of a given genotype as a proxy for distinctiveness of peptide repertoires (*4*) have been used as an outcome predictor in infectious disease (*5*), transplantation (*6*, *7*), and certain cancer treatment studies (*8*). However, a more accurate measurement of the impact of a larger peptide repertoire on disease risk or outcome would employ metrics that are directly peptide-based.

We describe a novel metric to measure the proportion of overlap between HLA allotypic pairs based on empirically determined peptide repertoires for each allotype, which we term "functional divergence". Analyses of functional divergence in a large cohort of well-defined untreated HIV-1 seroconverters in which heterozygous advantage was previously observed shows that greater functional divergence between pairs of allotypes at *HLA-A* and/or *HLA-B* correspondingly associates with slower disease progression. The association was far stronger than either the dichotomous assignment of genetic zygosity or metrics of allelic genetic distances. The impact of functional divergence on HIV disease was robustly and consistently reproduced in an independent cohort consisting of HIV-1 elite controllers and noncontrollers, as well as in a cohort of >3500 individuals where longitudinal viral loads were available for analysis. The peptide-centric algorithm described herein holds an advantage over genetic-based approaches and demonstrates the sizeable impact of functional divergence at *HLA* class I on outcome to HIV-1 infection.

Effect of allelic homozygosity on HIV disease progression

Our analysis focused initially on a cohort of up to 870 seroconverters, 60% of whom were included in the prior study that established the protective effects of *HLA* class I heterozygosity (*1*). Reanalysis of the enlarged cohort strengthens the observation that homozygosity at *HLA* class I results in faster progression to AIDS, as indicated by progression to CD4 counts below 200 cells per ml (CD4 < 200) [hazard ratio (HR) = 1.53 and *P*-value = 6×10^{-4}] (Fig. 1).

Effects of individual *HLA* class I loci are strongest for homozygosity at *HLA-B* (HR = 1.82, *P* = 0.004), intermediate for *HLA-A* (HR = 1.56, *P* = 0.003), and lowest for *HLA-C* (HR = 1.43, *P* = 0.036). The tight linkage disequilibrium between the *HLA-B* and *-C* loci leads to difficulty in resolving effects for each locus independently, but neither homozygosity at *HLA-C* nor *HLA-B* alone showed a significant association with progression to CD4 < 200 (table S1), the latter trending in the same direction (HR = 1.36) in a particularly small number of individuals (N = 15).

Effect of genetic distance on HIV progression

Dichotomously categorizing individuals as genetic homozygotes versus heterozygotes does not account for distinctions between the two alleles (or their products) present at each heterozygous locus. Interallelic genetic distances at *HLA* class I loci [*HLA* evolutionary divergence (HED)] have been used as a proxy for differential peptide binding across pairs of allotypes encoded by the corresponding genotype (*4*). A negative correlation between *HLA-B* genetic distances and HIV viral load (VL) was observed in a large cohort of PLWH but the observed correlation was low (Kendall t coefficient < 0.08) (*5*). No correlation was observed between HIV VL and *HLA-A* genetic distance and a positive correlation was found with *HLA-C* distances (*5*). We find no effect of the HED genetic distance metric for *HLA-A*, *-B*, or *-C* individually or combined on progression to low CD4 counts in our seroconverter cohort in analyses including or excluding genetic homozygotes (fig. S1). Analyses of other survival outcomes, including time to an AIDS-defining illness or death after seroconversion, also failed to show significant effects of HED. These data suggest that the relatively weak effects of HED on HIV VL cannot be mirrored in analyses of HIV disease progression.

Impact of HLA class I supertypes on HIV progression

Global analyses of various peptide pools eluted from class I allotypes or binding assays of peptides exogenously-bound to specific HLA allotypes have identified broad peptide motifs associating with groups of *HLA-A* and *HLA-B* alleles, known as supertypes (*9*). Similarity of preferred amino acids at position 2 or the C terminus of the peptide, the residues anchoring the peptide into the groove of the class I molecule, constitutes the primary basis for broadly grouping the alleles into 6 supertypes for *HLA-A* and

[1]Basic Science Program, Frederick National Laboratory for Cancer Research, National Cancer Institute, Frederick, MD, USA and Laboratory of Integrative Cancer Immunology, Center for Cancer Research, National Cancer Institute Bethesda, MD, USA. [2]Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA. [3]Division of Infectious Diseases, Department of Medicine, The Feinberg School of Medicine, Northwestern University, Chicago, IL, USA. [4]Bridge HIV, San Francisco Department of Public Health, San Francisco, CA, USA. [5]Department of Medicine, Epidemiology and Biostatistics, University of California, San Francisco, CA, USA. [6]Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA. [7]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. [8]US Military HIV Research Program, Walter Reed Army Institute of Research, Silver Spring, MD, USA. [9]Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN, USA. [10]Department of Medicine, University of California, San Francisco, CA, USA. [11]Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA.
***Corresponding author. Email: carringm@mail.nih.gov**
†These authors contributed equally to this work.

6 supertypes for *HLA-B* (*9*). As the supertypes are based on peptide binding data (of various types and sources) rather than genetic distances of allotypic pairs, we tested for an effect of supertype zygosity on progression to CD4 < 200 but observed no significant effect in a comparison of heterozygotes (across supertypes) versus homozygotes (within supertypes) at *HLA-A* (HR = 1.09, $P = 0.47$) or *HLA-B* (HR = 1.21, $P = 0.14$) on progression to CD4 < 200, even though these analyses included *HLA* allelic homozygotes. We next generated further subdivisions of HLA-A and HLA-B supertypes based on physico-chemical properties (size, charge, and hydrophobicity) of the motif-defining amino acids, which resulted in 19 distinct supertypes for HLA-B and 10 for HLA-A (see Materials and Methods). The increased specificity of the novel supertype definition indicated significant heterozygosity advantage across supertypes at both *HLA-A* (HR = 1.30, $P = 0.048$) and *HLA-B* (HR = 1.76, $P = 0.0028$). This significance is not sustained, however, upon the exclusion of allelic homozygotes at *HLA-A* (HR = 0.88, $P = 0.60$) or *HLA-B* (HR = 1.52, $P = 0.32$). Overall, these data suggest that accounting for greater scrutiny of peptide properties may be necessary in order to detect differential effects of specific pairs of HLA allotypes on HIV disease outcomes.

Determination of functional divergence through submotif overlap

A more in-depth analysis of complementarity between pairs of allotypes encoded by genetic heterozygotes at *HLA* class I was attempted using the well-standardized HLAthena dataset of peptides eluted from class I allotypes that were individually expressed on the otherwise *HLA* class I negative .221 cell line (*10*). Eluted peptides derived from each of 95 HLA allotypes ranged in length from 7 to 42 amino acids and peptide repertoire sizes varied from 692 to 4033 per allotype. To simplify analyses, we focused solely on 9-mer peptides, which account for about 60% of the characterized HLA class I bound peptides. Sarkizova *et al.* (*10*) performed a sequence-based clustering analysis of peptides eluted from each class I allotype, which we used to derive a specific logogram for each identified cluster (Fig. 2A). This approach allowed us to focus on aggregate properties of the peptides (i.e., those defining a cluster) presented by a given allotype rather than minor differences across peptides (i.e., variants within a cluster).

Only those amino acids with a positional entropy lower than 0.1 (i.e., high selectivity for one amino acid at a given position) were retained in the analysis in order to reduce noise contributed at positions with high entropy (Fig. 2A). The resulting simplified logograms, termed submotifs, allow compact representation of the different families of peptides that a given allotype can present. This approach condensed a universe of 111,898 nonamer peptides presented by 95 different HLA class I allotypes into 382 distinct submotifs, with a range of 2 to 33 9-mer clusters per HLA allotype. We compared submotifs of the allotypic pairs encoded by a given genotype at each class I locus and calculated the proportion of submotifs that differ between that pair of allotypes, allowing quantification of the functional divergence and redundancy between pairs of allotypes (Fig. 2B). Peptides behaving as singletons (shown in black in Fig. 2A) were included as a separate subgroup in order to avoid a complete lack of overlap between any given allotypic pair. By including this subgroup of singletons, the metric accounts for the total number of submotifs associated with each pair of allotypes.

Individuals in the cohort were then ranked by the extent of submotif redundancy between pairs of allotypes encoded at each of their *HLA* loci, and disease outcome was compared between the top- and bottom-ranked groupings of individuals (defined by appropriate quantiles) for each locus (Fig. 2C and Materials and Methods). The comparison of individuals from top and bottom quantiles is advantageous for delineating effects of functional divergence from noise. Functional divergence and its effects are continuous, varying according to the allotype combinations involved and not by effects of individual allotypes. That is, a single allotype may exist at both extremes of the functional divergence spectrum depending on the alternative allotype with which it is paired. Effects of intermediate levels of functional divergence can be masked by numerous other genetic, viral, or environmental factors that affect HIV outcome. Consequently, in order to determine whether functional divergence has an impact on disease outcome, we initially compared groups that differ substantially in divergence. Correction for the well-documented (including the cohorts used in this study) effects on progression of *HLA-B*57*, *-B*27*, certain *-B*35* alleles (*11*), as well as for racial background, was consistently applied in all analyses described.

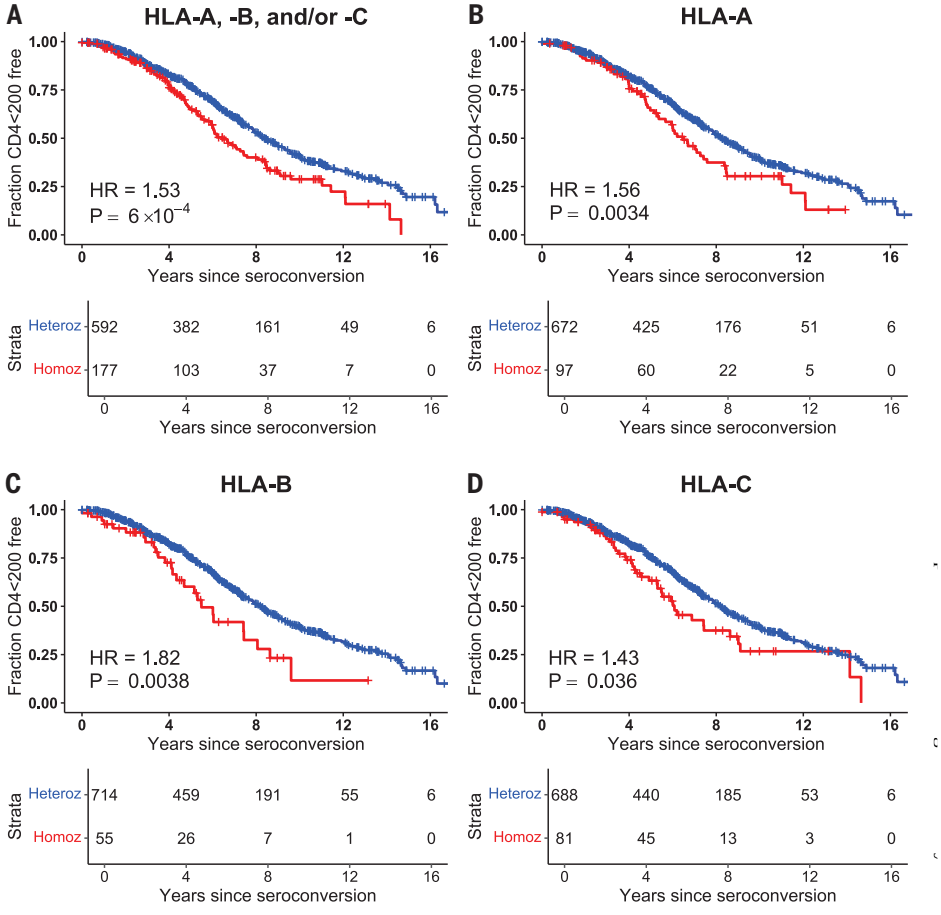
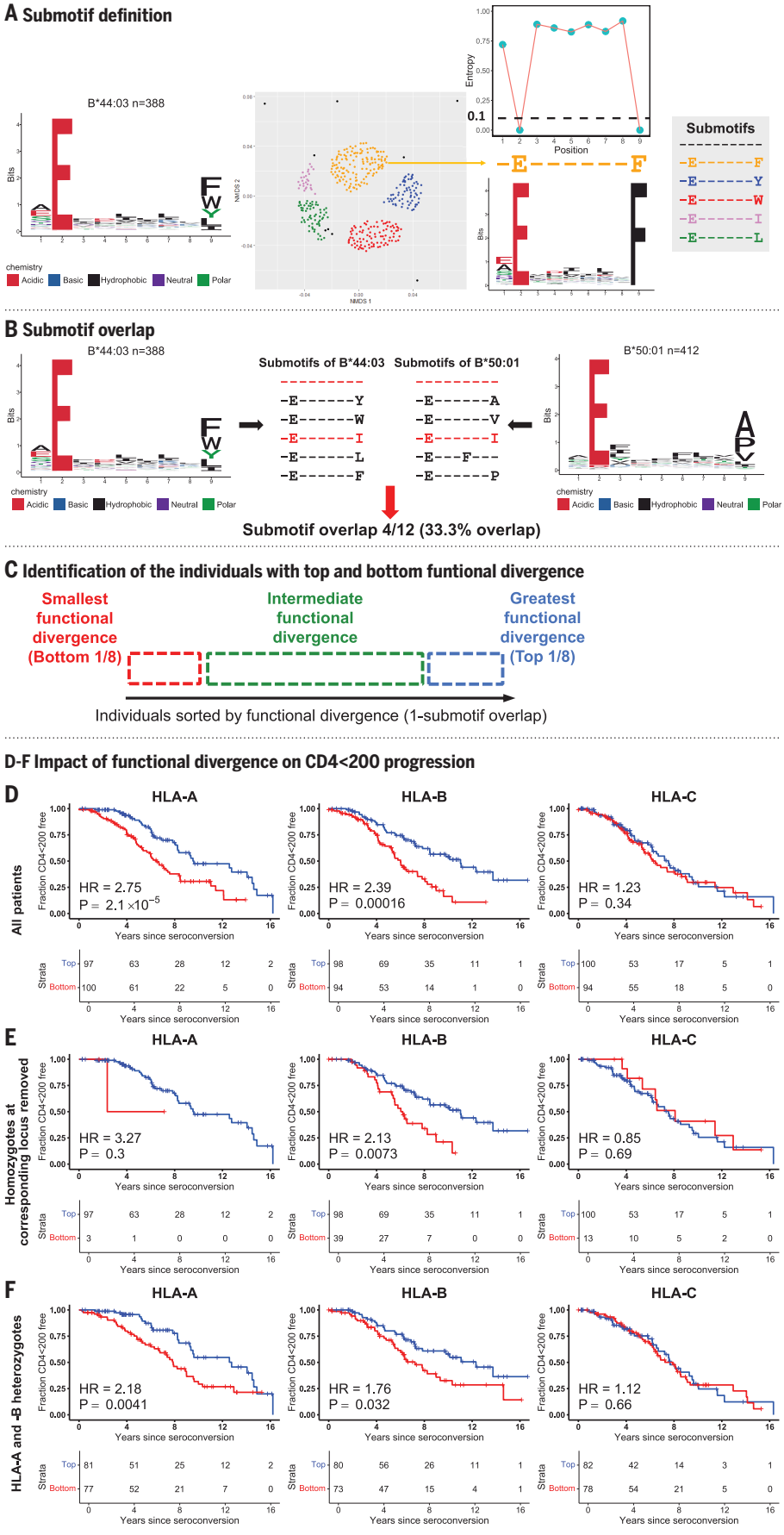


Fig. 1. Effect of genetic homozygosity on progression to CD4 < 200 after HIV infection. Progression to CD4 < 200 among (**A**) individuals homozygous at any *HLA-A*, *-B*, and/or *-C* locus/loci (red curve) compared with heterozygotes at all three loci (blue curve). (**B**) *HLA-A* homozygotes (red) compared with *HLA-A* heterozygotes (blue). (**C**) *HLA-B* homozygotes (red) compared with *HLA-B* heterozygotes (blue). (**D**) *HLA-C* homozygotes (red) compared with *HLA-C* heterozygotes (blue). All analyses were corrected for *HLA-B*57*, *B*27*, *B*35Px*, and racial background.

Fig. 2. Submotif overlap and its effect on progression to CD4 < 200 after HIV infection. (**A**) Representative motif logogram and nonmetric ordination plot as generated for each of 95 HLA allotypes studied previously (*10*). Submotifs corresponding to each cluster were determined by retaining only the amino acids at each position with an entropy lower then 0.1 (shown for B*44:03). (**B**) Determination of submotif overlap corresponding to the two HLA allotypes encoded by each genotype. Red indicates identical whereas black indicates distinct submotifs across the two allotypes. (**C**) Determination of the range in functional divergence from smallest to greatest at each *HLA* locus across individuals. Comparisons were made between individuals falling into the top-ranked versus bottom-ranked quantiles (shown in octiles). (**D** to **F**) Progression to CD4 < 200 was compared between individuals in the bottom octile (red curves) versus top octile (blue curves) of functional divergence at each HLA locus (D) among all individuals, (E) after removing genetic homozygotes corresponding to each individual locus, but keeping the same octile definition as in (D) (diminishing the number of individuals in the bottom octile), or (F) after removing genetic homozygotes at *HLA-A* and/or *HLA-B* prior to selection of the bottom and top octiles. Analyses were corrected for *HLA-B*57*, *B*27*, *B*35Px*, and racial background.



Effect of functional divergence on HIV disease progression

Analyses measuring the impact of functional divergence on HIV disease outcome were performed comparing extreme octiles as this quantile size was determined to represent a balance between the magnitude and significance of resultant hazard ratios (Materials and Methods, table S2). Low functional divergence (i.e., high submotif redundancy) associated with faster progression to low CD4 counts (Fig. 2D), where significant differences between the top versus bottom ranked octiles for HLA-A (HR = 2.75, $P = 2.1 \times 10^{-5}$) and HLA-B (HR = 2.39, $P = 1.5 \times 10^{-4}$) were observed. The lack of a significant effect of functional divergence for HLA-C (HR = 1.23, $P = 0.34$) was consistent with prior findings of reduced association between genetic homozygosity at *HLA-C* (Fig. 1) and disease progression, which likely reflected homozygosity at the closely linked *HLA-B* locus.

To eliminate the possibility that genetic homozygosity is predominantly responsible for the more rapid disease progression among individuals in the bottom octile (red curves in Fig. 2D), individuals homozygous at each respective locus were removed from the top and bottom octiles defined in Fig. 2D, diminishing the power of the analysis (Fig. 2E). Only three *HLA-A* genetic heterozygotes remained in the bottom functional divergence octile after removal of *HLA-A* homozygotes, resulting in a lack of significance (HR = 3.27, $P = 0.30$), but both the trend and significance were well conserved for HLA-B (HR = 2.13, $P = 0.0073$) after exclusion of *HLA-B* genetic homozygotes. Once again, functional divergence at HLA-C had no significant association with HIV disease progression (HR = 0.85, $P = 0.69$) after removal of *HLA-C* genetic homozygotes. Genetic homozygotes at *HLA-C* were not removed from further analyses given the apparent lack of an effect of HLA-C functional divergence on HIV disease progression. Analyses in which the top and bottom octiles were fully repopulated after exclusion of homozygotes at *HLA-A* and/or *HLA-B* (Fig. 2F) restored numbers and power, and significant protective effects of greater functional divergence were observed for HLA-A (HR = 2.18, $P = 0.0041$) and HLA-B (HR = 1.76, $P = 0.032$), but not for HLA-C (HR = 1.12, $P = 0.66$). We conclude that greater functional divergence at both HLA-A and HLA-B independently confer substantial protection against HIV disease progression.

Combination of HLA-A and HLA-B functional divergence on HIV disease progression

The linked *HLA-A* and *-B* loci are separated by roughly 1Mb and sufficient levels of recombination between the two loci has occurred (*12*) such that independent effects of each locus on HIV disease can be detected. Figure 2, D to F, indicated that functional divergence levels at both HLA-A and -B individually have an impact on HIV disease progression, implying that their combined effects might be more influential than considering each separately. We tested the combined effects of functional divergence at HLA-A and -B by comparing individuals with the greatest functional divergence (top octile) to those with the least functional divergence (bottom octile) at HLA-A and/or HLA-B based on genotypes for each person at these loci. Besides being biologically relevant, accounting for combined effects of these two loci enhances analytical power as a result of the enlarged sizes of the two comparison groups. In the first set of analyses, individuals having HLA-A and -B allotypic pairs in which one ranked in the top octile and the other ranked in the bottom octile were excluded. HLA-A and HLA-B combined functional divergence based on the following groups were then compared: the combined grouping of those whose HLA-A divergence was in the top octile and HLA-B was intermediate; plus those whose HLA-B divergence was in the top octile and HLA-A was intermediate; plus those whose HLA-A and HLA-B divergences were both in the top octile versus the combined grouping of those whose HLA-A divergence was in the bottom octile and HLA-B was intermediate; plus those whose HLA-B divergence was in the bottom octile and HLA-A was intermediate; plus those whose HLA-A and HLA-B divergences were both in the bottom octile. A particularly robust protective effect of greater functional divergence was observed when considering the combination of both loci (HR = 2.87, $P = 2.4 \times 10^{-8}$) (Fig. 3A), as opposed to HLA-A or HLA-B individually (Fig. 2D) and is among the strongest documented HLA effects on HIV disease progression. A second approach in which HLA-A and HLA-B functional divergence values were first summed followed by selection of the top and bottom octiles of divergence (as opposed to first selecting individuals from each of the extreme octiles for HLA-A and HLA-B separately and then combining them into the respective top/bottom divergence groupings) produced similar results but with less statistical support (fig. S2).

The impact of functional divergence on AIDS progression was strongest and most significant for early outcomes, including time to CD4 < 200 and time to AIDS93 (defined by the CDC as CD4 < 200 or an AIDS-defining illness, whichever is diagnosed first) (*13*), but showed less significant effects for the later outcomes, AIDS87 (AIDS-defining illness) (*14*), or death (table S3). These data reflect the same pattern observed for genetic zygosity. Notably, the effect size (HR) was always greater for functional divergence than for genetic zygosity across each outcome (table S3).

To ascertain whether the protective effect of greater HLA functional divergence between pairs of HLA allotypes tends to be continuous, we subdivided the cohort into further categories. Inclusion of individuals with intermediate functional divergence scores at both HLA-A and HLA-B indicates that this grouping progresses to loss of CD4 counts at a rate intermediate to the two extreme quantiles, as expected (mean HR with each stepwise category = 1.57, $P_{trend} = 2.2 \times 10^{-7}$; Fig. 3B). Further subdivisions splitting both the top and bottom groupings according to whether only one or both HLA-A and HLA-B were responsible for their respective rankings resulted in a total of 5 distinct categories. These groups consisted of individuals ranking as follows: (A) the top group (most divergent) at both HLA-A and HLA-B, (B) the top group at only one and intermediate at the other, (C) intermediate at both; (D) the bottom group (least divergent) at only one and intermediate at the other, and (E) the bottom group at both. These progressive rankings revealed a perfectly continuous effect of functional divergence on HIV disease progression (mean HR with each decreasing stepwise category = 1.49, $P_{trend} = 7.6 \times 10^{-8}$, Fig. 3C), despite low numbers in each of the two extreme groupings (i.e., greatest divergence at both HLA-A and HLA-B and least divergence at both HLA-A and HLA-B). Individuals having HLA-A and -B genotypes in which one ranked in the top octile and the other ranked in the bottom octile exhibit a mixed phenotype whereby upon an initial slow progression to CD4 < 200, subsequently, the rate of progression substantially increased (purple curve, fig. S3).

Significance of disease associations with level of functional divergence was maintained after removal of genetic homozygotes at HLA-A and -B from each of the respective analyses (Fig. 3, D to F), illustrating the impact of this property beyond genetic zygosity status alone. The strength of the effects upon exclusion of genetic homozygotes was equal to or slightly stronger than those when genetic homozygotes were included in the analyses (compare HR of Fig. 3A, B, and C to 3D, E, and F, respectively).

Effect of functional divergence on HIV elite controller status and mean VL in untreated individuals

An independent replication cohort consisting of HIV elite controllers (those who maintain undetectable VL for at least one year; $n = 308$) and noncontrollers (those with viral VLs >10,000 copies per ml; $n = 1895$) was used to confirm findings observed in the seroconverter cohort. Functional divergence based on the five progressive categories shown in Fig. 3, C and F, was strongly predictive of outcomes in comparisons of elite controllers versus noncontrollers, either in the presence [mean odds ratio with each stepwise category (OR_{trend}) = 1.53, $P = 6.6 \times 10^{-6}$ or absence ($OR_{trend} = 1.43$, $P = 6.3 \times 10^{-6}$) of genetic homozygotes (Table 1). VL provides an additional means to assess functional divergence

by measuring the strength and significance of the regression estimate (change in mean VL per unit increase in functional divergence). An analysis of the five progressive categories assigned to 3563 individuals with longitudinal VL measurements—2156 of whom overlapped with the elite controller versus noncontroller cohort—showed that greater functional divergence significantly associated with reduction in VLs in both the presence (estimate $P = 2.5 \times 10^{-7}$) or absence (estimate $P = 4.2 \times 10^{-4}$) of homozygotes at *HLA-A* and/or *HLA-B* loci (Table 1). These data soundly confirm the substantial impact of functional divergence on HIV control.

Functional divergence based only on submotifs to which HIV peptides correspond

The 382 submotifs used in our analyses were derived from self peptides eluted from .221 cells expressing single allotypes of HLA class I (*10*), and some of these submotifs may not be relevant when considering HIV-derived peptides. We took several approaches to test this possibility, stringently restricting the analyses to individuals fully heterozygous at both HLA-A and -B. We tested whether functional divergence based only on the set of submotifs corresponding to the HIV peptidome derived from clade B and clade C HIV 2004 consensus sequences

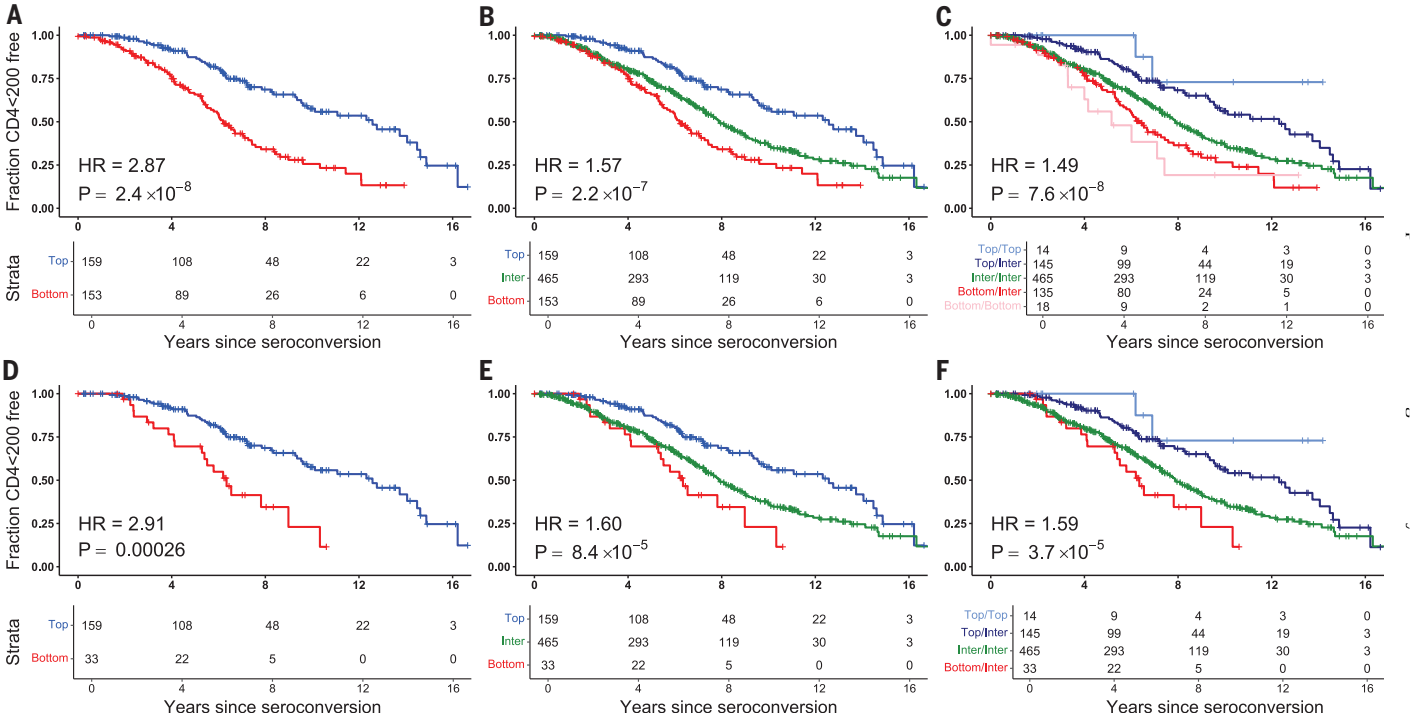


Fig. 3. Combined effect of HLA-A and HLA-B submotif overlap on progression to CD4 < 200. (**A**) Progression to CD4 < 200 among individuals in the bottom (red curve) versus top (blue curve) octile(s) of functional divergence at HLA-A and/or HLA-B. For individuals in whom only one of the two (HLA-A or HLA-B) was in the top octile, the other was necessarily in the intermediate range. (**B**) Individuals carrying intermediate levels of functional divergence at both *HLA-A* and *HLA-B* (green curve) were added to the analysis shown in (A). *P* = test for trend. (**C**) Individuals were further stratified as follows: bottom octiles at both HLA-A and HLA-B (pink curve); bottom octile at HLA-A or HLA-B and intermediate at the other (red curve); intermediate at both (green curve); bottom octile for only one and intermediate at the other (dark blue curve); top octiles at both (light blue curve). *P* = test for trend. The HR represents an average of the effect with each decreasing category. The analyses were performed among all individuals (A to C) or following removal of homozygotes at *HLA-A* and *HLA-B* (i.e., heterozygous at both *HLA-A* and *HLA-B*) (**D** to **F**). Analyses were corrected for *HLA-B*57*, *B*27*, *B*35Px*, and racial background.

Table 1. Effect of HLA-A and/or HLA-B submotif overlap on controller status and mean VL after HIV infection. Logistic regression of controller status (elite controller versus noncontroller) and linear regression of mean VL were used to estimate the effect of functional divergence at HLA-A and/or HLA-B structured in 5 categories as defined for Fig. 3C.

Study	Study Participants	Outcome	Modeling	N	Effect estimate†	*P*
Elite controllers (301) versus Noncontrollers (1844)	All	Controller status	Logistic regression	2145	1.53	6.6×10^{-6}
	HLA-A and *-B* genetic heterozygotes only			1759	1.43	6.3×10^{-4}
Individuals with mean VL (3462)	All	Mean VL	Linear regression	3462	0.14	2.5×10^{-7}
	HLA-A and *-B* genetic heterozygotes only			2865	0.11	4.2×10^{-4}

†The effect estimate corresponds to the OR_{trend} for the logistic regression and the regression coefficient for the linear regression. Analyses were adjusted for *HLA-B*57*, *B*27*, *B*35Px*, and racial background.

present in the Los Alamos National Laboratory (LANL) database (www.hiv.lanl.gov/content/sequence/HIV/CONSENSUS/Consensus.html) would produce similar results to those observed when the full set of 382 submotifs were employed. All possible overlapping nonamers derived from HIV consensus sequences (n = 5042 from the two clades combined) represented 289 of the 382 submotifs derived from self. Thus, 93 submotifs to which none of the HIV nonamer peptides corresponded were excluded from the analysis. Calculation of functional divergence using only the 289 submotifs corresponding to the HIV peptidome slightly improved the effect of functional divergence on HIV disease progression for HLA-B, and somewhat diminished the effect for HLA-A (fig. S4A). Thus, both a broad and more restricted calculation of HLA functional divergence reveals an impact of this property on HIV disease.

HIV sequences are highly variable, and the number of distinct nonamers greatly exceeds those derived from consensus HIV sequences. The 2021 LANL compendium of HIV sequences (www.hiv.lanl.gov/content/sequence/HIV/COMPENDIUM/2021compendium.html) contained a total of 1,933,474 unique nonamers, which corresponded to all but one of the 382 submotifs derived from self. We ranked submotifs based on the number of HIV nonamers corresponding to each submotif (ranging from 3 to 163,533). An iterative estimation of the effect on disease progression with stepwise elimination of submotifs represented by the fewest number of peptides was performed for each locus. Upon elimination of submotifs poorly represented among HIV nonamers at each HLA locus, functional divergence trended toward higher HR values (fig. S4B). This pattern may indicate improved accuracy of the estimation of functional divergence following elimination of redundant or irrelevant submotifs. Further elimination of submotifs over the course of 30 elimination steps eventually resulted in reduced HR values, presumably due to elimination of relevant submotifs (fig. S4B).

The inclusion of all potential HIV nonamers in these analyses is likely to somewhat diminish the accuracy in determining the strength of the effect of functional divergence on disease outcome as many of the nonamers included in the analyses do not represent true epitopes (i.e., cannot be processed and presented in vivo). To address this issue, we have used 2483 experimentally verified HIV nonamer epitopes recorded in the LANL database (www.hiv.lanl.gov/content/immunology/variants/ctl_variant.html), which represent 270 submotifs out of the 382 submotifs corresponding to self peptides. Although the LANL epitope data are incomplete and biased toward alleles known to have effects on natural history outcomes of HIV infection, our analyses of functional divergence based on the 270 epitope-relevant submotifs showed consistent differences in HR between the bottom and top octiles for both HLA-A and HLA-B (fig. S4C).

Overall, the use of pertinent, disease-specific submotifs may somewhat improve the detection of functional divergence on disease outcome, but when disease-specific submotifs are not defined or attainable, submotifs based on self peptides are nearly as effective.

Discussion

Protection against viral disease conferred by *HLA* class I heterozygosity is likely mediated by a broader peptide repertoire presented by distinct HLA allotypic combinations relative to the more limited repertoire presented by homozygotes. Previous reports have studied *HLA* class I zygosity dichotomously by comparing heterozygotes composed of any given pair of distinct alleles to homozygotes, providing evidence to support heterozygote advantage (*1*, *8*, *15*–*19*). This approach, however, ignores differences in the breadth of the peptide repertoire among heterozygotes, and any effect these differences may have on disease outcome, response to vaccination, or other clinically relevant phenotypes. We probed more deeply into the impact of complementarity across peptide binding repertoires between pairs of HLA class I allotypes, developing a metric termed functional divergence that takes into account HLA allotype-specific peptide binding data directly. We find that untreated PLWH who carry any given pair of HLA allotypes characterized by greater functional divergence in their peptide binding profiles (i) progress more slowly to AIDS, (ii) are more likely to be an elite controller, and (iii) sustain lower VL levels over time, even when restricting analyses to genetic heterozygotes only.

Properties of HLA class I that affect their ability to present a broader range of antigenic epitopes have been shown to associate with HIV control. Dependence on tapasin for peptide loading varies markedly across HLA allotypes (*20*–*26*). Those HLA class I allotypes that are less dependent on tapasin have been shown to elicit T cell responses with broader epitope specificity and to associate with protection against HIV disease (*25*). Likewise, greater functional divergence, by definition, extends the number of distinct peptides capable of being presented by allotypes encoded by the specific *HLA* genotypes carried in each individual. Properties of HLA allotypes/allotypic pairs, such as tapasin independence and functional divergence, that enhance the repertoire of antigen presentation decrease the chances of viral immune escape, likely explaining their observed associations with control of HIV in natural history cohorts. We predict that these properties will also influence efficacy of vaccination.

Previous studies have employed metrics such as Grantham distance to estimate genetic distance between *HLA* allelic pairs of each given genotype (*4*), an approach that has been a useful initial step in the study of HLA class I functional divergence. Studies of a large HIV cohort (>6000 individuals) indicated that greater genetic divergence between *HLA-B* allelic pairs of any given genotype significantly associated with decreased HIV VLs, but this same trend was not observed for *HLA-A* or *HLA-C* (*5*). A method based on standard peptide binding prediction algorithms, which are indirect and limited in that they ignore antigen processing effects on peptide availability, can somewhat improve the association between divergence and VLs relative to the method based solely on genetic distances (*5*). Although we find highly significant effects of functional divergence for HLA-B and HLA-A, particularly the combination of HLA-A and HLA-B, in analyses of three outcomes to HIV infection (disease progression in seroconverters, elite versus noncontrol of HIV, and mean VL over time), we found no evidence that genetic distance predicts outcome in any of these cohorts. This was true for analyses performed by dichotomizing the top one-fourth of the genetic divergence spectrum versus the remainder of the study participants, as was used in cancer immunotherapy studies (*8*) or using divergence as a continuous variable, as done in HIV studies (*5*). It is possible that our cohorts were not of sufficient size to detect the relatively weaker effect of genetic divergence compared with the more direct and sensitive approach used herein. The strength of the effect of functional divergence as measured using our metric was among the greatest observed in predicting HIV outcome, underscoring the advancement gained in measuring divergence using metrics employing endogenous peptides that have been processed and presented on the cell surface through physiological means.

Previous data have shown that HIV escape mutations occur most frequently at anchor residues (*27*) and these mutations were computationally predicted to severely affect binding of the peptide to the corresponding HLA allotype. A key feature of the metric described herein includes the grouping of peptides by motif similarity where motifs are primarily defined by the anchor residues and noise contributed by minor peptide variation is eliminated by applying low entropy cutoffs. Given that immune escape generally occurs at the same positions defining the submotif categories, it is not surprising that this metric of functional divergence is strongly predictive of HIV outcomes. In our cohort, simple quantifiers such as aggregate peptide repertoire sizes do not show a clear association with HIV progression in heterozygotes, most likely as a consequence of the large proportion of peptides in the repertoire that differ slightly in sequence but are functionally and immunologically

equivalent. Removal of this "noise" is necessary to enable focus on peptide groupings that contribute distinctly to immunological response.

Our findings in an extended cohort of HIV seroconverters confirm that genetic heterozygosity at both *HLA-A* and *HLA-B* associate with protection against HIV progression, but there is no clear effect of *HLA-C* zygosity. Likewise, greater functional divergence at both HLA-A and HLA-B independently or combined significantly associates with slower disease progression and there is no apparent effect of functional divergence at HLA-C. Comparison of individuals with the greatest functional divergence at both HLA-A and HLA-B to those with the lowest divergence levels shows the greatest differentiation in protection, indicating the importance of the two loci acting in unison. The superior ability of HLA-A and HLA-B to present a broad repertoire of peptides as compared with HLA-C was previously reported (*28*, *29*). Furthermore, HLA-A and HLA-B are thought to present a complementary set of peptides in which HLA-A is more fastidious at the F pocket and HLA-B is more fastidious at the B pocket (*29*), which may explain the strength of the effects reached when considering functional divergence at HLA-A and HLA-B combined. Disease-specific refinement of the submotif pool used to estimate functional divergence shows promise as an approach to improve accuracy to some extent but attaining the complete, unbiased disease-specific epitope pool will be challenging. Use of new mass spectrometry technologies to characterize HLA allotype-specific HIV epitopes (*30*) should provide a broader and less biased definition of true epitopes and lead to improved accuracy in tests of functional divergence. However, submotifs determined from the self peptide pool successfully detect the impact of functional divergence and can be applied across disease types.

Our metric is likely to be applicable to other viral diseases, particularly those that undergo immune escape through mutation of epitopes. It is unclear, however, whether the same metric will be predictive of disease outcome when the epitopes are much more similar to self, as in the case of cancer neoantigens where generally only a single amino acid change from self is often observed (*31*, *32*), possibly requiring a more granular approach to defining peptide overlap. Furthermore, the class of the neoantigens (e.g., genomic variants, transcriptomic variants, proteomic variants, or viral-derived variants) can vary depending on tumor type (*33*), which may also result in the necessity to adjust the properties of the metric. The work described herein represents a foundation for determining functional divergence that can be readily modified to suit the particular outcome in question. The metric may be applicable in predicting clinical outcomes such as efficacy of vaccination or immunotherapy as well as the natural history of disease pathogenesis.

REFERENCES AND NOTES

1. M. Carrington *et al*., *Science* **283**, 1748–1752 (1999).
2. A. L. Hughes, T. Ota, M. Nei, *Mol. Biol. Evol.* **7**, 515–524 (1990).
3. E. K. Wakeland *et al*., *Immunol. Res.* **9**, 115–122 (1990).
4. F. Pierini, T. L. Lenz, *Mol. Biol. Evol.* **35**, 2145–2158 (2018).
5. J. Arora *et al*., *Mol. Biol. Evol.* **37**, 639–650 (2020).
6. A. M. Daull *et al*., *Front. Immunol.* **13**, 841470–841470 (2022).
7. C. Féray *et al*., *Ann. Intern. Med.* **174**, 1385–1394 (2021).
8. D. Chowell *et al*., *Nat. Med.* **25**, 1715–1720 (2019).
9. J. Sidney, B. Peters, N. Frahm, C. Brander, A. Sette, *BMC Immunol.* **9**, 1 (2008).
10. S. Sarkizova *et al*., *Nat. Biotechnol.* **38**, 199–209 (2020).
11. M. P. Martin, M. Carrington, *Immunol. Rev.* **254**, 245–264 (2013).
12. M. Cullen, S. P. Perfetto, W. Klitz, G. Nelson, M. Carrington, *Am. J. Hum. Genet.* **71**, 759–776 (2002).
13. *MMWR Recomm. Rep.* **41**, 1–19 (1992).
14. CDC, *MMWR Suppl.* **36**, 1S–15S (1987).
15. C. L. Roark *et al*., *Hum. Immunol.* **83**, 730–735 (2022).
16. N. Shah *et al*., *Leukemia* **25**, 1036–1039 (2011).
17. J. Tang *et al*., *AIDS Res. Hum. Retroviruses* **15**, 317–324 (1999).
18. S. S. Wang *et al*., *Cancer Res.* **78**, 4086–4096 (2018).
19. P. C. Doherty, R. M. Zinkernagel, *Nature* **256**, 50–52 (1975).
20. R. Greenwood, Y. Shimizu, G. S. Sekhon, R. DeMars, *J. Immunol.* **153**, 5525–5536 (1994).
21. C. A. Peh *et al*., *Immunity* **8**, 531–542 (1998).
22. C. A. Peh, N. Laham, S. R. Burrows, Y. Zhu, J. McCluskey, *J. Immunol.* **164**, 292–299 (2000).
23. A. W. Purcell *et al*., *J. Immunol.* **166**, 1016–1027 (2001).
24. D. Zernich *et al*., *J. Exp. Med.* **200**, 13–24 (2004).
25. A. A. Bashirova *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **117**, 28232–28238 (2020).
26. S. M. Rizvi *et al*., *J. Immunol.* **192**, 4967–4976 (2014).
27. J. M. Carlson *et al*., *J. Virol.* **86**, 13202–13216 (2012).
28. M. Rasmussen *et al*., *J. Immunol.* **193**, 4790–4802 (2014).
29. D. Di, J. M. Nunes, W. Jiang, A. Sanchez-Mazas, *Mol. Biol. Evol.* **38**, 1580–1594 (2021).
30. S. Sengupta *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2123406119 (2022).
31. T. N. Schumacher, R. D. Schreiber, *Science* **348**, 69–74 (2015).
32. L. P. Richman, R. H. Vonderheide, A. J. Rech, *Cell Syst.* **9**, 375–382.e4 (2019).
33. N. Xie *et al*., *Signal Transduct. Target. Ther.* **8**, 9 (2023).

ACKNOWLEDGMENTS

We would like to thank G. Nelson for helpful discussions. **Funding:** This project has been funded in whole or in part with federal funds from the Frederick National Laboratory for Cancer Research, under contract no. 75N91019D00024. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products or organizations imply endorsement by the US Government. This research was supported in part by the Intramural Research Program of the NIH, Frederick National Laboratory, Center for Cancer Research. See the extended acknowledgments in Supplementary Materials for full details. **Author contributions:** Conceptualization and study design: M.C., M.V., and C.O. Bioinformatic and statistical analyses: M.V., C.O. HLA genotyping: Y.Y. Supervision: M.C. Cohort sources: S.W., S.B., G.D.K., J.J.G., N.L.M., D.W.H., S.G.D., B.W., X.Y. Writing – original draft: C.O., M.V., A.A.B, M.C. Writing – review and editing: all authors **Competing interests:** Authors declare that they have no competing interests. **Data and materials availability:** All results described in the paper are presented in the main manuscript or the supplementary materials. The functional divergence between every allotypic pair at each locus is provided in data S1 and is available at https://github.com/matviard/FD/. All clinical data underlying the study are available from the cohorts upon application to the cohort curation websites indicated by links in the supplementary materials *Study participants* section. The peptide dataset is publicly available (*10*). The HLA allelic sequences are available from the IPD-IMGT/HLA database (https://www.ebi.ac.uk/ipd/imgt/hla/).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adk0777
Materials and Methods
Extended Acknowledgements
Figs. S1 to S4
Tables S1 to S4
References (*34*, *35*)
Data S1

Submitted 1 August 2023; accepted 7 December 2023
10.1126/science.adk0777

NANOMATERIALS

Near-unity NIR phosphorescent quantum yield from a room-temperature solvated metal nanocluster

Wan-Qi Shi,[1]† Linlin Zeng,[2]† Rui-Lin He,[1] Xu-Shuang Han,[1] Zong-Jie Guan,[1,3] Meng Zhou,[2]* Quan-Ming Wang[1]*

Metal nanoclusters have emerged as promising near-infrared (NIR)–emissive materials, but their room-temperature photoluminescence quantum yield (PLQY), especially in solution, is often low (<10%). We studied the photophysics of $Au_{22}(^tBuPhC{\equiv}C)_{18}$ (Au_{22}) and its alloy counterpart $Au_{16}Cu_6(^tBuPhC{\equiv}C)_{18}$ ($Au_{16}Cu_6$) (where ^{t}Bu is *tert*-butyl and Ph is phenyl) and found that copper (Cu) doping suppressed the nonradiative decay (~60-fold less) and promoted intersystem crossing rate (~300-fold higher). The $Au_{16}Cu_6$ nanocluster exhibited >99% PLQY in deaerated solution at room temperature with an emission maximum at 720 nanometers tailing to 950 nanometers and 61% PLQY in the oxygen-saturated solution. The approach to achieve near-unity PLQY could enable the development of highly emissive metal cluster materials.

Luminophores with emissions at the near-infrared (NIR) region (first NIR window 700 to 900 nm and second NIR window 1000 to 1700 nm) have applications in areas such as biological imaging and optical communication (*1*–*3*). Colloidal quantum dots and organic fluorescent dyes are two typical kinds of NIR-emitting materials that have been extensively studied (*4*–*7*). Recently, ultrasmall gold nanoclusters (Au NCs) have emerged as a new class of NIR-emissive materials (*8*–*11*). These structurally well-defined compounds can enable a fundamental understanding of the underlying photophysical mechanisms (*12*, *13*), which is helpful for the rational synthesis of clusters with improved and tailored optical properties. Given their relatively low toxicity (*14*) and large Stokes shift (*15*), excellent photostability (*16*), and solution processibility (*17*), Au NCs have potential applications in deep-tissue bioimaging (*18*, *19*) and solution-processed light-emitting diodes (*20*). However, the photoluminescence quantum yields (PLQYs) of these NCs usually are low, with only a few reaching 10% in normal solvents under ambient conditions, including Au_{13} protected by N-heterocyclic carbene (~16% at 730 nm) (*21*), alkynyl-protected Au_{24} (~12% at 925 nm) (*8*), $Au_{42}(PET)_{32}$ (PET is 2-phenylethanethiolate, ~12%, dual emission at 875 and 1040 nm) (*22*), and $Au_{38}S_2(S\text{-}Adm)_{20}$ (S-Adm is 1-admantanethiolate, ~15% at 900 nm) (*23*). Most NCs to date have a low PLQY (<1%) in the NIR region, which limits their usefulness.

It is challenging to obtain a high PLQY in the NIR region in solution because emitters suffer from high nonradiative decay rates according to the energy gap law (*24*, *25*). Materials with high quantum efficiency, especially those near unity, have long been pursued (*26*, *27*). Highly emissive materials could enable underexplored applications, such as efficient luminescent solar concentrators (*28*) and optical refrigeration (*29*, *30*).

Here, we report an alloy metal NC $Au_{16}Cu_6$ ($^tBuPhC{\equiv}C)_{18}$ ($Au_{16}Cu_6$) (where ^{t}Bu is *tert*-butyl and Ph is phenyl) that is isostructural to Au_{22} ($^tBuPhC{\equiv}C)_{18}$ (Au_{22}) and displayed a near-unity phosphorescence quantum yield in the NIR region in solution at room temperature (RT). The PLQY of $Au_{16}Cu_6$ of >99% in deaerated solutions was 10-fold higher than that of Au_{22}. We performed single-crystal x-ray diffraction, photophysical measurements, and theoretical calculations to investigate the doping effects to determine the relation between geometric and electronic structures and luminescence properties. Ultrafast intersystem crossing (ISC) and suppressed nonradiative decay were found to underlie the distinctive photophysical behavior of $Au_{16}Cu_6$.

Synthesis and crystal structures of Au_{22} and $Au_{16}Cu_6$

Au_{22} clusters protected by thiolate or alkynyl ligands are among the most extensively studied luminescent clusters. $Au_{22}(SG)_{18}$ (SG is glutathione thiolate) with 8% PLQY (*31*) was first reported by Xie *et al.*, and the predicted structural information of $Au_{22}(SR)_{18}$ (where R is a methyl group) (*32*) was experimentally confirmed with alkynyl-protected Au_{22} by Han *et al.* (*33*) and Tsukuda and colleagues (*34*) independently. These studies demonstrated the importance of weak reduction conditions in the synthesis of Au_{22} clusters. We use a methanol solution of sodium hydroxide as the reducing agent to obtain Au_{22} with 4-(*tert*-butyl)phenylacetylene ligands. For $Au_{16}Cu_6$, borane *tert*-butylamine or sodium borohydride with greater reducing power was needed because of the difficulty in achieving copper reduction.

The composition of Au_{22} and $Au_{16}Cu_6$ was determined by electrospray ionization mass spectrometry (ESI-MS) in positive-ion mode. The prominent peak at a mass-charge ratio m/z of 3714.5 (Fig. 1A) was in agreement with the dicationic ion $[Au_{22}(^tBuPhC{\equiv}C)_{18}Cs_2]^{2+}$. The isotopic distribution pattern matched the simulated one (Fig. 1A, inset). A small peak at m/z = 4908.3 was also observed that corresponded to $[Au_{44}(^tBuPhC{\equiv}C)_{36}Cs_3]^{3+}$, which corresponds to a structure in which two $Au_{22}(^tBuPhC{\equiv}C)_{18}$ molecules were combined by three Cs^+ ions under MS conditions.

We also determined the composition of $Au_{16}Cu_6$. The prominent peak observed at m/z = 3314.3 (Fig. 1B) corresponded to $[Au_{16}Cu_6(^tBuPhC{\equiv}C)_{18}Cs_2]^{2+}$. The observed isotopic pattern matched the simulated one (Fig. 1B, inset). The small peaks at m/z = 3636.4 and 4374.8 were assigned to $[Au_{17}Cu_6(^tBuPhC{\equiv}C)_{20}Cs_3]^{2+}$

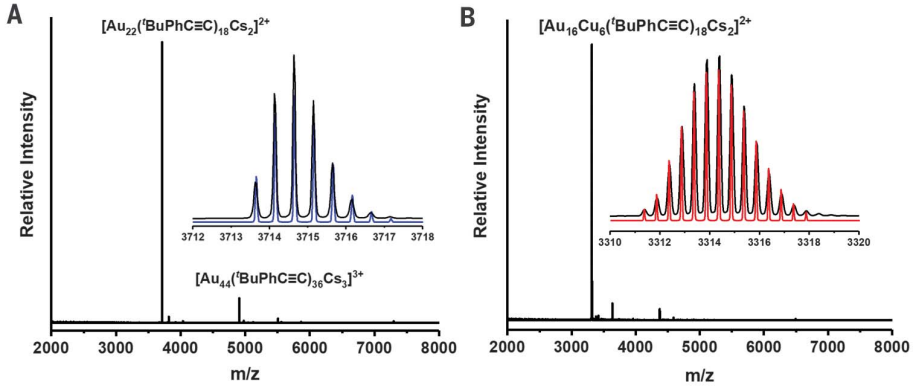


Fig. 1. Mass spectra. (**A**) Mass spectra of $[Au_{22}(^tBuPhC{\equiv}C)_{18}Cs_2]^{2+}$. Inset: The measured (black trace) and simulated (blue trace) isotopic patterns of $[Au_{22}(^tBuPhC{\equiv}C)_{18}Cs_2]^{2+}$. (**B**) Mass spectra of $[Au_{16}Cu_6(^tBuPhC{\equiv}C)_{18}Cs_2]^{2+}$. Inset: The measured (black trace) and simulated (red trace) isotopic patterns of $[Au_{16}Cu_6(^tBuPhC{\equiv}C)_{18}Cs_2]^{2+}$.

[1]Department of Chemistry, Engineering Research Center of Advanced Rare Earth Materials (Ministry of Education), Tsinghua University, Beijing 100084, P. R. China. [2]Hefei National Research Center for Physical Sciences at the Microscale, Department of Chemical Physics, University of Science and Technology of China, Hefei, Anhui 230026, P. R. China. [3]Department of Chemistry, College of Chemistry and Chemical Engineering, Hunan University, Changsha 410082, P. R. China.
*Corresponding author. Email: qmwang@tsinghua.edu.cn (Q.-M.W.); mzhou88@ustc.edu.cn (M.Z.)
†These authors contributed equally to this work.

and $[Au_{32}Cu_{12}(^tBuPhC{\equiv}C)_{36}Cs_3]^{3+}$, respectively. The former form by adding a $[AuCs_3(^tBuPhC{\equiv}C)_2]^{2+}$ to $Au_{16}Cu_6(^tBuPhC{\equiv}C)_{18}$, and the latter by combining two $Au_{16}Cu_6(^tBuPhC{\equiv}C)_{18}$ molecules with three Cs^+ ions. On the basis of both the ESI-MS and x-ray data, the formulas of Au_{22} and $Au_{16}Cu_6$ were determined to be $Au_{22}(^tBuPhC{\equiv}C)_{18}$ and $Au_{16}Cu_6(^tBuPhC{\equiv}C)_{18}$, respectively.

Single-crystal structural analysis revealed that Au_{22} and $Au_{16}Cu_6$ are isostructural (Fig. 2, A and B). The structure of Au_{22} consisted of an hourglass-shaped double tetrahedra Au_7 core, a $[Au_6(^tBuPhC{\equiv}C)_6]$ ring surrounding the waist of the Au_7, and three $[Au_3(^tBuPhC{\equiv}C)_4]$ staples. The trimeric staples bridged the upper and lower triangles of the middle Au_7 core through π-coordination from the alkynyl groups. The Au_7 core combined with the six Au atoms of the ring could also be viewed as a centered cuboctahedron Au_{13} core that was flattened along the C_3 axis. The alkynyl ligands show H-type (π-M-π, M is Au/Cu) and bridging coordination modes in the staples (fig. S3). The structure of $Au_{16}Cu_6$ was almost identical to that of Au_{22}, except that six Cu atoms replaced Au (Fig. 2B). Because the structure has threefold symmetry, the six doping sites could be classified in two types, one type in the Au_3Cu_3 ring and the other in the Au_2Cu staples (Fig. 2, C and D).

The doping of Cu led to structural contraction of $Au_{16}Cu_6$ compared to Au_{22} as reflected in the several bond-length changes. The average M-Au distance from the Au_6 ring to the center Au was 3.418 Å, whereas this distance shortened to 3.121 Å in $Au_{16}Cu_6$. The distance from the vertex M to the adjacent Au on the Au_2M staples changed from 3.391 to 3.151 Å as the staples changed from Au_3 to Au_2Cu. The M-Au distance from the vertex M to the Au_3M_3 ring was reduced from 3.391 to 2.976 Å. Even considering the difference in atomic radius between Au (1.44 Å) and Cu (1.28 Å), the change in the Cu-Au distance is much larger than the radius difference (0.30, 0.24, and 0.42 Å versus 0.16 Å, respectively). Notably, the introduction of Cu also decreased the lengths of Au-Au bonds in the Au-Au bond lengths within the Au_7 core (2.720 Å versus 2.705 Å), the distance from Au in the Au_3M_3 ring to the central Au (3.295 Å versus 3.159 Å), and the Au-Au distance between Au_2M staple and the $Au_{10}M_3$ core (3.477 Å versus 3.273 Å). The detailed bond-length data are listed in table S1.

Optical properties and theoretical calculations

The absorption spectra of Au_{22} and $Au_{16}Cu_6$ in CH_2Cl_2 are shown in Fig. 3A. Au_{22} had two prominent absorption peaks at 470 and 520 nm, and a shoulder at 385 nm. By contrast, $Au_{16}Cu_6$ showed two prominent peaks at 530 and 580 nm and a shoulder at 390 nm, and the Cu doping red-shifted the absorption peaks. Thus, the Au_{22} solution was orange, whereas the solution of $Au_{16}Cu_6$ was pink. Doping with Cu narrowed the optical energy gap from 2.19 to 1.92 eV (fig. S6). Both clusters show high stability and remained intact at RT in the solid state for at least 2 years.

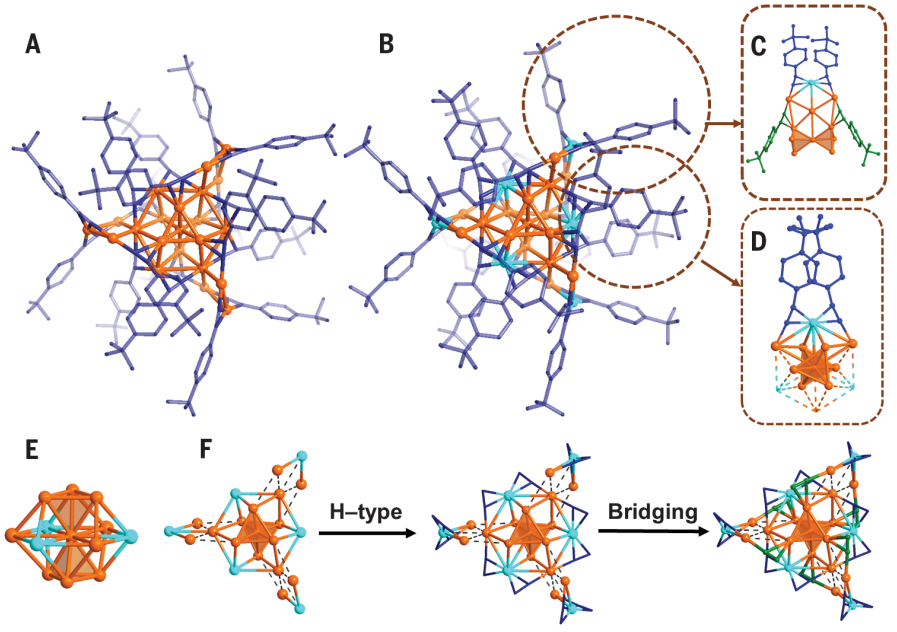


Fig. 2. Molecular structures of Au_{22} and $Au_{16}Cu_6$ and structure anatomy of $Au_{16}Cu_6$. Molecular structures of $Au_{22}(^tBuPhC{\equiv}C)_{18}$ (**A**) and $Au_{16}Cu_6(^tBuPhC{\equiv}C)_{18}$ (**B**); H atoms have been omitted for clarity. (**C** to **F**) The structure anatomy of $Au_{16}Cu_6(^tBuPhC{\equiv}C)_{18}$. Coordination modes of Cu atoms in Au_2Cu staples (C) and Au_6Cu_6 ring (D). $Au_{10}Cu_3$ kernel comprises a covertex bitetrahedral Au_7 core and a Au_3Cu_3 ring surrounded on the waist (E). The structural evolution of $Au_{16}Cu_6(^tBuPhC{\equiv}C)_{18}$ (F); *tert*-butyl phenyl groups have been omitted for clarity. Orange: Au atoms; light blue, Cu atoms; blue, deep blue, or green, C atoms.

Time-dependent density functional theory (TD-DFT) calculations were performed to explore the electronic structures of Au_{22} and $Au_{16}Cu_6$. The simulated absorption spectra for the clusters correlated well with the experimental results (Fig. 3, B and C), although peaks were more pronounced in the simulated spectra. In Au_{22}, a small peak (peak a) at 500 nm was observed, corresponding to the highest occupied molecular orbital–lowest unoccupied molecular orbital (HOMO→LUMO) transition (93%).

This peak could barely be resolved at RT because of the presence of strong electron–phonon interactions. Thus, temperature-dependent optical measurements were conducted for verification. The absorption peaks were sharper and slightly blue-shifted at 80 K, and new vibronic features appear at around 570 and 422 nm (fig. S8A). The main absorption (peak b) at 475 nm arose primarily from HOMO → LUMO + 1 (45%) and HOMO-7 → LUMO (24%) transitions and exhibited pronounced ligand-to-metal charge-transfer characteristics.

The first absorption (peak α) of $Au_{16}Cu_6$ is assigned to HOMO → LUMO (85%), and this transition has much larger oscillator strength than that in Au_{22} (0.081 versus 0.018). The Kohn-Sham molecular orbital energy level diagram reveals that the LUMO of $Au_{16}Cu_6$ predominantly consists of 6sp atomic orbitals from gold. By contrast, HOMO is composed of a large percentage of 5d orbitals of gold.

Notably, $Au_{16}Cu_6$ has a different electronic structure from that of Au_{22}. Substantial involvement of Cu in the frontier orbitals of $Au_{16}Cu_6$ was observed (Figs. 3, D and E). The participation of Cu atoms leads to a modification in the energy level of $Au_{16}Cu_6$ compared to Au_{22}. Specifically, the LUMO level of $Au_{16}Cu_6$ is distinctly shifted toward a lower energy range, whereas the energy of HOMO increases slightly, in comparison to Au_{22} (Fig. 3, D and E, and fig. S9). The incorporation of Cu results in the clear delocalization of electrons from Au to the Au_2Cu staples and the π orbital of alkynyl ligands, and the Au_3Cu_3 ring in $Au_{10}Cu_3$ core has a large contribution to the LUMO of $Au_{16}Cu_6$ (fig. S10). Thus, replacing some of the gold atoms with copper increases the energy level of the HOMO and lowers the energy level of the LUMO, which is responsible for the narrowing of energy gaps and the red-shift of absorptions. Consequently, excited-state dynamics could be influenced after doping with Cu.

The luminescence properties of Au_{22} and $Au_{16}Cu_6$ in CH_2Cl_2 at RT are shown in Fig. 4, A and B (black lines). Au_{22} shows an emission peak at 690 nm, with an absolute PLQY of

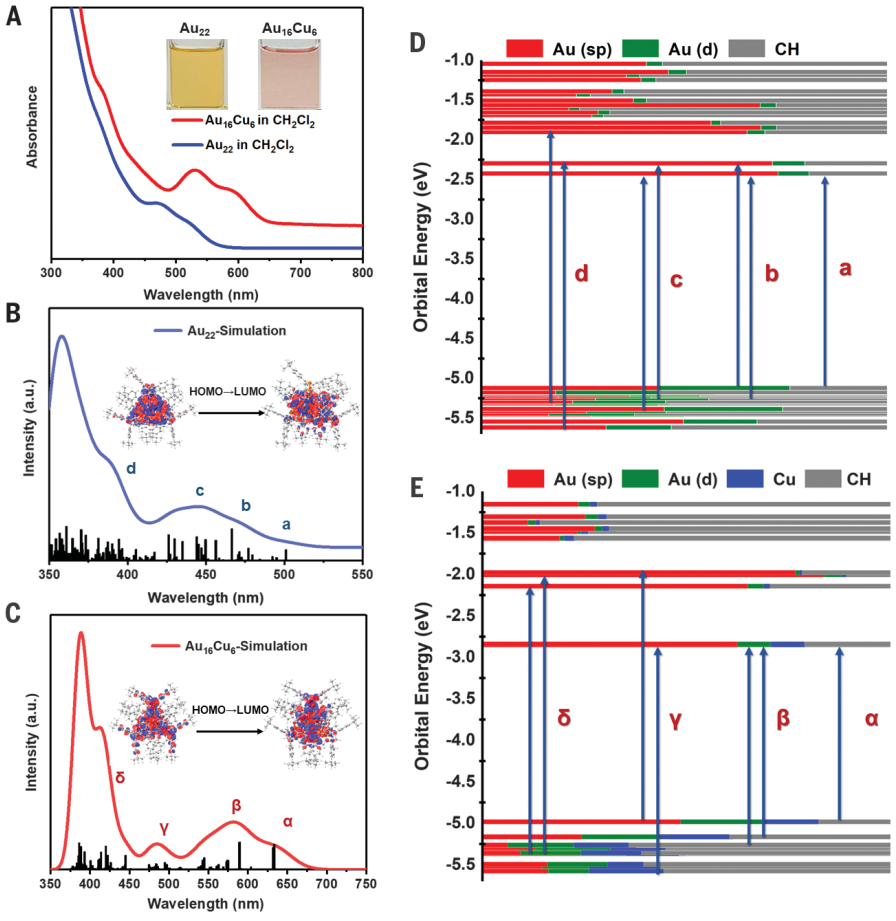


Fig. 3. Absorption and TD-DFT calculation results of Au_{22} and $Au_{16}Cu_6$. (**A**) The optical absorption spectra of Au_{22} (blue trace) and $Au_{16}Cu_6$ (red trace). Inset: Photos of Au_{22} and $Au_{16}Cu_6$ in solution. Simulated ultraviolet–visible spectra of Au_{22} (**B**) and $Au_{16}Cu_6$ (**C**). Inset: Parts of their frontier orbitals, respectively. Kohn–Sham molecular orbital energy level diagram and the associated populations of atomic orbitals in each molecular orbital for Au_{22} (**D**) and $Au_{16}Cu_6$ (**E**).

~9% in the air. $Au_{16}Cu_6$ exhibited an intense emission at ~720 nm, and the PLQY in air was measured to be ~95% by the absolute method (fig. S11) and ~91% relative to Nile blue in ethanol (Fig. 4C). The PL intensity of $Au_{16}Cu_6$ was much higher than that of Au_{22} as revealed in the excitation–emission contour maps (Fig. 4, D and E). In deaerated solution, the PLQY of $Au_{16}Cu_6$ reached 100% as determined by both absolute and relative methods. When the solution was saturated with O_2, PL intensities of Au_{22} and $Au_{16}Cu_6$ were quenched to ~7% and ~60% (fig. S15), respectively, implying the involvement of the spin-forbidden triplet states. Furthermore, the time-resolved PL spectra of the Au_{22} and $Au_{16}Cu_6$ measured by time-correlated single-photon counting revealed that PL lifetimes (τ) are 485 ns and 1.64 μs, respectively (Fig. 4F).

To unravel the underlying mechanism of PL enhancement, we calculated the radiative decay rate (k_r) and nonradiative decay (k_{nr}) (table S2). With the k_r slightly promoted (~3-fold), k_{nr} sharply declined (~60-fold slow), indicating that doping with copper greatly suppressed the nonradiative relaxation channels (*35*). We attributed this change to the more compact structure of $Au_{16}Cu_6$. In addition, we varied the excitation wavelengths and found that the emission peak positions of Au_{22} and $Au_{16}Cu_6$ remained unchanged (fig. S16), suggesting that the PL originated from the same excited state at RT.

Excited-state dynamics

We probed the excited-state dynamics of Au_{22} and $Au_{16}Cu_6$ using femtosecond and nanosecond transient-absorption spectroscopy (fs and ns-TA). The ns-TA of two NCs showed only one decay component (figs. S18 and S19), with an excited-state lifetime of 600 ns for Au_{22} and 1.3 μs for $Au_{16}Cu_6$, which are comparable with their PL lifetimes. In contrast to the simple nanosecond dynamics, different femtosecond dynamics were observed in two NCs. In both NCs, we observed broad excited-state absorption (ESA). For Au_{22} (Fig. 5A), with 380-nm excitation, the ESA peak exhibited weak ultrafast decay in the first 1 ps but increased between 1 and 300 ps in the fs-TA data map. For $Au_{16}Cu_6$ (Fig. 5B), ultrafast decay occurred during the first 1 ps and was followed by a nondecay relaxation component that remained constant between 2 ps and 2 ns. The TA kinetics probed at 700 nm and the different dynamics of the two NCs are displayed in Fig. 5C.

In gold and other noble-metal NCs, spin-orbit coupling (SOC) can give rise to ultrafast ISC to form a triplet state. To identify the origin of the long-lived emissive states in two NCs, we performed triplet sensitization experiments. We probed the ns-TA dynamics of a solution of NCs and perylene by selectively exciting the NCs without pumping perylene molecules. With exclusive excitation of Au_{22} and $Au_{16}Cu_6$, we observed the generation ESA signal of the triplet state (T_1) of perylene between 0 and 100 ns (fig. S20). Triplet energy transfer (TET) occurred in both NCs, and the excited-state energy was transferred from the triplet state of NCs to the triplet state of perylene molecules, indicating that both NCs emitted through triplet states.

On the basis of the TET experiment, ISC should form long-lived triplet emissive states in both NCs. However, Au_{22} and $Au_{16}Cu_6$ showed drastically different ISC rates (Fig. 5C). In Au_{22}, a very slow rise in the ESA signal indicated slow ISC with a time constant of 148 ps. In $Au_{16}Cu_6$, ultrafast relaxation (0.5 ps) indicated that internal conversion occurred with ISC at the same time. On the basis of the time-resolved data, the excited-state relaxation diagram can be summarized in Fig. 5D. In contrast to the very slow ISC (148 ps) in Au_{22}, a rapid ISC (~0.5 ps) occurred in $Au_{16}Cu_6$ and more efficiently generated the triplet state with much higher PLQY (~100%). With Cu doping, the ISC rate was accelerated by two orders of magnitude from Au_{22} to $Au_{16}Cu_6$.

The ISC rate k_{isc} is mainly related to SOC and vibration contribution, as given by the following equation (*36*) according to first-order perturbation theory:

$$k_{isc} = \frac{2\pi}{\hbar}\left\langle \Psi(S)\middle|\hat{H}_{SOC}\middle|\Psi(T)\right\rangle^2 \rho$$

where $\Psi(S)$ and $\Psi(T)$ are the electronic wave functions of the singlet and triplet states; H_{SOC} is the SOC Hamiltonian; ρ is the Franck–Condon weighted density of states, representing the vibrational contribution; and $\hbar$ is Planck's constant divided by 2π. ρ can be evaluated in the form of a Boltzmann distributed population as

$$\rho = \frac{1}{\sqrt{4\pi\lambda RT}}\exp\left[-\frac{(\Delta E_{ST}+\lambda)^2}{4\lambda k_B T}\right]$$

where ΔE_{ST} is the singlet–triplet splitting energy, λ is the Marcus reorganization energy (due

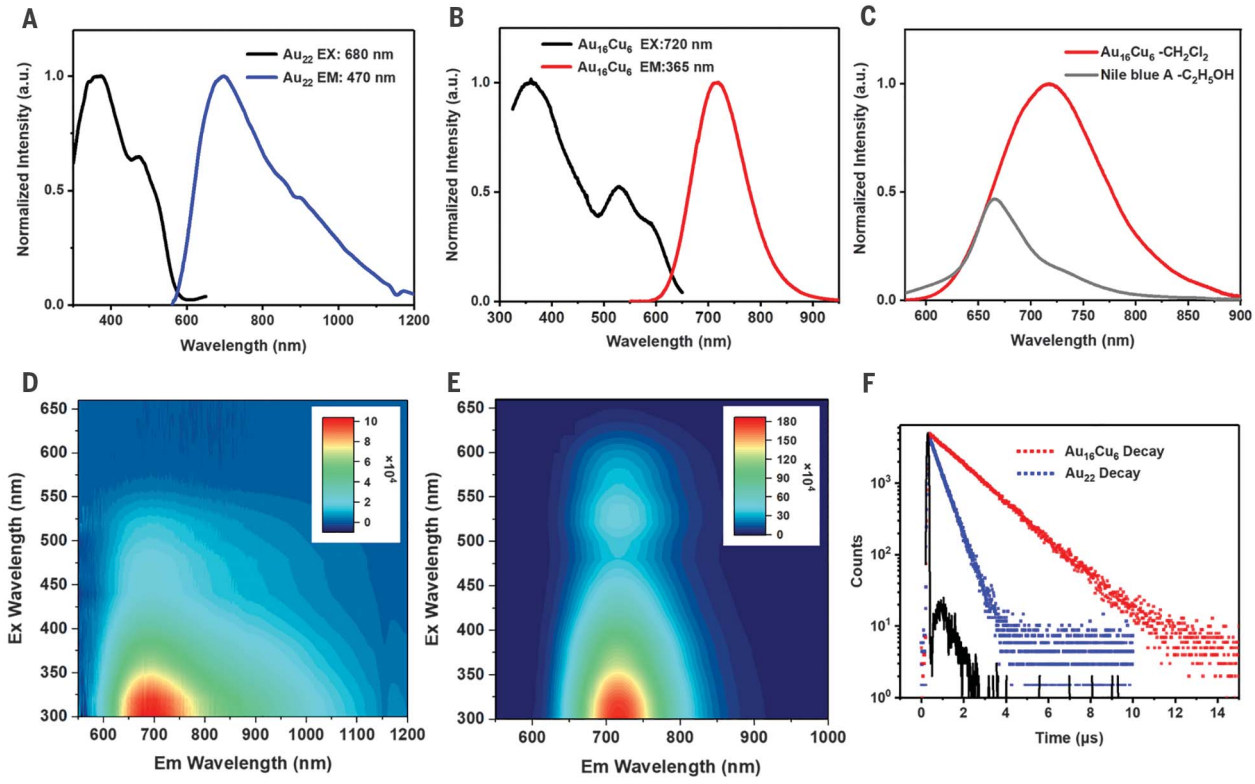


Fig. 4. Luminescent properties of Au_{22} and $Au_{16}Cu_6$ in CH_2Cl_2. Excitation and emission spectra of Au_{22} (**A**) and $Au_{16}Cu_6$ (**B**). (**C**) Comparison of the emission spectra of $Au_{16}Cu_6$ and Nile blue A. $Au_{16}Cu_6$ is in CH_2Cl_2 and Nile blue A in C_2H_5OH. The value of Nile blue A is 0.27 in C_2H_5OH (*38*). Excitation–emission contour maps of Au_{22} (**D**) and $Au_{16}Cu_6$ (**E**). PL decay traces of Au_{22} (blue) and $Au_{16}Cu_6$ (red) (**F**).

to intramolecular, low-frequency vibrations and solvent-induced relaxation) (*37*), k_B is the Boltzmann constant, R is the gas constant, and T is temperature.

Because Au_{22} and $Au_{16}Cu_6$ are isostructural, the doping of Cu made the largest contribution to the increase of k_{isc}. Replacing six Au atoms with Cu in Au_{22} had a negligible effect on the SOC term, whereas the doping resulted in small ΔE_{ST} and λ terms and in turn a much greater ρ. As a result, the drastically increased k_{isc} led to a fast population of electrons in the triplet state. In addition, the doping led to a more compact structure, and this geometric effect helped suppress nonradiative decay. Both the fast ISC rate and the suppressed nonradiative decay gave rise to the near-unity PLQY.

Discussion

Gold NCs have demonstrated high NIR luminescence from the involvement of both sp and d bands in the frontier orbitals. We achieved near-unity phosphorescent quantum yield by doping copper in a gold NC. Excited-state dynamics revealed that the doping led to a 300-fold increase in the ISC rate from S_1 to T_1, which made the near-unity PLQY possible. This work demonstrates that near-unity PLQY can be attained with an alloy of gold-copper NCs even in solution at RT, which will enable applications

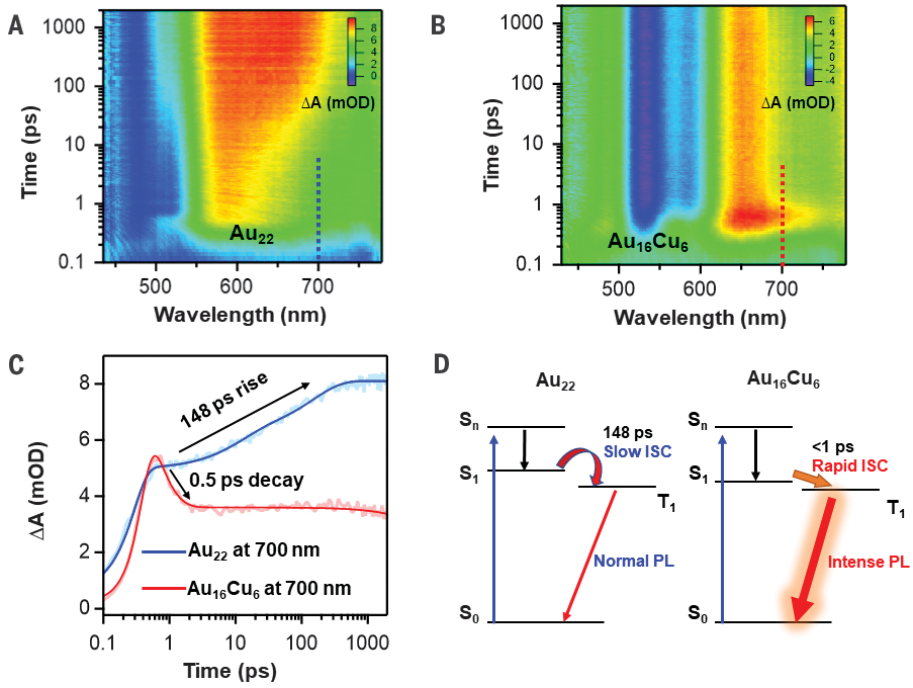


Fig. 5. Ultrafast excited-state dynamics results of Au_{22} and $Au_{16}Cu_6$. (**A**) Fs-TA data map of Au_{22} NCs with 380-nm excitation. (**B**) Fs-TA data map of $Au_{16}Cu_6$ NCs with 380-nm excitation. (**C**) TA kinetic traces probed at 700 nm and corresponding fits of two NCs. (**D**) Excited-state relaxation diagram obtained from the time-resolved data.

ranging from biological imaging to luminescent devices. In the future, it will be interesting to see the doping effects of other dopant metals, which may lead to unexpected optical properties. It is also desirable to regulate the emission energy region of metal NCs in a systematic way, as the ligand shell and metal core structure can be modified through various synthetic strategies.

REFERENCES AND NOTES

1. G. Hong, A. L. Antaris, H. Dai, *Nat. Biomed. Eng.* **1**, 0010 (2017).
2. G. Hong *et al*., *Nat. Photonics* **8**, 723–730 (2014).
3. A. Zampetti, A. Minotto, F. Cacialli, *Adv. Funct. Mater.* **29**, 1807623 (2019).
4. Y. Chen, S. Wang, F. Zhang, *Nat. Rev. Bioeng.* **1**, 60–78 (2023).
5. H. Bahmani Jalali, L. De Trizio, L. Manna, F. Di Stasio, *Chem. Soc. Rev.* **51**, 9861–9881 (2022).
6. L. Dai, Q. Zhang, Q. Ma, W. Lin, *Coord. Chem. Rev.* **489**, 215193 (2023).
7. E. H. Sargent, *Adv. Mater.* **17**, 515–522 (2005).
8. X.-K. Wan *et al*., *Angew. Chem. Int. Ed.* **54**, 9683–9686 (2015).
9. D. Li *et al*., *Small* **16**, e2003851 (2020).
10. G. Wang, T. Huang, R. W. Murray, L. Menard, R. G. Nuzzo, *J. Am. Chem. Soc.* **127**, 812–813 (2005).
11. W.-D. Si *et al*., *Sci. Adv.* **9**, eadg3587 (2023).
12. M. Zhou *et al*., *Science* **364**, 279–282 (2019).
13. C. M. Aikens, *Acc. Chem. Res.* **51**, 3065–3073 (2018).
14. K. Zheng, M. I. Setyawati, D. T. Leong, J. Xie, *ACS Nano* **11**, 6904–6910 (2017).
15. Q. Li, C. J. Zeman4th, Z. Ma, G. C. Schatz, X. W. Gu, *Small* **17**, e2007992 (2021).
16. H. Ma *et al*., *Sci. Adv.* **9**, eadh7828 (2023).
17. M. Zhou, Z. Lei, Q. Guo, Q.-M. Wang, A. Xia, *J. Phys. Chem. C Nanomater. Interfaces* **119**, 14980–14988 (2015).
18. G. Yang *et al*., *J. Am. Chem. Soc.* **145**, 11879–11898 (2023).
19. H. Liu *et al*., *Adv. Mater.* **31**, e1901015 (2019).
20. T. Li, Z. Wang, Y. Zhang, Z. Wu, *Nanomaterials* **12**, 3837 (2022).
21. M. R. Narouz *et al*., *J. Am. Chem. Soc.* **141**, 14997–15002 (2019).
22. L. Luo, Z. Liu, X. Du, R. Jin, *J. Am. Chem. Soc.* **144**, 19243–19247 (2022).
23. Q. Li, C. J. I. V. Zeman 4th, G. C. Schatz, X. W. Gu, *ACS Nano* **15**, 16095–16105 (2021).
24. R. Englman, J. Jortner, *Mol. Phys.* **18**, 145–164 (1970).
25. K. Kwak, V. D. Thanthirige, K. Pyo, D. Lee, G. Ramakrishna, *J. Phys. Chem. Lett.* **8**, 4898–4905 (2017).
26. R. Hamze *et al*., *Science* **363**, 601–606 (2019).
27. M. Amani *et al*., *Science* **350**, 1065–1068 (2015).
28. F. Meinardi, F. Bruni, S. Brovelli, *Nat. Rev. Mater.* **2**, 17072 (2017).
29. M. Sheik-Bahae, R. I. Epstein, *Nat. Photonics* **1**, 693–699 (2007).
30. T. Sadi, I. Radevici, J. Oksanen, *Nat. Photonics* **14**, 205–214 (2020).
31. Y. Yu *et al*., *J. Am. Chem. Soc.* **136**, 1246–1249 (2014).
32. Y. Pei, J. Tang, X. Tang, Y. Huang, X. C. Zeng, *J. Phys. Chem. Lett.* **6**, 1390–1395 (2015).
33. X.-S. Han *et al*., *Angew. Chem. Int. Ed.* **59**, 2309–2312 (2020).
34. S. Ito, S. Takano, T. Tsukuda, *J. Phys. Chem. Lett.* **10**, 6892–6896 (2019).
35. R. Czerwieniec, M. J. Leitl, H. H. H. Homeier, H. Yersin, *Coord. Chem. Rev.* **325**, 2–28 (2016).
36. K. Schmidt *et al*., *J. Phys. Chem. A* **111**, 10490–10499 (2007).
37. J.-L. Brédas, D. Beljonne, V. Coropceanu, J. Cornil, *Chem. Rev.* **104**, 4971–5004 (2004).
38. A. M. Brouwer, *Pure Appl. Chem.* **83**, 2213–2228 (2011).

ACKNOWLEDGMENTS

We thank T. Huang and L. Duan of Tsinghua University for help with luminescence measurements. **Funding:** This work was supported by the National Natural Science Foundation of China 91961201, 21971136 (Q.-M.W.), 22273095 (M.Z.), the National Key R&D Program of China 2022YFA1503900 (Q.-M.W.), and the Chinese Academy of Sciences XDB0450202, YSBR-007 (M.Z.). **Author contributions:** Project concept: Q.-M.W., W.-Q.S. Synthesis: W.-Q.S., X.-S.H. Measurements and data analysis: W.-Q.S., L.Z., X.-S.H., Z.-J.G., M.Z. Theoretical calculation: R.-L.H. Z.-J.G. Writing – original draft: W.-Q.S., L.Z., X.-S.H. Writing – review and editing: Q.-M.W., M.Z., W.-Q.S., L.Z. **Competing interests:** A Chinese patent application (ref. CN202311716266.1) was submitted by Tsinghua University that covers the highly NIR phosphorescent gold/copper nanocluster. **Data and materials availability:** All data are available in the main text or the supplementary materials. The X-ray crystallographic structure reported in this article has been deposited at the Cambridge Crystallographic Data Centre (CCDC) with deposition number (CCDC 2289983-2289984).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adk6628
Materials and Methods
Supplementary Text
Figs. S1 to S22
Tables S1 to S6
References (*39*–*48*)

Submitted 4 September 2023; accepted 7 December 2023
10.1126/science.adk6628

WORKING LIFE

By **Maj-Britt Buchholz**

My lab's lighter footprint

I went into our lab meeting feeling dejected. For months, I had been trying to recycle some of our lab's plastic waste into 3D-printed lab tools, but it wasn't going well. "I've only been able to print three small items," I told my colleagues. They knew how committed I was to reducing the environmental footprint of our research. So one asked, "Maybe we should be thinking about how we can avoid having so much plastic waste in the first place?" Another chimed in: "Could we figure out how much energy we are consuming and how to reduce that?" Their questions buoyed my spirits—and led to a series of changes in how we operate in the lab.

When I started my Ph.D., I knew my work in cancer research was generating a lot of waste, but I had accepted it as a necessary evil. That all changed one day when I saw a huge pile of plastic waste—pipette boxes, tubes, packaging—stacked against the wall. Until then, the lab had operated at a reduced capacity because of the COVID-19 pandemic. But as we ramped up our research and returned to full capacity, our waste production skyrocketed visibly.

I went to my Ph.D. adviser and pitched the idea of creating 3D-printed lab tools out of some of our waste. She encouraged me and helped me write a proposal. We secured funding to purchase recycling equipment and a dedicated 3D printer, but the project didn't go quite how I envisioned. The equipment required me to individually load each item to be recycled, which was incredibly time consuming. It was also tricky to get the recycled material into a consistent shape that could be fed into the 3D printer. After months of struggling, I feared that my vision might not be feasible.

"I encourage scientists ... to reduce the environmental impact of their own research."

It was frustrating to put so much energy into a project and not be able to get it fully off the ground. So it was a welcome relief when my colleagues were eager to brainstorm other ways to reduce our environmental footprint.

We decided to create a group dedicated to pushing for "greener" ways to operate, inviting researchers in other labs at our institution as well as our own. Many of my colleagues, including Ph.D. students, lab technicians, and principal investigators, were eager to join. We began meeting monthly to come up with solutions.

Over the following months, we launched a series of initiatives that we feel have made a difference. We introduced new recycling bins to each lab after learning that some of our plastic waste could be recycled if it was sorted properly. We took a hard look at our lab protocols to devise ways to reduce our use of plastic items. And we purchased meters to track the energy consumption of various pieces of lab equipment. That helped us create rules for each device, which we outlined on stickers: "Turn me off at the end of the day," "Never turn me off," or "Turn me off if empty or not in use."

Our most energy-hungry devices were the ultra–low-temperature freezers, each of which consumed more energy than a typical household. So, we started an institutewide freezer challenge, asking labs to remove samples that were not needed anymore and consolidate what remained in fewer freezers. Through those efforts, we managed to turn off multiple freezers. We also increased the temperature of several ultra–low-temperature freezers after realizing some samples didn't have to be stored so cold.

I'm still disappointed that I was not able to fully realize my 3D-printing vision, and I am stubborn enough to keep trying. But I've also learned that if one approach isn't working, sometimes it's best to take a step back and look for other solutions. I'm thankful that with the help of the academic community around me, we were able to do that.

I encourage scientists everywhere to think about ways to reduce the environmental impact of their own research. The work we do is important. But we should also do our part to be as sustainable as we can. After all, what significance does our science hold if, ultimately, there remains no viable planet we can call our home? ■

Maj-Britt Buchholz is a Ph.D. student at the Princess Máxima Center for Pediatric Oncology. Do you have an interesting career story to share? Send it to SciCareerEditor@aaas.org.