

nature

FIELD GUIDE

Magnetic structure of the interstellar medium reveals how stars start to form

Moving target
Is precision health care the panacea that some practitioners claim?

Coronavirus
What the world can learn from Africa's experience of COVID

Cycle of life
Earth's eccentric orbit influences evolution of marine phytoplankton

Vol 597 No 7412
6 January 2022

Nature.2022.01.08

[Sat, 08 Jan 2022]

- [This Week](#)
- [News in Focus](#)
- [Opinion](#)
- [Work](#)
- [Research](#)
- [Amendments & Corrections](#)

This Week

- **How researchers can help fight climate change in 2022 and beyond** [05 January 2022]
Editorial • COP26 energized the global effort to halt global warming. Research is now crucial to monitoring progress and creating solutions.
- **Molecular biologists: let's reconnect with nature** [03 January 2022]
World View • A New Year's resolution for bench scientists is to step out of the lab to study how life really works.

- EDITORIAL
- 05 January 2022

How researchers can help fight climate change in 2022 and beyond

COP26 energized the global effort to halt global warming. Research is now crucial to monitoring progress and creating solutions.



Devastating floods that hit Germany last July were made more likely by the warming climate. Credit: Christof Stache/AFP/Getty

Late last year, the major climate summit in Glasgow, UK — the 26th Conference of the Parties to the United Nations climate convention (COP26) — injected much-needed momentum into the political and business community in the fight to stop climate change. The year ahead represents an

opportunity for scientists of all stripes to offer up expertise and ensure that they have a voice in this monumental effort.

Science is already baked into the UN's formal climate agenda for 2022. In February, the Intergovernmental Panel on Climate Change (IPCC) is scheduled to release its assessment of the latest research into how climate warming is affecting people and ecosystems; a month later, the panel is set to provide an analysis of the options for curbing emissions and halting global warming. Combined with [last year's report on climate science](#), the governments of the world will have a solid review of the state-of-the-art of research on climate change. But the research community's work stretches far beyond the IPCC.

At the top of governments' climate agenda is innovation. Existing technologies such as wind and solar power, whose price has plummeted over the past decade, and more-efficient lighting, buildings and vehicles will help to reduce emissions. But if green energy is to push out fossil fuels and fulfil the rising demand for reliable power in low-income countries, scientists and engineers will be needed to solve a range of problems. These include finding ways to cut the price of grid-scale electricity storage and to address technical challenges that arise when integrating massive amounts of intermittent renewable energy. Research will also be required to provide a new generation of affordable vehicles powered by electricity and hydrogen, and low-carbon fuels for those that are harder to electrify, such as aircraft.

Even in the most optimistic scenarios, such clean-energy deployments are unlikely to be enough to enable countries to keep their climate commitments. More innovation will also be needed — for example, in the form of technologies that can pull carbon dioxide out of the atmosphere. These have yet to be tested and demonstrated at any significant scale. Governments and funders also need to support scientists in efforts to understand the safety and efficacy of various [controversial geoengineering technologies](#) — methods for artificially cooling the planet, such as the addition of particles to the stratosphere to reflect sunlight back into space — if only to determine whether there is sense in even contemplating such alternatives.

There are signs of renewed support for research and innovation in helping to address climate change. In Glasgow, 22 countries, as well as the European Commission (EC), announced plans to cooperate on innovation focused on greening cities, curbing industrial emissions, promoting CO₂ capture and developing renewable fuels, chemicals and materials. The EC has also announced efforts to drive new funds into demonstration projects to help commercialize low-carbon technologies. And China, currently the world's largest emitter of greenhouse gases, is [creating a vast research infrastructure](#) focused on technologies that will help to eliminate carbon emissions.

In the United States, under President Joe Biden, the Democrats have also made innovation a linchpin of efforts to address climate change. A bipartisan bill enacted in November will expand green-infrastructure investments, as well as providing nearly US\$42 billion for clean-energy research and development at the US Department of Energy over the next 5 years, roughly doubling the current budget, according to the Information Technology and Innovation Foundation, a think tank in Washington DC. Another \$550 billion for climate and clean-energy programmes is included in a larger budget bill that Democrats hope to pass this year. Economic modelling suggests that the spending surge could help to lower emissions in the coming decade while teeing up technologies that will be crucial to eliminating greenhouse-gas emissions in the latter half of the century.

In addition to enabling green innovation, scientists have an important part to play in evaluating climate policies and tracking commitments made by governments and businesses. Many of the initiatives that [gained traction at COP26](#) need science to succeed. That includes evaluating how [climate finance](#) — money that wealthy nations have committed to help low-income nations to curb emissions and cope with climate change — is spent. Research is also needed to understand the impacts of carbon offsets and carbon trading, for which new rules were agreed at COP26.

Climate science, too, must continue apace, helping governments and the public to understand the impact of climate change. From floods in Germany to fires in Australia, the evolving field of climate attribution has already made it clear that global warming is partly to blame for numerous tragedies. Attribution science will also feed into an ongoing geopolitical debate about who should pay for the rising costs of climate-related natural disasters, as

many low-income countries seek compensation from wealthy countries that are responsible for the bulk of the greenhouse-gas emissions so far.

These and other issues will be discussed again in November at COP27 in Sharm El-Sheikh, Egypt, where it will be crucial to make sure that everyone has a voice and that research supports climate monitoring and innovation everywhere, not just in richer nations.

A new agreement made at COP26 that requires governments to report annually on their climate progress should help to maintain pressure on them to act on climate change. But science and innovation will be equally important to driving ever-bolder climate policies.

Nature **601**, 7 (2022)

doi: <https://doi.org/10.1038/d41586-021-03817-4>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03817-4>

| [Section menu](#) | [Main menu](#) |

- WORLD VIEW
- 03 January 2022

Molecular biologists: let's reconnect with nature



A New Year's resolution for bench scientists is to step out of the lab to study how life really works.

- [Edith Heard](#) 0

Charles Darwin's voyage on HMS *Beagle* led to a treasure trove of observations: the behaviour of cuttlefish, a parasitic ichneumon wasp feasting inside live caterpillars, fossils of extinct giant sloths and 'mastodons'. The result, of course, was his theory of natural selection.

Darwin needed the complex natural world to inspire his theory. Today's molecular biologists usually focus on specific organisms in isolation and in carefully controlled environments that have as few variables as possible. To be sure, this has yielded impressive discoveries: RNA vaccines against

COVID-19, bioluminescence to monitor tumours, genomic sequencing to produce better crops, and more.

Molecular biologists, myself included, study the world at the smallest of scales: chromosomes, subcellular structures, proteins, metabolites. But, too often, this focus and our well-controlled labs deprive us of the fullest picture. We miss the range of genetic variation and how that mediates physiological and behavioural responses to environmental fluctuations, in individuals as well as in populations of organisms and across ecosystems. How Darwin's 'struggle for life' happens has been largely unexplored at the molecular level. In my view, molecular and cellular biologists must go back out into the world to study life in its natural context.

This year, the European Molecular Biology Laboratory (EMBL), of which I am director-general, is launching a programme called Molecules to Ecosystems that will pursue new ways of doing exactly this. For example, we will collaborate with ecologists, zoologists, environmental scientists and epidemiologists. And we plan to launch mobile labs with state-of-the-art molecular technologies to explore land–water interfaces across Europe — areas that harbour pollutants and pockets of antimicrobial resistance.

Molecular biologists are used to multidisciplinarity: we deploy X-ray physics and cryo-electron microscopy to study DNA, RNA and protein structures; chemistry to understand metabolic pathways; and informatics to analyse variation, including genomes and their epigenetic modifications. We can measure metabolites at the single-cell level, and we use fluorescence to identify cell organelles and macromolecules in multicellular systems. These data integrate genetic variation with phenotypic variation in individual cells, revealing associations that show how microbes (and other cells) function in different conditions.

For example, in oceanic plankton, some hosts of symbiotic microalgae trigger expansion of symbiont photosynthetic machinery to increase carbon production, which boosts ecological success in nutrient-poor waters. The ability to culture free-living symbionts mimicking the host microhabitat, and understand how their metabolism and morphology shift, could prompt fresh thinking around carbon fixation.

Technological advances will also allow researchers to explore organisms from volcanic coasts to the ocean depths. Sampling at sites that vary in pH, pollution, nutrients and salinity will offer insights into biodiversity and how natural and human-made changes influence it. Metabolic pathways are often at the heart of environmentally induced change. Such work can and should inspire metabolomics analysis to assess how toxins work, or prompt high-throughput biological imaging to catalogue morphological effects.

All of this means applying tools of basic research — in the wild and in the lab — to decipher molecular mechanisms that underlie organisms' variability and survival. Collaborative research at EMBL includes how nutrition affects a sea anemone's phenotype, the potential impact of nitrogen-fixing microbes called diazotrophs on oceans, and how antibiotics and other drugs alter the gut microbiome and human health.

Exploring 'life in context' is more urgent than ever, because that context is changing quickly. In 2009, US synthetic biologist James Collins pointed out a tragedy of timing: the first generation of scientists with the tools to address the dimensions of biodiversity on Earth might also be the last with the opportunity to do so.

Multidisciplinary collaboration enhances everyone's research. For instance, the artificial-intelligence system AlphaFold, which predicts 3D protein structures from amino-acid sequences, was the fruit of work involving structural biologists and deep-learning specialists. It gives researchers a head start on any problem involving protein structures. Likewise, cooperative efforts between machine-learning experts, biologists and clinicians are building tools to guide customized treatments for cancers.

Such successes mean that all scientists should work together to scrutinize genetic and environmental variation, and how human changes to the planet affect it. Molecular biologists can contribute to solving global challenges such as dealing with a changing climate, polluted water supplies and unsustainable food systems. For example, greater molecular understanding of microbial communities or crops could inspire alternative food sources, or biomaterials from fungal mycelium could replace conventional materials that contribute to carbon emissions.

Gaining insights about how organisms function under different conditions requires us to move beyond our usual comfort zone of lab research — and to look at areas such as arid soils and polluted rivers and cities, where there is a real ‘struggle for life’. These crucial sites need to be tackled.

As 2022 rolls in, I would like to ask every scientist reading this to consider how to interrogate the changing natural context in their research.

Nature **601**, 9 (2022)

doi: <https://doi.org/10.1038/d41586-021-03818-3>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03818-3>

| [Section menu](#) | [Main menu](#) |

News in Focus

- **[Afghanistan's academics despair months after Taliban takeover](#)** [17 December 2021]
News • Research has stalled, funds have evaporated and many scientists are still struggling to get out.
- **[Webb telescope blasts off successfully — launching a new era in astronomy](#)** [25 December 2021]
News • Hundreds of engineering steps must now take place as the observatory unfurls and travels to its new home.
- **[Journals adopt AI to spot duplicated images in manuscripts](#)** [21 December 2021]
News • A few publishers are using automated software to catch flaws in submitted papers.
- **[The science events to watch for in 2022](#)** [17 December 2021]
News • Omicron, Moon missions and particle physics are among the themes set to shape research in the coming year.
- **[Is precision public health the future — or a contradiction?](#)** [04 January 2022]
News Feature • Some public-health researchers are embracing data and technology to target small groups with precise health interventions. Others fear that these tactics could fail millions.

- NEWS
- 17 December 2021

Afghanistan's academics despair months after Taliban takeover

Research has stalled, funds have evaporated and many scientists are still struggling to get out.

- [Smriti Mallapaty](#)



People queue for supplies distributed by the World Food Program in Afghanistan as a Taliban fighter looks on. Credit: Hector Retamal/AFP/Getty

Four months after the Taliban took control of Afghanistan, academics say they feel marooned, [and abandoned](#) by the international community. With limited prospects for research, many scientists have left or are still trying to find routes out, so they can continue their careers.

Researchers say they have been stripped of their finances and academic freedoms and do not feel valued by the new government. Many continue to fear being persecuted for their international connections, ethnicity or gender — or because they have been critical of the government — and some say they have been threatened with death or retribution by the Taliban.

“The current government has had a completely destructive effect on research,” says Shohra Qaderi, who is from Afghanistan but is currently studying clinical science and public health at Shahid Beheshti University of Medical Sciences in Tehran. “Research is a form of freedom of speech for scholars, but this freedom of speech has been taken by the Taliban.”

Qaderi was due to return to a university in Afghanistan earlier this year, but decided to extend her stay in Iran because of the dire situation for researchers.

The Taliban seized power in mid-August as the last of the US-led coalition troops that had been supporting the previous government pulled out. But the broader international community has not yet recognized the current government. Billions of dollars in funding, assets and loans earmarked for government agencies and development and humanitarian services remain frozen. And research funding from international sources is no longer flowing.

Desperate situation

Some international donors are looking for ways to hand out funds through non-governmental channels, but that has taken time. On 10 December, donors agreed to release US\$280 million to the United Nations children’s agency UNICEF and the World Food Programme to directly deliver essential health services and food assistance to more than 10 million people in the country.

Understandably, “research is not really a priority”, says Jaffer Shah, a clinical researcher at Drexel University in Philadelphia, Pennsylvania, but who is based in New York City and collaborates with colleagues in Afghanistan.

The nation’s roughly 40 public universities and higher-education institutions officially remain open, but are effectively not functioning. Researchers say government officials have told staff to continue registering their attendance, but classes are not running. University staff members have not received their salaries since before August, and international funding for research projects has been put on hold. Female faculty members are still allowed to work, with some restrictions.

“We are just wasting our time here,” says a researcher who asked to remain anonymous. After the Taliban takeover, they were still weighing their options, but now they are considering leaving the country because their research has stalled and their teenage daughter has not been able to go to school.

In August, the Taliban banned co-education of male and female university students — and many public universities are still working out the practicalities of implementing segregated teaching. The Taliban government has also placed representatives at public universities. In one widely reported case, the government appointed a vice-chancellor for the prestigious Kabul University, but staff members deemed the candidate unqualified for the position, leading to widespread resignations. He was later replaced by another candidate.

Research stalled

Many of the country’s more than 120 private universities are still operating. But these institutions rely on student fees for funding, and many students are struggling to pay, have left the country or have abandoned their studies. Academics say salaries and research funding have been slashed.

“Our research activities have stopped,” says Sayed Hamid Mousavi, a medical researcher and director of research at the private Kateb University

in Kabul, who says he hasn't been paid in four months. He and others are trying to keep busy with desk-based work, such as writing up retrospective analyses of data; laboratory-based research has stalled and it is hard to gather new data.

Reports have also surfaced of academics being mistreated at both public and private universities. Ramin Mansoory, a legal scholar who was an adviser to the governor of Afghanistan's Balkh province before he fled to Poland in August, has heard reports of faculty members from minority ethnic groups being fired and replaced with those of Pashtun ethnicity, the group to which many Taliban members belong. There have also been some news articles reporting the kidnapping and killing of academics from minority ethnic groups.

In September, Qaderi co-authored a [short correspondence](#) in *Nature* on how the Taliban's takeover could imperil the health of pregnant women and new mothers. The article listed a university in Afghanistan as one of her affiliations — but after it was published, some of her colleagues received threats and asked her to remove the affiliation.

Leaving the country

In August, *Nature* reported on the [tide of researchers](#) seeking refuge abroad, reflected by a record number of applications to programmes that arrange placements for scholars at foreign universities. Since then, [applications have continued to rise](#), mostly from scholars still stranded in Afghanistan, but only a handful have found institutions to take them.

"Everyone is trying to get out," says Eraj Haidari, a visiting researcher in law at City, University of London. Haidari is one of only four individuals to have received a UK university fellowship supported by the Council for At-Risk Academics, which is based in London. Executive director Stephen Wordsworth says that the council has received some 540 applications from scholars from Afghanistan.

Scholars at Risk, a similar organization based in New York City, has received more than 1,300 applications. Only 20 applicants have so far been

connected with institutions in the United States and Europe, and the organization is seeking spots for another 35.

The Institute of International Education's Scholar Rescue Fund in New York City has also been flooded with a record number of applications. Requests have risen since August — “a trend we only expect to continue,” says director James King.

But even when researchers do secure positions, getting out of Afghanistan is tricky. Many embassies in the country remain closed, which means that scholars often first have to get to neighbouring countries, such as Pakistan and Iran, to process their papers.

Researchers say they have little hope that the situation in Afghanistan will improve in the near future. “We do not have the outside world’s support and the number of researchers and scholars has decreased significantly. We are facing a brain drain,” says Qaderi.

Nature **601**, 11-12 (2022)

doi: <https://doi.org/10.1038/d41586-021-03774-y>

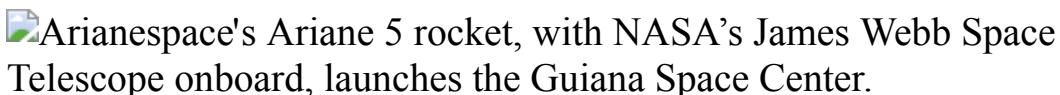
This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03774-y>.

- NEWS
- 25 December 2021

Webb telescope blasts off successfully — launching a new era in astronomy

Hundreds of engineering steps must now take place as the observatory unfurls and travels to its new home.

- [Alexandra Witze](#)



The James Webb Space Telescope launched at 9:20 a.m. local time from the European Space Agency's spaceport in Kourou, French Guiana. Credit: NASA/NASA TV via Reuters

The James Webb Space Telescope — humanity's biggest gamble yet in its quest to probe the Universe — soared into space on 25 December, marking the culmination of decades of work by astronomers around the world. But for Webb to begin a new era in astronomy, as many scientists hope it will, hundreds of complex engineering steps will have to go off without a hitch in the coming days and weeks.

“Now the hard part starts,” says John Grunsfeld, an astrophysicist and former astronaut and head of science for NASA.

The US\$10-billion Webb is the [most complicated and expensive space observatory in history](#), and the successor to NASA's Hubble Space Telescope, which has studied the Universe since 1990. Following its launch, Webb will now embark on the riskiest part of its mission — deploying all

the parts required for its enormous mirror to peer deep into the cosmos, back towards the dawn of time.

Not until all the equipment works and the first scientific observations have been made, probably in July, will astronomers be able to relax. Until then, “there’s going to be a lot of nervousness”, says Heidi Hammel, an interdisciplinary scientist for Webb and vice-president for science at the Association of Universities for Research in Astronomy in Washington DC.

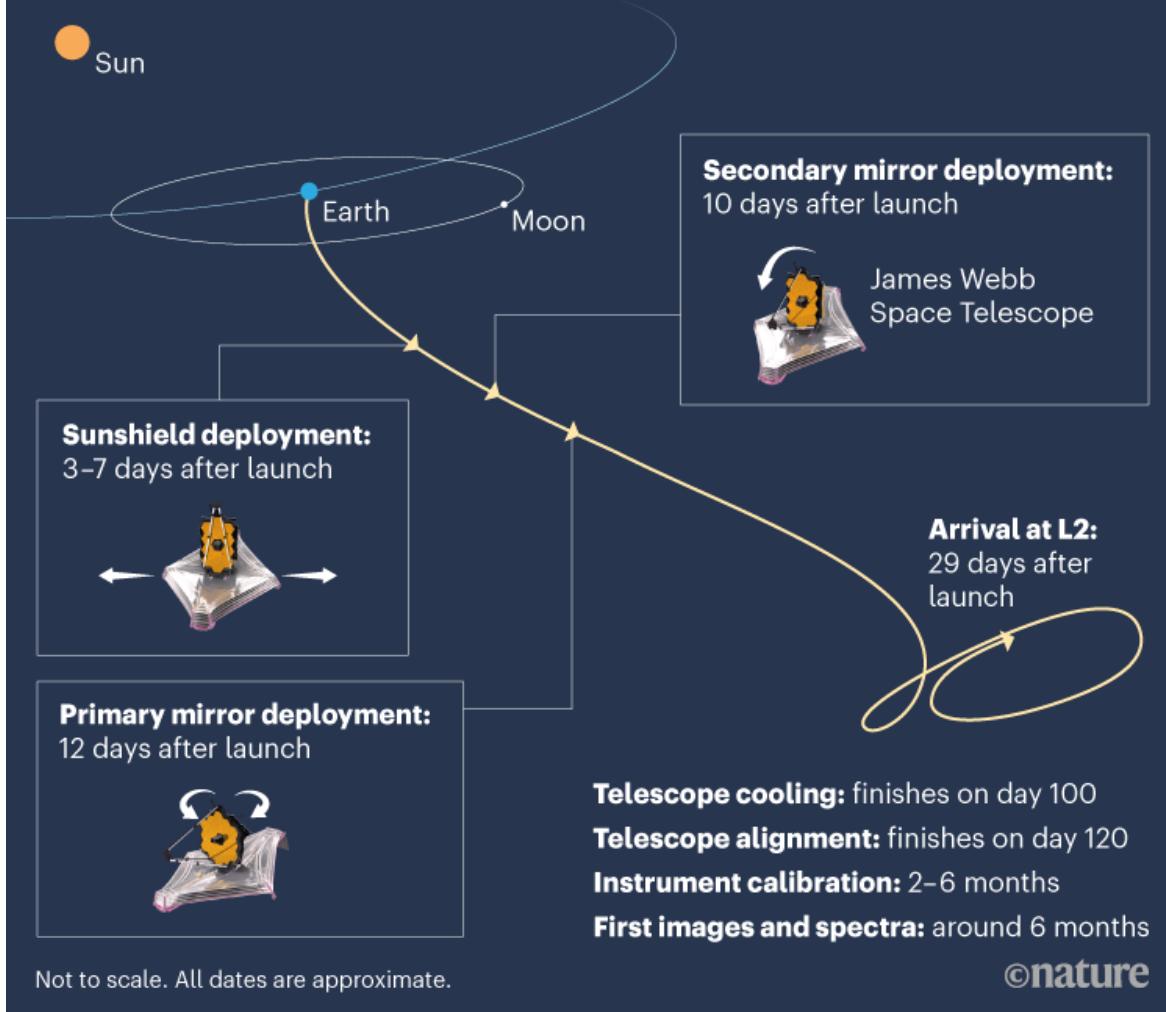
The NASA-built Webb launched at 9.20 a.m. local time from the Kourou spaceport in French Guiana, on an Ariane 5 rocket provided by the European Space Agency (ESA). The project’s third international partner is the Canadian Space Agency.

“What an emotional day,” said Thomas Zurbuchen, NASA’s head of science, on a webcast from the launch site. “It’s the beginning of one of the most amazing missions that humanity has conceived.”

“I’m feeling very emotional right now after seeing something we’ve all waited for for so long finally happen,” adds Jeyhan Kartaltepe, an astronomer at the Rochester Institute of Technology in New York who has been awarded observing time on Webb. “I am super grateful for everyone that worked so hard to make it a success.”

WEBB'S JOURNEY

After launch, the James Webb Space Telescope will begin unfolding its equipment and deploying its mirrors as it travels to its final home at the point L2, 1.5 million kilometres from Earth.



Telescope images: NASA

Setting a course

The Ariane 5 carried Webb on an apparently flawless trajectory into space, which conserves more fuel for the telescope to use for science in the coming years. After separating from the launch vehicle 27 minutes after launch, Webb unfolded its solar panels, a crucial step that allowed electrical power to begin flowing. “That was half an hour on the edge of my seat,” says

Grunsfeld. Hours later, it was expected to burn its engines to set it on course toward its ultimate destination, a position in space known as the second Lagrange point, or L2 (see ‘Webb’s journey’). There, 1.5 million kilometres from Earth, it will always be on the opposite side of the planet from the Sun, looking into the dark reaches of outer space with its sensitive optics shielded from sunlight.

The journey to L2 will take 29 days, with more than 300 ‘points of failure’ at which something could go wrong. The entire process is something like a butterfly emerging from its chrysalis, says Günther Hasinger, ESA’s director of science — a very expensive and very complex butterfly that is three storeys tall.

First and most crucial is the deployment of Webb’s kite-shaped sunshield, which is the size of a tennis court. The shield’s job is to shade Webb from radiation, and cool the environment from 110 °C on the Sun-facing side to –235 °C on the shaded side. Webb requires frigid temperatures for its optics to be able to pick up the glimmers of distant galaxies and other cosmic objects in infrared wavelengths.



The telescope launched on a European Ariane 5 rocket. Credit: Bill Ingalls/NASA

“The unique aspect about Webb is that it is a cold telescope,” says Hasinger. If the sunshield does not deploy properly, Webb’s science will be severely degraded.

Three days after launch, two rectangular pallets are meant to unfold on either side of the observatory. Over the course of four more days, the pallets will open to reveal the five membrane-like layers of the sunshield, and then stretch them and fix them in place, like tightening a sheet over a mattress.

Mirror alignment

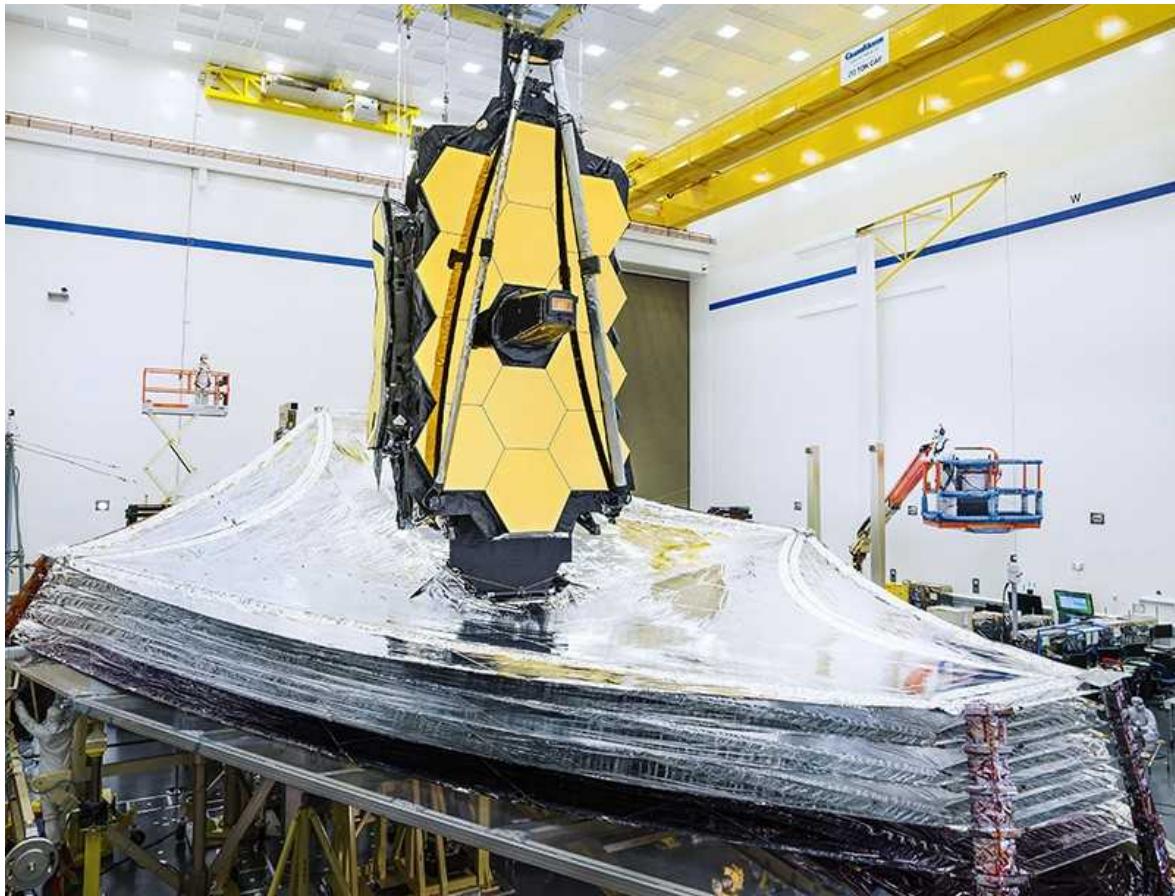
Ten days after launch, if all goes well, Webb will move its small secondary mirror to face its giant primary mirror, which will still be folded up. Two days later, the primary mirror will begin to swing two hinged sections into place to create the complete 6.5-metre-wide mirror. At that point, the mirror’s 18 hexagonal segments, made of beryllium and coated in gold, will resemble a gigantic glimmering honeycomb — and Webb will become a true telescope, Hammel says, because it will be able to capture light.

Roughly two weeks later, Webb should reach its final location at L2. That is too far from Earth for astronauts to visit and fix the telescope if anything goes wrong, as they did after Hubble launched with flawed optics.

(Astronauts, including Grunsfeld, ultimately flew to Hubble five times to upgrade its scientific instruments and keep it as a world-class observatory.) But if the sunshield and the mirrors deploy properly, then the worst is behind Webb. From then on, it’s a matter of aligning the mirrors and calibrating the instruments. Those steps are complicated but have been done before, on ground-based telescope mirrors.

One reason it will be six months before science can begin is that the telescope needs time to cool down to its operating temperature. Only about four months after launch will the primary mirror segments be sufficiently cooled and aligned to assume the correct shape. Mission engineers will then begin calibrating the instruments.

And one of the telescope's four instruments, which will work in mid-infrared wavelengths, requires another cooler to get to -266°C , just 7°C above absolute zero. Once it reaches that temperature, Webb should be 100 times more sensitive than Hubble.



The telescope's beryllium and gold mirror sits in the middle of a five-layered sunshield.Credit: Chris Gunn/NASA

Infrared eyes

Webb's infrared vision will allow it to peer back more than 13.5 billion years, towards distant stars and galaxies whose light has been stretched to infrared wavelengths by the expansion of the Universe. It will also be able to penetrate dust-swathed regions such as the places where stars are born, and to probe the atmospheres of planets beyond the Solar System. "It's designed to answer outstanding questions in all fields of astrophysics," says astronomer Antonella Nota, ESA's project scientist for Webb.

If everything goes as planned over the next six months, mission scientists will be under intense pressure to release stunning images and data from Webb as soon as possible. A small committee of astronomers at the Space Telescope Science Institute in Baltimore, Maryland, which operates Webb, has drawn up a secret list of which objects to observe first. The first tranche of results will probably include spectacular images of and data on planets, stars and galaxies, to show off the telescope's capabilities. After that, science will begin in earnest for the other astronomers who are queuing up to use Webb.

For now, they can only watch and wait to see whether the telescope works as its designers intend. "I know we have done everything we can," says Hammel.

Nature **601**, 12-13 (2022)

doi: <https://doi.org/10.1038/d41586-021-03655-4>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03655-4>

| [Section menu](#) | [Main menu](#) |

- NEWS
- 21 December 2021

Journals adopt AI to spot duplicated images in manuscripts

A few publishers are using automated software to catch flaws in submitted papers.

- [Richard Van Noorden](#)



AI software that spots duplicated images in research papers can work faster and on a larger scale than manual checkers — but still needs editorial oversight. Credit: Laurence Dutton/Getty

Just before a study appears in any of ten journals published by the American Association for Cancer Research (AACR), it undergoes an unusual extra check. Since January 2021, the AACR has been using artificial intelligence

(AI) software on all manuscripts it has provisionally accepted after peer review. The aim is to automatically alert editors to duplicated images, including those in which parts have been rotated, filtered, flipped or stretched.

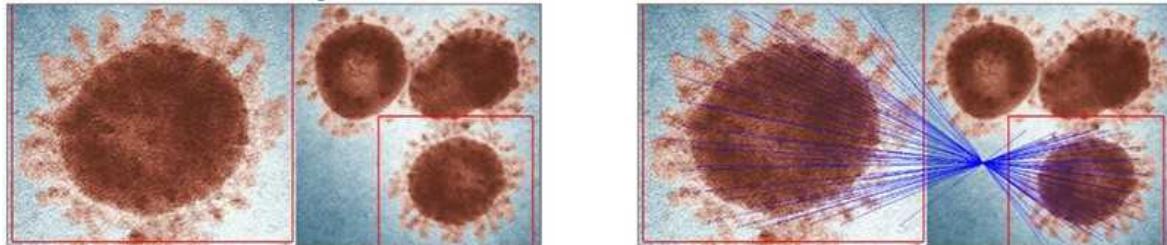
The AACR is an early adopter in what could become a trend. Hoping to avoid publishing papers with images that have been doctored — whether because of outright fraud or inappropriate attempts to beautify findings — many journals have hired people to manually scan submitted manuscripts for issues, often using software to help check what they find. But *Nature* has learnt that in the past year, at least four publishers have started automating the process by relying on AI software to spot duplications and partial duplications before manuscripts are published.

The AACR tried numerous software products before it settled on a service from Proofig, a firm in Rehovot, Israel, says Daniel Evanko, director of journal operations at the association in Philadelphia, Pennsylvania. “We’re very happy with it,” he adds. He hopes the screening will aid researchers and reduce problems after publication.

Professional editors are still needed to decide what to do when the software flags images. If data sets are deliberately shown twice — with explanations — then repeated images might be appropriate, for instance. And some duplications might be simple copy-and-paste errors during manuscript assembly, rather than fraud. All this can be resolved only with discussions between editors and authors. Now that AI is getting sufficiently effective and low-cost, however, specialists say a wave of automated image-checking assistants could sweep through the scientific publishing industry in the next few years, much as using software to check manuscripts for plagiarism became routine a decade ago. Publishing-industry groups also say they are exploring ways to compare images in manuscripts across journals.

Other image-integrity experts welcome the trend, but caution that there has been no public comparison of the various software products, and that automated checks might throw up too many false positives or miss some kinds of manipulation. In the long term, a reliance on software screening might also push fraudsters to use AI to dupe software, much as some tweak text to evade plagiarism screening. “I am concerned that we are entering an

arms race with AI-based tech that can lead to deepfakes that will be impossible to find,” says Bernd Pulverer, chief editor of *EMBO Reports* in Heidelberg, Germany.



This constructed example from the image-checking software firm Proofig shows how its program compares parts of images (red rectangles, left) and flags up identical parts even in stretched or rotated pictures. The blue lines indicate that the AI sees hundreds of identical features. Credit: Adapted from CDC/Proofig

Software's moment?

Researchers have been developing image-checking AI for years because of concerns about errors or fraud — which are probably polluting the scientific literature to a much greater extent than the limited numbers of retractions and corrections suggest. In 2016, a manual analysis¹ of around 20,000 biomedical papers led by microbiologist Elisabeth Bik, a consultant image analyst in California, suggested that as many as 4% might contain problematic image duplications. (Typically only about 1% of papers receive corrections each year, and many fewer are retracted.)

“I am aware of around 20 people working on developing software for image checking,” says Mike Rossner, who runs the consultancy firm Image Data Integrity in San Francisco, California, and introduced the first manual screening of manuscripts at the *Journal of Cell Biology*, 20 years ago. Last year, publishers joined together to form a working group to set standards for software that screens papers for image problems; the group issued guidelines this year on how editors should tackle doctored images, but hasn’t yet produced guidance on software.

Several academic groups and companies have told *Nature* that journals and government agencies are trialling their software, but Proofig is the first to name clients publicly. Besides the AACR, the American Society for Clinical Investigation started using Proofig's software for manuscripts in the *Journal of Clinical Investigation (JCI)* and *JCI Insight* in July, says Sarah Jackson, executive editor of those journals in Ann Arbor, Michigan. And SAGE Publishing adopted the software in October for five of its life-sciences journals, says Helen King, head of transformation at SAGE in London.

Proofig's software extracts images from papers and compares them in pairs to find common features, including partial duplications. A typical paper is checked in a minute or two; the software can also correct for tricky issues such as the compression artefacts that can arise when high-resolution raw data are compressed into smaller files, says Dror Kolodkin-Gal, the firm's founder. "The computer has an advantage over human vision," he says. "Not only does a computer not get tired and run much faster, but it is also not affected by manipulations in size, location, orientation, overlap, partial duplication and combinations of these."

The cost of image checking is much higher than that of plagiarism checking, which specialists say runs to less than US\$1 per paper. Kolodkin-Gal declined to discuss pricing in detail, but said that contracts with publishers tend to charge on the basis of the number of images in a paper, but also depend on the volume of manuscripts. He says they equate to per-paper charges "closer to tens of dollars than hundreds of dollars".

At the *JCI*, says Jackson, the software picks up more problems than did previous manual reviews by staff members. But staff are still essential to check Proofig's output, and it was important that the journal already had a system of procedures for dealing with various image concerns. "We really feel that rigorous data is an absolute hallmark of our journals. We have decided this is worth the time and money," Jackson says. At the AACR, Evanko says many authors are happy that duplication errors are brought to their attention before publication.

Meanwhile, the publisher Frontiers, in Lausanne, Switzerland, has developed its own image-checking software as part of a system of automated checks called AIRA (Artificial Intelligence Review Assistant). Since August

2020, an internal research-integrity team has been using AIRA to run image checks on all submitted manuscripts, a spokesperson says. The majority of papers that it flags up don't actually have problems: only around 10% require follow-up from the integrity team. (Frontiers declined to say what fraction of papers AIRA flags.)

Image-integrity specialists including Bik and Rossner say they haven't tried AIRA or Proofig themselves, and that it is hard to evaluate software products that haven't been publicly compared using standardized tests. Rossner adds that it's also important to detect image manipulation apart from duplication, such as removing or cropping out parts of an image, and other photoshopping. "The software may be a useful supplement to visual screening, but it may not be a replacement in its current form," he says.

"I am convinced, though, that eventually this will become the standard in manuscript screening," adds Bik.

Industry caution

Publishers that haven't yet adopted AI image screening cite cost and reliability concerns — although some are working on their own AIs. A spokesperson for the publisher PLOS says that it is "eagerly" monitoring progress on tools that can "reliably identify common image-integrity issues and that could be applied at scale". Elsevier says it is "still testing" software, although it notes that some of its journals screen all accepted papers before publication, checking for concerns around images "using a combination of software tools and manual analysis".

In April 2020, Wiley introduced an image-screening service for provisionally accepted manuscripts, now used by more than 120 journals, but this is currently manual screening aided by software, a spokesperson says. And Springer Nature, which publishes *Nature*, says that it is assessing some external tools, while collating data to train its own software that will "combine complementary AI and human elements to identify problematic images". (*Nature*'s news team is editorially independent of its publisher.)

Pulverer says that EMBO Press still mostly uses manual screening because he's not yet convinced by the cost–benefit ratio of the commercial offerings, and because he is part of the cross-publisher working group that is still defining criteria for software. “I have no doubt that we will have high-level tools before long,” he says.

Pulverer worries that fraudsters might learn how the software works and use AI to make fake images that neither people nor software can detect.

Although no one has yet shown that such images are appearing in research papers, one preprint² posted on bioRxiv last year suggested that it was possible to make fake versions of biological images such as western blots that were indistinguishable from real data. But researchers are working on the problem: computer scientist Edward Delp at Purdue University in West Lafayette, Indiana, leads a team that is spotting media faked by AIs, in a [programme funded by the US Defense Advanced Research Projects Agency](#), and is focusing on fake biological imagery such as microscope images and X-rays. He says his team “has one of the best” sets of detectors for GANs, or generative adversarial networks — a way of pitting AIs against each other to create realistic images. A paper describing his system is under review.

Cross-journal image checks

For the moment, AI image checking is generally done within a manuscript, not across many papers, which would make it increasingly computationally intensive. But commercial and academic software developers say that this is technically feasible. Computer scientist Daniel Acuña at Syracuse University in New York last year [ran his software on thousands of COVID-19 preprints to find duplications](#).

Crossref, a US-based non-profit collaboration of more than 15,000 organizations that organizes plagiarism checking across papers, among other things, is currently running a survey to ask its members about their concerns on doctored images, what software they are using and whether a “cross-publisher service” that could share images could be viable and helpful, says Bryan Vickery, Crossref’s director of product in London.

And in December, STM Solutions — a subsidiary of the STM, an industry group for scholarly publishers in Oxford, UK — announced that it was working on a “cloud-based environment” to help publishers collaborate “to check submitted articles for research integrity issues” — while maintaining privacy and confidentiality. Detecting image manipulation, duplication and plagiarism across journals is “high on our road map”, says Matt McKay, an STM spokesperson.

Nature **601**, 14-15 (2022)

doi: <https://doi.org/10.1038/d41586-021-03807-6>

References

1. 1.

Bik, E. M., Casadevall, A. & Fang, F. C. *mBio* **7**, e00809-16 (2016).

2. 2.

Qi, C., Zhang, J. & Luo, P. Preprint at bioRxiv
<https://doi.org/10.1101/2020.11.24.395319> (2021).

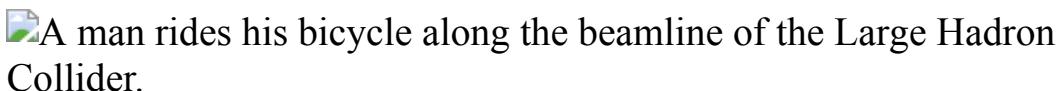
This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03807-6>

- NEWS
- 17 December 2021

The science events to watch for in 2022

Omicron, Moon missions and particle physics are among the themes set to shape research in the coming year.

- [Davide Castelvecchi](#)



The beamline of the Large Hadron Collider, which will restart operations in 2022. Credit: Valentin Flauraud/AFP/Getty

COVID continues

As the world enters the third year of the COVID-19 pandemic, with no end in sight, an immediate challenge is to better understand the impact of Omicron — the [fast-spreading coronavirus variant](#) that was first spotted at the end of November — and the [threat it poses](#). Early results indicate that [vaccines are less effective against Omicron](#); scientists are still racing to find out more about the [severity of the disease it causes](#).

In 2022, researchers and public-health authorities will also continue to monitor the [rise of new SARS-CoV-2 variants](#), as well as the [long-term effects](#) on people who have recovered from infection.

Wealthy countries have begun giving their populations [booster shots of existing vaccines](#), and these roll-outs are likely to continue amid concerns about Omicron. But nearly half the world's population has not yet received a

single dose of a vaccine. One big question is whether pharmaceutical companies will waive patents or take other steps to help make their vaccines more affordable for lower-income countries, to begin filling the [huge gap in global coverage](#). Meanwhile, discussions about the [origins of the virus](#) will probably continue. The World Health Organization has renewed its efforts to solve the riddle, by appointing [a team of 26 scientists](#).

Vaccines upgraded

Vaccine developers have set their sights on the next generation of vaccines designed to protect against the rapidly evolving coronavirus. Next year could see the development of [messenger RNA vaccines](#) that are targeted to specific variants, and some public-health officials are hoping for an increased role for vaccines using other technologies. [Protein-based vaccines](#) are a more conventional kind of immunization — some have been used for decades against diseases including hepatitis and shingles — and in 2021 they have shown promise in phase III COVID–19 clinical trials. [Vaccines based on DNA](#) are cheaper to manufacture than mRNA vaccines and do not require cold storage, so could be good alternative for lower-income countries.

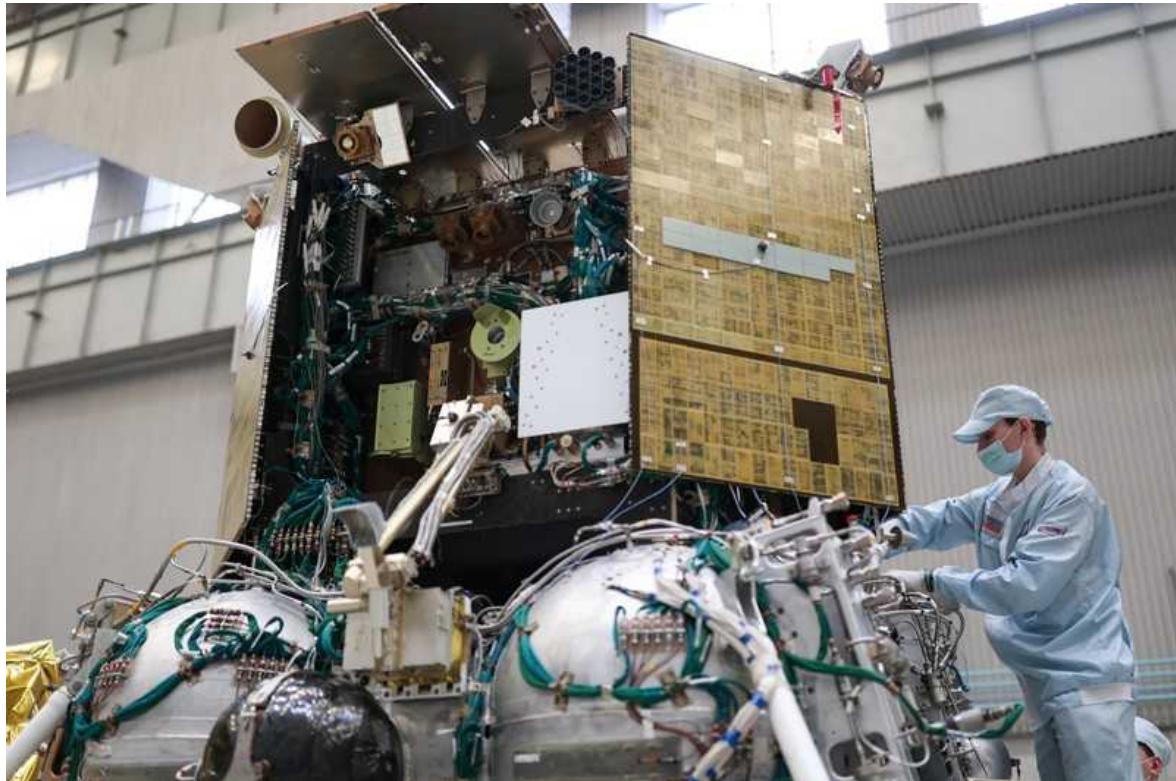
Progress on vaccines is also expected for other major viruses and diseases, including [HIV](#), [malaria](#) and [Lyme disease](#).

Big physics bonanza

After a multi-year shutdown and extensive maintenance work, the Large Hadron Collider is [scheduled to restart operations](#) at CERN, the European particle-physics laboratory outside Geneva, Switzerland, in June. The LHC's major experiments ATLAS and CMS were upgraded and expanded with additional layers of detector components. This will enable them to collect more data from the 40 million collisions of protons each of them produces every second.

And after their own upgrades, the world's four gravitational-wave detectors — one in Japan, one in Italy and two in the United States — will [begin a new observing run](#) in December.

At Michigan State University in East Lansing, the [Facility for Rare Isotope Beams](#) is expected to start operations in early 2022. The US\$730 million multi-stage accelerator aims to synthesize thousands of new isotopes of known elements, and it will investigate nuclear structure and the physics of neutron stars and supernova explosions.



Russia's Luna 25 lander will begin its journey to the Moon next year.Credit: Sergei Bobylev/TASS/Getty

Moon missions

A veritable armada of orbiters and landers from space agencies and private companies is scheduled to leave for the Moon in 2022. NASA will launch the [Artemis I](#) orbiter in the first test of the long-overdue launch system that is intended eventually to take astronauts back to the surface of the Moon. And the agency's [CAPSTONE](#) orbiter will conduct experiments in preparation for [the Gateway](#), the first space station to orbit the Moon.

India's third lunar mission, Chandrayaan-3, aims to be its first to make a soft landing (one that doesn't damage the craft) and will carry its own rover. Japan will also attempt its first soft landing on the Moon, with the [SLIM mission](#), and Russia is aiming to revive the glory of the Soviet lunar programme with the [Luna 25 lander](#). The [Korea Pathfinder Lunar Orbiter](#) will inaugurate South Korea's own Moon exploration.

On the private side, Tokyo-based company ispace is launching the Hakuto-R lander, which will carry the [United Arab Emirates' Rashid Moon rover](#). Two US companies, Astrobotic Technology in Pittsburgh, Pennsylvania, and Intuitive Machines in Houston, Texas, are readying probes that will carry NASA instruments to the lunar surface.

To Mars and the stars

Another epic space journey to watch will be the joint Russian–European [ExoMars](#) mission, which is scheduled to blast off in September and will carry the European Space Agency's Rosalind Franklin rover to Mars, where it will search for signs of past life. The launch was originally scheduled for 2020, [but was delayed](#) partly because of issues with the parachutes needed to touch down safely.

China also plans to complete its [space station](#), [Tiangong](#), and has lined up more than 1,000 experiments for it, ranging from astronomical and Earth observation to the effects of microgravity and cosmic radiation on bacterial growth.



Climate change could exacerbate droughts, such as the one that dried up this lake in California. Credit: David Swanson/Reuters

Climate action

Energized by this year's [COP26 summit](#) in Glasgow, UK, delegates from around the world will converge on Sharm El-Sheikh, Egypt, in November 2022 for COP27, another round of United Nations climate talks. Countries are expected to come up with climate commitments consistent with the 2015 Paris agreement goal of limiting global warming to well below 2 °C above pre-industrial temperatures. In the meantime, researchers will be monitoring greenhouse-gas emissions following pledges made at COP26 — which included promises to reduce the use of coal and cut methane emissions. After a pandemic-induced dip in 2020, [carbon emissions have rebounded in 2021](#).

Push to save biodiversity

Countries are working on a new set of targets to slow down the [loss of biological diversity](#). The Aichi Biodiversity Targets, established in 2010, were mostly missed by their 2020 deadline. The next meeting of parties to the [UN Convention on Biological Diversity](#) — originally planned for 2020 — is scheduled to take place in Kunming, China, from 25 April to 8 May, but concern over the Omicron variant of SARS-CoV-2 might scupper those plans yet again. Habitat loss and other factors linked to human activity have put an estimated one million plant and animal species at risk of extinction.

Nature **601**, 16–17 (2022)

doi: <https://doi.org/10.1038/d41586-021-03772-0>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03772-0>

| [Section menu](#) | [Main menu](#) |

- NEWS FEATURE
- 04 January 2022

Is precision public health the future — or a contradiction?

Some public-health researchers are embracing data and technology to target small groups with precise health interventions. Others fear that these tactics could fail millions.

- [Carrie Arnold](#) ⁰



A man has a nasal swab at a mobile COVID-19 testing site in Manhattan, New York City. Credit: Todd Heisler/*The New York Times*/Redux/eyevine

From their offices in a high-rise building in Queens, epidemiologist Sharon Greene and her colleagues watched the COVID-19 pandemic sweep through New York City in April 2020. Using an open-source data-analytics program called SaTScan, her team mapped outbreaks as they unfolded across individual neighbourhoods, almost in real time. This sophisticated approach relied on detailed data from hospitals and laboratories, and showed that the virus wasn't affecting all New Yorkers equally. That knowledge helped Greene's team at the New York City Department of Health and Mental Hygiene to distribute testing resources and protective gear such as masks and gloves to the right places.

It was a different approach from New York City's typical pandemic response plan, which advised largely blanket policies such as lockdowns and mass testing.

“Instead of just parking a testing van somewhere in an affected zip code, we can park it at an intersection in the middle of the cluster,” Greene says. “It’s hyper-local public health.” By the middle of the year, cases in the city began to drop.

The tech-centric, targeted approach used by Greene and other epidemiologists to address COVID-19 is part of a burgeoning field known as precision public health. The concept is a modernization of the 150-year-old field of epidemiology, similar to how precision medicine has transformed health care, says Muin Khoury, director of the Office of Genomics and Precision Public Health at the US Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, and one of the idea’s biggest advocates.

The definition of precision public health is sprawling and variable: for most researchers in the field it includes a sweep of data-driven techniques, such as sequencing pathogens to detect outbreaks and turbo-charging data collection to monitor harmful environmental exposures. It also encompasses an ambition to target interventions to specific people who need them.

For Caitlin Allen, a PhD student of public health at Emory University in Atlanta, who organized a [meeting on precision public health](#) in October last year, the kernel of the idea is simple. “You’re doing all the things you normally do in public health, but the unique aspect is that we’re using big data and predictive analytics to be more targeted and tailored in these efforts,” she says. The concept promises to save money and lives by targeting interventions to the right people.

To Sandro Galea, an epidemiologist and dean of the Boston University School of Public Health, Massachusetts, and others, however, it sounds too good to be true. “We’re all looking for the silver bullet, but there isn’t one,” says Galea.

The debate over the merits of precision public health has typically taken place in the pages of academic journals. But funders are putting hundreds of millions of dollars behind precision-public-health initiatives, and some researchers worry about the implications for conventional public health. Spending on public health is already sliding: although national health

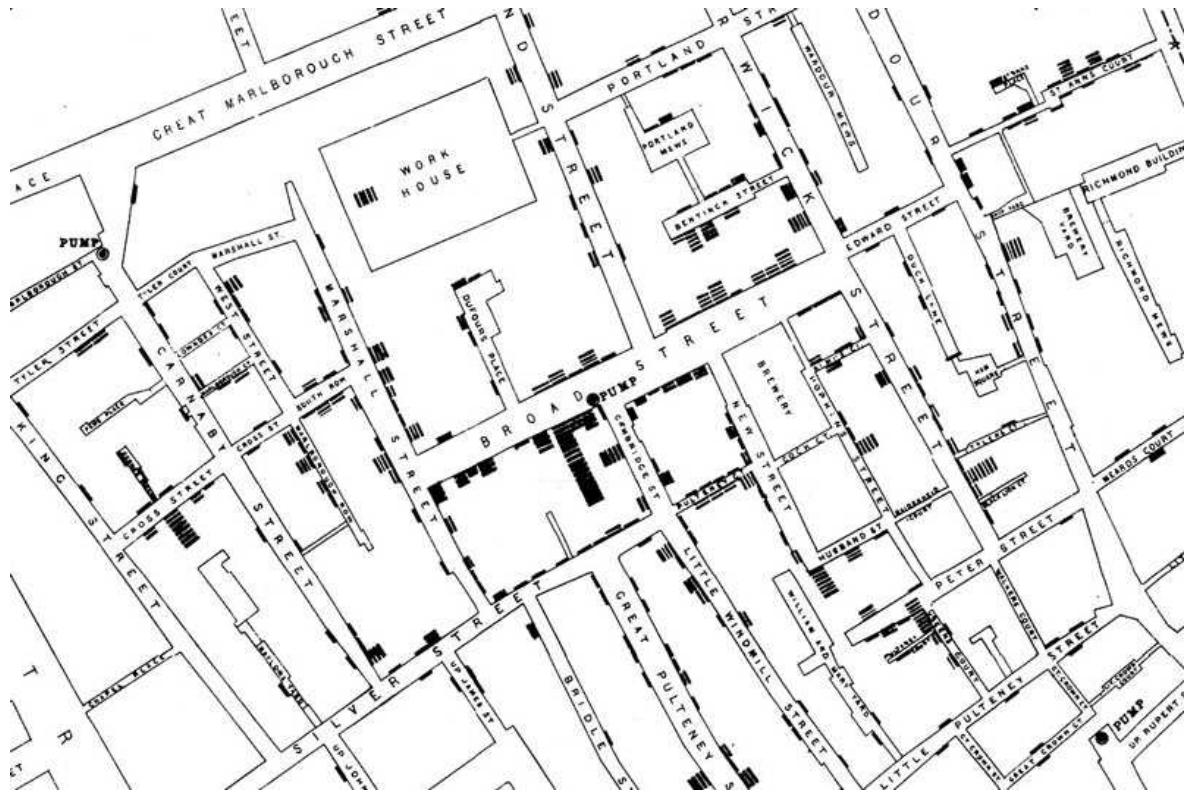
expenditure in the United States grew by 4.3% from 2008 to 2018, researchers found no change in public-health spending¹.

Galea is concerned that the precision approach is diverting attention away from regular public health. “I worry that this is becoming the great sucking sound where we focus all our energy on technological approaches and we don’t focus on more foundational issues that will make a difference in the lives of millions,” Galea says.

Think local

For John Quackenbush, a biostatistician at the Harvard School of Public Health in Boston, the push for precision public health has existed since the birth of epidemiology.

In the 1850s, citizens of London stared down a different outbreak: cholera. The disease killed millions of people in the nineteenth century. In 1854, a string of cases appeared in Soho, in central London. Physician John Snow worked just a few streets away. Snow went from door to door and began to plot cholera cases on a map. He found that people who got their water from a pump on Broad Street were much more likely to develop cholera.



Physician John Snow helped to track cholera cases (black) with precision during an outbreak in London in 1854. Credit: ‘On the mode of Communication of Cholera’, John Snow 1854

Snow’s statistics might seem pedestrian to modern epidemiologists, but they were state-of-the-art in 1854. If Snow had had access to SaTScan, he would have used that, Quackenbush says, and maybe brought the outbreak to a halt much more quickly.

“We just have access to unprecedented quantities of data,” Quackenbush says.

More than a century ago, the first municipal public-health departments in cities such as London and New York aimed to improve the health of large swathes of the population by building sewage systems, removing rubbish, purifying drinking water and collecting basic information on births and deaths. The efforts were broad both because all residents would benefit from these interventions, and because early sanitation workers had little information on where to target their resources.

As time passed, epidemiology grew more sophisticated and precise as researchers were able to home in on specific clusters of disease and ever more detailed risk factors for future illnesses².

In that sense, the concept of precision public health is not new — but the phrase is recent. Khoury coined it in a [March 2015 blog post](#) as he watched the genomics revolution take hold in medicine. He wanted to see that same energy spill over into public health. To some extent, it already had. In 1996, the CDC launched [PulseNet](#), which used DNA fingerprinting of bacteria that caused food poisoning to detect large, diffuse outbreaks across county and state lines. Large outbreaks of *Salmonella*, *Listeria* and *Escherichia coli* were [investigated and halted](#) when they might otherwise have gone unnoticed.

Khoury saw similar potential in other genetic technologies. As a paediatrician and geneticist, he had spent years studying a genetic condition called familial hypercholesterolaemia (FH), which causes extremely high cholesterol levels and can lead to heart attack and stroke. Despite being readily recognized with basic screening tests and a proven therapy (high-dose statins) that reduces the risk of heart disease by 80%, many of those with FH go undiagnosed. With less than 1% of the population affected, public-health investments to screen everyone would be impractical and wasteful.

To Khoury and others, FH seemed like the perfect place for precision public health to shine. A US team is piloting a machine-learning method that scans health records (including standard blood cholesterol measurements) and identifies those likely to have FH³. This strategy could help to ensure that individuals with FH get appropriate screening and treatment without wasting resources on the 99% of people without FH.

Soon after Khoury's initial blog post, the term start popping up everywhere. In June 2016, the University of California, San Francisco, hosted the first precision-public-health conference. In the same year, the Zika virus swept across the Western Hemisphere. Health officials in Miami, Florida, began preparing for it to reach the United States. Using a detailed geographic information system to map locally acquired cases, they were [able to target](#)

[mosquitoes](#) with insecticides sprayed across just two blocks of Miami, rather than dousing whole neighbourhoods or even the entire city.



An official in Miami, Florida, sprays insecticide in targeted neighbourhoods to kill mosquitoes and stem the spread of Zika virus in 2016. Credit: Joe Raedle/Getty

Precision approaches are taking off in many more guises. The Bill & Melinda Gates Foundation has given a total of \$271 million to the Child Health and Mortality Prevention Surveillance network led by Emory University to map out areas of greatest maternal-health problems and childhood malnutrition across Africa and Asia. The goal is to help governments, charities and other advocates to create evidence-based policies to target childhood disease where it is most severe.

And in September 2019, the Rockefeller Foundation launched a [\\$100-million Precision Public Health Initiative](#) dedicated to using predictive analytics to prevent health threats, and exploiting big data to address the social factors that lead to poor health, such as discrimination and poverty. During the pandemic, the foundation created an interactive dashboard,

containing detailed statistics on various COVID-19 testing strategies, to help educators determine the safest way to reopen schools.

The idea seemed so promising that advocates began to wonder whether precision public health had any downsides. Its critics would find plenty.

Woolly definition

When you ask Galea what his problems are with precision public health, he laughs as if to say, “Where to start?”

His biggest complaint is that no one has defined exactly what precision public health is. One common description is ‘the right intervention to the right population at the right time’. But critics say this isn’t any different from what John Snow did more than 150 years ago. Dubbing it ‘precision’ is at best extraneous and at worst deceptive, Galea says. “This is what public health does anyway.”

What makes precision public health problematic, he says, is that it focuses on new technologies instead of the bread-and-butter methods that have made the field so successful.

With decades of public-health funding cuts, Galea and his colleague Merlin Chowkwanyun at the Columbia University Mailman School of Public Health in New York City say they understand why epidemiologists and other professionals need to use snappy terms to secure even basic funding. However, they worry that the seemingly subtle change to ‘precision’ approaches conceals a broader shift away from public health’s historical ideals of improving the well-being of even the most marginalized populations.

The core of public health, says David Taylor-Robinson, a health-equities scientist at the University of Liverpool, UK, is to improve the health of populations. The ‘precision’ in precision public health, notes Taylor-Robinson, refers to individuals, not populations. Improving the health of individuals is clinical medicine, not public health. In that sense, he says, “precision public health is an oxymoron”.

The danger of this contradiction is that it becomes easier to lose sight of the lower-tech, more challenging strategies that researchers know would improve people's health, such as universal health care, welfare support and reducing income inequalities, according to Mira Vegter, a social scientist at Wageningen University in the Netherlands.

"We can be blinded by the data opportunities at the expense of some of the more dynamic social questions," she says.

But in an emergency, there might not be time to rise to the challenge of addressing complex social questions. As the COVID-19 pandemic hit, some data-heavy techniques got the chance to prove their worth.

Hi-tech health

Even before the pandemic, many public-health departments, including New York City's, had been honing real-time data analytics to trace infectious-disease outbreaks^{4,5}.

As the coronavirus spread through New York, Greene's goal was to focus on areas where transmission was growing faster than the citywide average. Over time, the work allowed her to begin making short-term predictions (called nowcasts) about future cases⁶.

"It's hard to shift resources around in real time," Greene says. "If you can alert people earlier, they have more of a chance to protect themselves, and the more likely you are to prevent another infection."

As well as tracking cases of COVID-19, epidemiologists around the world also began to track how the SARS-CoV-2 virus itself was spreading and changing, to help inform what measures were likely to be most effective. Over the past 15 years, public-health laboratories around the world have been sequencing the genomes of infectious pathogens — which change slightly as they are passed between people — as a way to connect the dots on who was infecting whom.

It's also how scientists noticed variants of SARS-CoV-2 emerging and spreading. Several large databases of its genome sequences allow virologists

to watch in near real time as coronavirus variants popped up and began to circulate. “It’s like opening a window into a whole new world,” says Emma Thomson, a virologist at the University of Glasgow, UK. “You see so much more detail,” she says.

Pathogen sequencing in public health has been a clear success, even to those wary of precision methods. “To not use these technologies strikes me as unambitious,” says Chowkwanyun. “I think there’s a way to marry this whizz-bang stuff with old-school public health.”

Galea agrees. But he also points out that the first wave of the pandemic overwhelmed New York City before it could scramble adequate resources to mount a defence. Galea says that investments in basic public health, such as improving housing and requiring paid sick leave for workers, would have benefits during a pandemic — and ripple effects beyond.

Besides, says Thomson, pathogen sequencing and other techniques rely on the ability of public-health systems to step in and do something to stem the outbreak. “It’s great and it’s important but it has its limitations, especially when public-health departments are overwhelmed,” she says.

It’s easy to forget that access to basic health data, such as birth and death certificates and communicable-disease reporting, is the fuel that drives the public-health machine, says Angeline Ferdinand, a public-health researcher at the University of Melbourne, Australia. Even basic data are lacking in some places, she points out — let alone the ability to perform whole-genome sequencing and the informaticians to put it to use. This could exacerbate inequalities both within and between countries.

Targeted treatment

For Allen and Khoury, however, there are many more ways in which new technologies and genomic information could pinpoint those most in need of help.

Take cancer, for example. Allen is piloting a project that analyses existing scans with machine learning to help identify individuals at high genetic risk of developing cancer, even if they don’t have a family history of the disease.

These people could benefit from more-intensive cancer screenings such as mammograms and colonoscopies. She recently completed a study that showed information collected by chatbots on medical websites could accurately predict whether a person met criteria for genetic testing for several cancers.⁷

Whether or not researchers see a conflict between public health's fundamental mission and its new 'precision' arm, both sides agree on one thing: public health will continue to be of utmost importance in the coming decades.

"The goal is to improve the health of the whole population, using all the tools that we have, whether we call them precision public health or not," Khoury says.

Nature **601**, 18-20 (2022)

doi: <https://doi.org/10.1038/d41586-021-03819-2>

References

1. 1.

Alfonso, Y. N., Leider, J. P., Resnick, B., McCullough, J. M. & Bishai, D. *Health Aff.* **40**, 664–671 (2021).

2. 2.

Institute of Medicine. in *The Future of Public Health* Ch.3 (National Academies Press, 1988).

3. 3.

Myers, K. D. *et al. Lancet Digit. Health* **1**, E393–E402 (2019).

4. 4.

Latash, J. *et al. Morb. Mortal. Wkly Rep.* **69**, 815–819 (2020).

5. 5.

Greene, S. K., Peterson, E. R., Kapell, D., Fine, A. D. & Kulldorff, M.
Emerg. Infect. Dis. **22**, 1808–1812 (2016).

6. 6.

Greene, S. K. *et al.* *JMIR Public Health Surveill.* **7**, e25538 (2021).

7. 7.

Ritchie, J. B. *et al.* *Hered. Cancer Clin. Pract.* **19**, 31 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03819-2>

| [Section menu](#) | [Main menu](#) |

Opinion

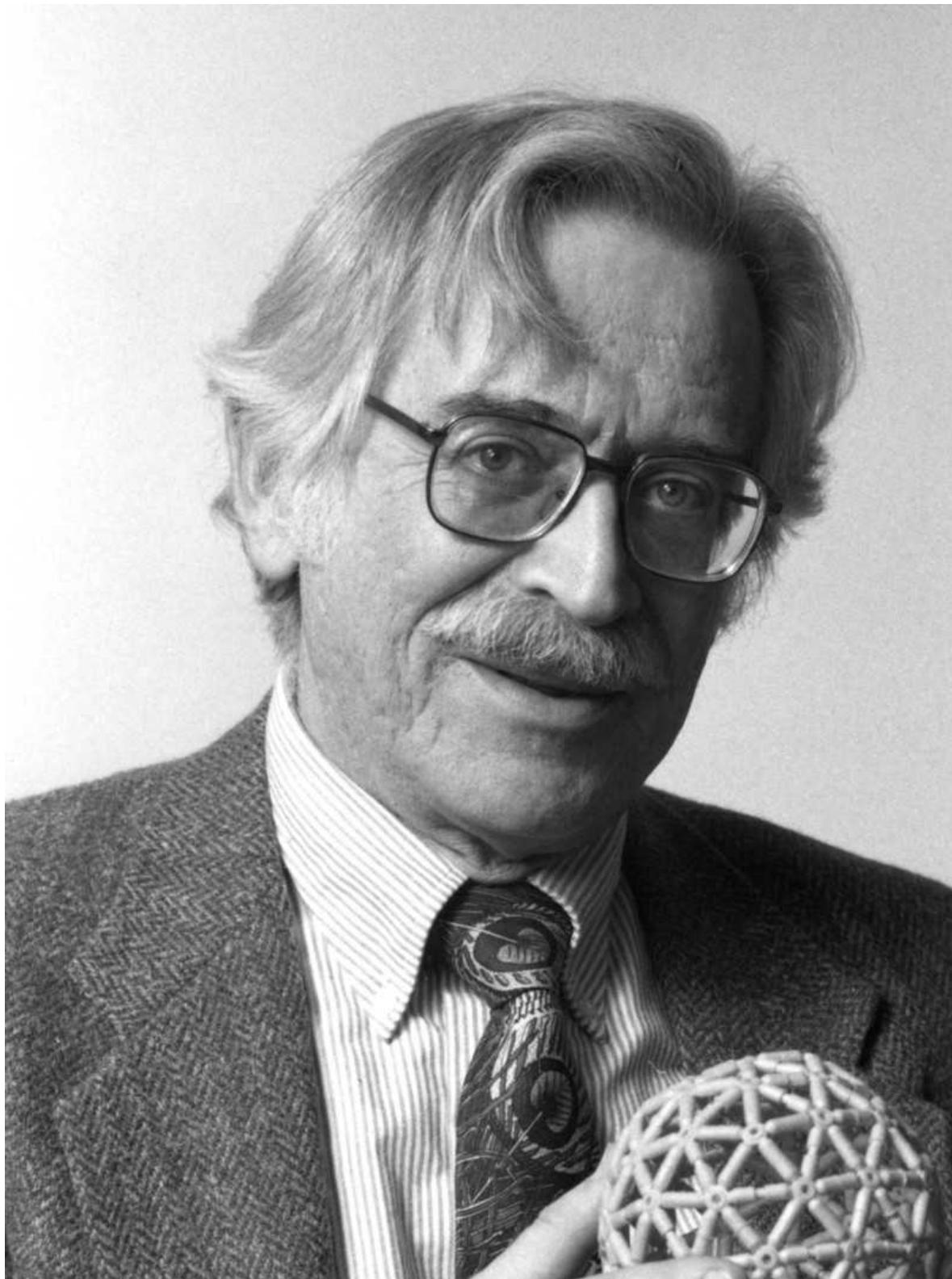
- **[Donald Caspar \(1927–2021\)](#)** [22 December 2021]
Obituary • Crystallographer who defined rules of virus structure.
- **[Two years of COVID-19 in Africa: lessons for the world](#)**
[03 January 2022]
Comment • Africa urgently needs to guarantee its own health security.
- **[Birth equity: US Supreme Court must support Black people](#)** [04 January 2022]
Correspondence •
- **[Push for ethical practices in geoscience fieldwork](#)** [04 January 2022]
Correspondence •
- **[COVID-19: boost mental-health resilience](#)** [04 January 2022]
Correspondence •
- **[Engaging with the science of fetal origins](#)** [04 January 2022]
Correspondence •

- OBITUARY
- 22 December 2021

Donald Caspar (1927–2021)

Crystallographer who defined rules of virus structure.

- [Lee Makowski](#) 0



Credit: AIP Emilio Segrè Visual Archives, Physics Today Collection

Donald Caspar defined the rules that govern the self-assembly of simple viruses. This laid the foundations for a new way of thinking about the molecular systems that regulate and drive all living cells. These rules made it straightforward to characterize other viruses, and then to design strategies to combat them. The same rules are also essential in designing viral vectors to deliver gene therapy.

Viruses typically consist of a strand of genetic material — DNA or RNA — enclosed in a coat of protein molecules. Using the laws of thermodynamics and the constraints of symmetry, Caspar catalogued the ways in which proteins can come together to form the icosahedral shells of spherical viruses (these include rhinovirus and poliovirus) and the helical lattices of rod-shaped viruses (such as Ebola).

In the mid-twentieth century, few protein structures had been solved at atomic resolution. Caspar used limited data from X-ray crystallography, electron microscopy and fibre diffraction to tease out features of prototypical viruses. With an artist's eye, unbounded curiosity and a passion for 'connecting the dots', he created exquisite sketches of virus structures and membranes, which still grace the pages of many texts.

Caspar was born in 1927 in Ithaca, New York, when his father was a graduate student in chemistry at Cornell University. When he was ten years old, a family friend, the crystallographer Isidor Fankuchen, told him of the recent discovery that particles of the rod-shaped tobacco mosaic virus (TMV) had a semi-crystalline structure, suggesting that the virus comprised many identical units. This set him on a lifelong journey to understand its structure and behaviour. At the Polytechnic Institute of Brooklyn in New York City, he was the youngest student on a two-week course on X-ray crystallography run by Fankuchen. After graduating in physics at Cornell, Caspar did his PhD on the structure of TMV at Yale University in New Haven, Connecticut. He then held postdoc positions at the California Institute of Technology in Pasadena, and at the MRC Laboratory of Molecular Biology in Cambridge, UK.

At Cambridge in 1956, he set out to prove the theory advanced by James Watson and Francis Crick that spherical virus particles had cubic symmetry (in other words, that the protein coat was made of identical subunits

arranged in equivalent ways). He demonstrated that tomato bushy stunt virus had icosahedral symmetry. It soon became clear that other viruses were also icosahedral. This, the most complex of the cubic symmetries, requires 60 identical subunits. But biochemical data indicated that many virus particles had a lot more.

Caspar resolved this paradox in the late 1950s, working with Aaron Klug at Birkbeck College in London, and taking a clue from architect Buckminster Fuller's geodesic domes. He introduced quasi-equivalence — the idea that equivalent protein subunits could vary slightly in conformation to tightly seal the viral shells. This opened up a universe of possibilities for the controlled assembly and disassembly of everything inside a living cell. Twenty years later, Caspar's own lab demonstrated that viral proteins can switch between non-equivalent conformations where necessary.

Caspar's PhD work on TMV was the basis for a paper published back-to-back with one by Rosalind Franklin in 1956 ([D. L. D. Caspar *Nature* 177, 928 \(1956\)](#); [R. E. Franklin *Nature* 177, 928–930; 1956](#)). These pinpointed the viral RNA deep within the helical protein rods and marked the start of a friendship that lasted until Franklin's death two years later. In 2008, their relationship was dramatized in Anna Ziegler's play *Photograph 51*, with romantic overtones that Don insisted were fictional.

In 1958, Caspar established what would become one of the most influential centres for the study of macromolecular structure: the Laboratory for Structural Biology at the Children's Cancer Research Foundation at Boston Children's Hospital. The lab was moved to Brandeis University in Waltham, Massachusetts, in 1972. In 1994, Caspar moved to the Institute for Molecular Biophysics at Florida State University in Tallahassee, where he spent the remainder of his career.

Writing was a struggle for Don. Editing a manuscript often devolved into an infinite loop. He wanted to say everything first — every statement was somehow a prerequisite for every other. He was much better at talking. In fact, to those of us who collaborated with Don, it was never obvious when he actually got any work done. He never stopped talking. He could eat his entire lunch without missing a beat. Group members would set a lab timer

before going to ask him a question. When the timer went off, they could plead the need to check an experiment.

Don's many passions included gardening and Grecian urns. In a walk around his garden in Cataumet, on Cape Cod in Massachusetts, he would relate the family, genus and common name of every plant and tree and expound on why he chose a particular variety and colour for each location. Go with him to a museum, and it was never time to leave without a visit to the urns. He was married for 50 years to Gwladys Caspar, an early biosafety officer at Harvard University in Cambridge, Massachusetts. Her quick guide to biosafety levels 1–4 (don't eat it, don't touch it, don't breathe it and don't do it here) is still taught.

Don was brilliant, kind and always enthusiastic about new ideas — his own and others'. His support for young scientists was extraordinary, as was his advocacy for women in science. Don's deep insight and broad perspective will be sorely missed.

Nature **601**, 21 (2022)

doi: <https://doi.org/10.1038/d41586-021-03820-9>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03820-9>

| [Section menu](#) | [Main menu](#) |

- COMMENT
- 03 January 2022

Two years of COVID-19 in Africa: lessons for the world

Africa urgently needs to guarantee its own health security.

- [Christian T. Happi](#) ⁰ &
- [John N. Nkengasong](#) ¹



Researchers receive diagnostics training at the African Centre of Excellence for Genomics of Infectious Diseases (ACEGID) in Ede, Nigeria. Credit: ACEGID

In the early months of the COVID-19 pandemic, Africa's rapid and coordinated response, informed by emerging data, was remarkable. Now, in 2022, as vast vaccination campaigns have enabled the global north to gain some control over the pandemic, Africa lags behind.

Since the 1960s, when many African countries gained independence, the continent has largely depended on the outside world for its health-security commodities: diagnostics, therapeutics and vaccines, as well as personal protective equipment and other medical supplies. The COVID-19 pandemic has exposed how easily international cooperation and multilateral agreements can dissolve, especially in the face of a global crisis — and just how vulnerable this dependence leaves Africa. In 2017 and 2018, more than 120 disease outbreaks were reported on the continent¹.

In our view, leaders of the 55 African Union member states face a stark choice.

In principle, Africa could build on the astonishing gains it has made in surveillance and public-health responsiveness to outbreaks in recent years. It could sufficiently invest in commodities to ensure its health security, and position itself as a world leader in fighting infectious diseases.

The alternative? There really isn't one. If the continent does not work towards guaranteeing self-sufficiency, it will fail to address the infectious-disease threats of the twenty-first century. It will thereby be unable to achieve the development goals encapsulated in the African Union's Agenda 2063: The Africa We Want. This is a blueprint for transforming the continent into a global powerhouse, laid out in 2013.

What Africa got right

In 2017, following the 2014–16 Ebola outbreak in West Africa, the African Union launched the Africa Centres for Disease Control and Prevention (Africa CDC) to help prepare the continent for epidemics and pandemics. Millions of dollars of investment by members of the African Union and organizations such as the Africa Development Bank, the World Bank, partner nations and foundations have bolstered the capability of the Africa

CDC and its Regional Collaborating Centres to handle disease outbreaks. The regional centres are networks of public-health institutions hosted in Egypt, Kenya, Zambia, Gabon and Nigeria for the northern, eastern, southern, central and western regions, respectively.

Just days after the first case of COVID-19 was reported in Egypt in mid-February 2020, the African Union and the Africa CDC convened an emergency meeting of all ministers of health in Addis Ababa, Ethiopia. Attendees agreed to a Joint Continental Strategy, whereby African Union member states would cooperate, collaborate, coordinate and communicate their efforts. The Africa CDC had already established the Africa Task Force for Coronavirus (AFTCOR) in early February 2020 to help achieve this.

In early March 2020, as cases of COVID-19 spread across the continent, countries took immediate and drastic measures, including implementing lockdowns and other restrictions designed to reduce transmission. This continent-wide collaboration was sustained. Since March 2020, AFTCOR has conducted bi-weekly meetings. The Bureau of the Assembly of the African Union Heads of State and Government (a body that coordinates the affairs of the African Union between the union's yearly summits) has also convened nearly every month to review Africa's response to the pandemic, with the Africa CDC providing technical guidance.

So far, 12 initiatives have been launched because of this coordination. These include the Partnership to Accelerate COVID-19 Testing in Africa, which has helped to secure diagnostics; the Africa Medical Supplies Platform, which has enabled the purchasing of crucial medical supplies; and the African Vaccine Acquisition Task Team (AVATT), a centralized purchasing agent established in November 2020 to make it easier for African Union member states to secure vaccines. So far, AVATT has enabled the continent to obtain 400 million doses of vaccines.

Meanwhile, Africa's centres for genomics have achieved impressive feats in surveillance. Coordination has come from the Africa CDC and national public-health institutions, such as the Nigeria Centre for Disease Control in Abuja. In March 2020, the first African SARS-CoV-2 genome was sequenced in Nigeria, only 48 hours after a sample (from a patient who had travelled from Italy) arrived at the African Centre of Excellence for

Genomics of Infectious Diseases (ACEGID) in Ede, with results released after 72 hours (see go.nature.com/3ywdobt). In October that year, the Beta variant² was discovered within days of a sample reaching the Network for Genomic Surveillance South Africa consortium, a collaboration of laboratories, researchers and academic institutions across the country. In November 2021, the heavily mutated Omicron variant was spotted in genome-sequencing data from Botswana and identified by researchers in South Africa.

Indeed, the best-resourced genomic centres have supported many other countries besides their own. So far, ACEGID (which C.T.H. directs), for instance, has sequenced samples from around 30 African countries. Researchers working at the centre have also trained more than 1,300 geneticists and public-health workers and officials from other countries in diagnostics and genomics for infectious diseases. After three weeks, trainees return to their own countries and apply what they have learnt, even though their equipment might be more basic than ACEGID's.

All of these sequencing efforts continue to guide local, regional, national and international public-health responses. Between April and July 2020, for instance, sequencing information obtained in Nigeria revealed that people were obeying lockdown restrictions during the day, but did so less at night. Most recently, alerting the world to the Omicron variant has spurred intensified surveillance in Africa and an explosion of new immunology and sequencing research, as well as changes to public policy in the global north.

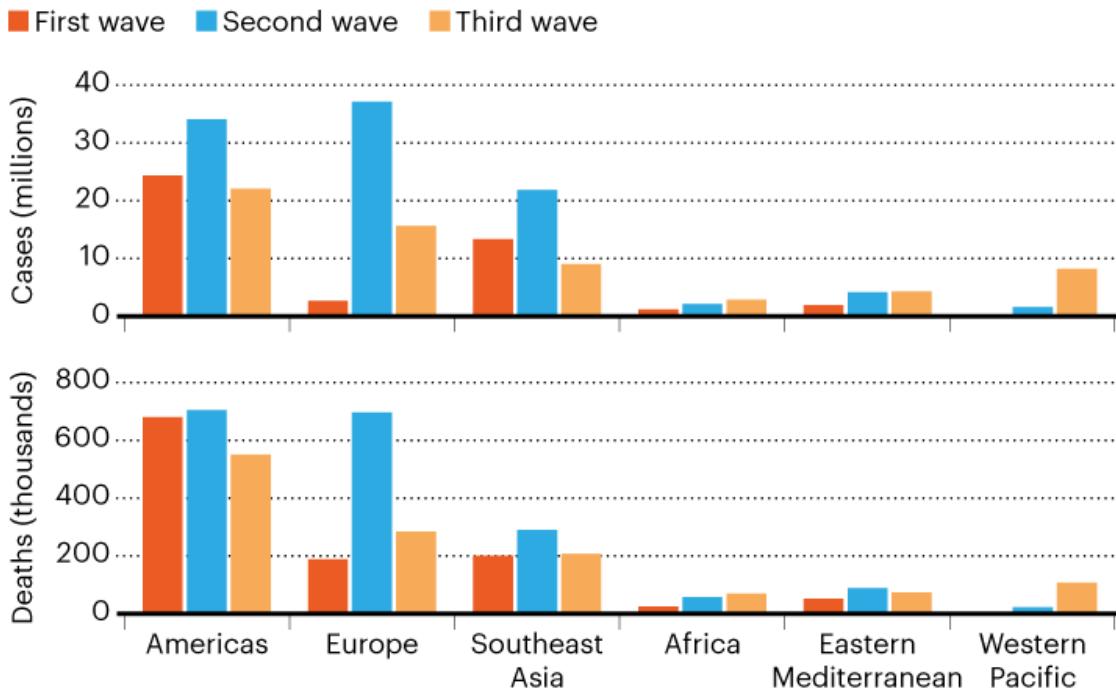
Such feats in coordination and collaboration — which have enabled many African nations to make the best use of available resources — stand in contrast to the inflexibility seen in other parts of the world.

Indeed, Africa's achievements, especially in the early phase of the pandemic, are at least part of the reason the continent was able to mitigate the first and, to some degree, the second waves of the COVID-19 pandemic (see 'The power of experience'). As of 13 December 2021, around 8.9 million cases of COVID-19 and nearly 225,000 deaths had been reported across Africa (<https://africacdc.org/covid-19>). This is in contrast to early models from public-health experts and epidemiologists, which had predicted

that up to 70 million Africans would be infected with SARS-CoV-2 by June 2020, with more than 3 million people dying³.

THE POWER OF EXPERIENCE

It is difficult to compare COVID-19 statistics across regions* with different demographics and varying systems for data collection. But Africa's previous experience of outbreaks at least partly explains why the continent has so far had fewer COVID-19 cases and deaths than many predicted.



©nature

*Regions as defined by the World Health Organization.

Source: WHO

What's happened since?

In the early months of the pandemic, more than 70 countries imposed restrictions on the export of medical materials — including the raw materials needed to make diagnostics. It wasn't until the African Union established the Africa Medical Supplies Platform in June 2020 that the continent was able to start procuring medical supplies at a pace that better met demand.

Now, leaders of African countries are desperately trying to access COVID-19 vaccines. Around 47% of people have been fully vaccinated globally, and many countries in the world are giving a further (booster) dose to their citizens. Yet Africa is still struggling — it has achieved full coverage in only about 7% of eligible people (currently, those who are 18 or older).

This is despite the promises of donors to make two billion vaccine doses available to people living in low- and middle-income countries by the end of 2021 through COVAX — an initiative launched in April 2020 by the World Health Organization, Gavi, the Vaccine Alliance and CEPI, the Coalition for Epidemic Preparedness Innovations.

The inequity of the COVID-19 pandemic is reminiscent of the moral catastrophe that occurred during the HIV/AIDS pandemic in the mid-1990s, as one of us (J.N.N.) warned in these pages more than a year ago⁴. Before the establishment of programmes such as the US President's Emergency Plan for AIDS Relief (PEPFAR) and the Global Fund to Fight AIDS, Tuberculosis and Malaria, more than ten million Africans died needlessly because they couldn't obtain the antiretroviral drugs that were readily available in wealthy nations.

What has happened during the COVID-19 pandemic should be a stark reminder to African leaders of the fragility of international cooperation and multilateralism⁴. But the lessons it holds for Africa go deeper even than this.

The downsides of aid

Historically, efforts to assist Africa have tended to be siloed. They take a top-down approach, with decision-making coming from a central body outside the continent, not from African institutions and experts. Efforts have generally focused on short-term crisis management, not on the kinds of sustainable systems, such as manufacturing capability for diagnostics, that could help Africa to take charge of its health security.

Take the 2014–16 Ebola epidemic. The outbreak wasn't formally classified as a global health emergency until nearly 2,000 cases and nearly 1,000 deaths had already been reported in West Africa. Many more lives might

have been saved had there been institutions in Africa, working directly with first responders on the ground, that were empowered to alert the world. When international aid did arrive from numerous countries, the many organizations involved — with their different messaging and ways of working — heightened people's confusion about the source of the virus and how best to combat it.

Ultimately, Africa's dependence on the outside world sustains a lack of confidence in Africa — both within and outside the continent. Certainly, Africa's successes in the COVID-19 pandemic have not been reported in the mainstream media in a way that could enable the world to learn from them. Instead, the conversation has focused mainly on whether problems with data collection, demographics, climate or other factors could explain why far fewer cases and deaths have been reported than epidemiologists and others predicted^{[3,5](#)}.



People in Nairobi, Kenya, receive a dose of the Oxford–AstraZeneca COVID-19 vaccine. Credit: Brian Inganga/AP/Shutterstock

Five steps

Multilateralism will always be crucial to preventing and responding to epidemics and pandemics. But Africa will be able to benefit from the advances it has made in disease surveillance only if its approach to public health is reconfigured towards self-reliance. This could be achieved in five steps.

Invest in health and disease. African leaders must honour their commitments to allocate at least 15% of their annual budgets to the health sector, as they agreed to do in 2001 at a meeting on HIV and other diseases in Abuja. Today, spending on health across the continent ranges from 0.1% to 2% of a nation's gross domestic product.

Renewing this commitment is especially urgent as the goals articulated in the Agenda 2063 begin to be realized. For instance, the African Continental Free Trade Area agreement, which came into force in January this year, creates a trade area including 1.3 billion people across 55 countries. Likewise, the African Union's Protocol on Free Movement of Persons in Africa will enable people to move around the continent without a visa. Both of these steps are crucial to economic development. They also come with major public-health implications.

Build regional control and capability. This means strengthening national public-health institutions, such as the Nigeria Centre for Disease Control and the Zambia National Public Health Institute in Lusaka. It also means strengthening and empowering the Africa CDC and its five Regional Collaborating Centres. These are the 'Africa-owned institutions' that are crucial to prevent, detect and respond to public-health threats. Ultimately, both national public-health institutions and regional centres must be the command centres for the control of disease outbreaks.

Accelerate translational research and development. Currently, Africa imports 70–90% of its drugs; there is almost no biotechnology sector. For comparison, China and India, with comparable populations, import 5% and 20%, respectively. There is also a severe shortage of human resources. There are currently only around 1,900 epidemiologists in Africa — far fewer than

the 6,000 needed, as stipulated by the Global Health Security Agenda. (The Global Health Security Agenda is an international effort to accelerate the implementation of health regulations, particularly in developing countries, to improve capacity in preventing and responding to infectious-disease threats.)

Governments, philanthropists and the private sector, such as the African Export–Import Bank and Africa Development Bank, must provide sustainable funding for research and development, with a focus on diagnostics, therapeutics and vaccines for infectious and non-communicable diseases. Such investment will be crucial to lure back the thousands of African scientists, clinicians, nurses and other skilled medical health workers who have gone overseas for training or employment. Accelerating translational research and development will also require continued investment in the existing genomics hubs of academic excellence in Africa.

Invest in early-warning systems. Ultimately, Africa — and the world — needs effective surveillance to detect and characterize deadly pathogens before they spread across the globe.

Some promising initiatives are being developed in Africa, such as SENTINEL (see ‘A sentinel for disease’), which one of us (C.T.H.) is leading. To be effective in the long term, such schemes must be integrated into Africa’s public-health institutions. Too many projects on health and disease in Africa are pursued in silos and funded only for as long as the principal investigator spearheading the project can persuade investors.

The Regional Integrated Surveillance and Laboratory Network (RISLNET) was launched by the Africa CDC in 2017 to enable the rapid detection and prevention of emerging public-health threats. Surveillance networks consisting of genomics laboratories and National Public Health Laboratories have already been established in Central Africa, and efforts are ongoing to expand these in southern, eastern and western Africa. In principle, systems such as SENTINEL could be incorporated into RISLNET to ensure that they are in line with what the African Union members envision for the continent, and that they have institutionalized and politicized backing.

Such systems must integrate the surveillance of disease in people, animals and the environment. Since February 2020, those involved in the One Health

programme at the Africa CDC have been trying to foster this holistic approach — mainly by bringing together people working in these different areas.

Build centralized governance. An African Pandemic Preparedness and Response Authority, as proposed by the African Union in October 2021, could empower the Africa CDC to coordinate pandemic responses across borders. This agency could be modelled on the European Health Emergency preparedness and Response Authority (HERA). A signed treaty would mean leaders have to cooperate, share data and so on. Likewise, the continent could capitalize on practices and tools born out of the COVID-19 crisis. The African Union COVID-19 Response Fund, established in March 2020, has enabled countries to pool funds to buy medical commodities, such as personal protective equipment. In principle, this could be upgraded to an African Disease Threat Fund. Similarly, AVATT could be used as a platform for acquiring other vaccines for the continent.

The way ahead

Unlike in previous disease outbreaks, during the COVID-19 pandemic, Africa has been a key player in the acquisition of scientific knowledge that has guided the global response.

For two years, hundreds of geneticists across the continent have worked seven days a week, often through the night, to sequence strains of SARS-CoV-2. And companies worldwide have used these data (most of which are available in public repositories such as GISAID) to develop COVID-19 vaccines and therapeutics. For those contributing so much to the global effort to curb this pandemic, it is galling to watch Africa continue to struggle in the acquisition and roll-out of COVID-19 vaccines.

The new way of doing public health that we set out here offers a different future. Embracing it is an imperative for Africa's security and economic survival. It will also benefit the world — as so powerfully demonstrated in November 2021 by the discovery of the Omicron variant.

A sentinel for disease

An early-warning system being developed in Africa could transform people's ability to prevent and respond to disease outbreaks throughout the world.

One of us (C.T.H.) has been working on the SENTINEL initiative⁶, in collaboration with geneticist Pardis Sabeti at the Broad Institute in Cambridge, Massachusetts, since January 2020. The project has three goals.

Detection. The idea is to use a series of three new diagnostic techniques to detect diseases. The first uses CRISPR–Cas9 gene editing to detect known viruses in a simple, inexpensive test for the highest-priority pathogens in a region. This can be used wherever a patient is receiving care — even in their own home. If this test fails to identify a pathogen, a blood sample taken at a primary-care facility and sent to a local hospital would be subject to a more expensive test. This detects hundreds of viruses and viral strains simultaneously. If this also fails to detect a known pathogen, a sample would then be sent to a regional sequencing centre for metagenomic sequencing. There, every virus — known and unknown — can be identified and tracked by sequencing all the DNA in the sample.

So far, researchers at ACEGID and the Broad have validated three CRISPR-based diagnostics — for Ebola, Lassa fever virus and SARS-CoV-2 (ref. 7). These technologies have also been incorporated into a single test that can detect up to 169 viruses simultaneously in a single sample^{7,8}.

Connection. Alongside better tools for the detection and surveillance of disease, researchers at the Broad and at ACEGID are building tools to enable the integration of all sorts of data sources — from diagnostic and clinical data obtained by health-care workers, to genomic surveillance data gathered across a continent. The aim is to make all these data streams available (and connected in real time) on a continent- or region-specific cloud-based dashboard. Given the limited resources in Africa, and in many of the other regions where SENTINEL could operate, many of these tools are being designed specifically for scientists, laboratory staff and public-health workers who might have limited computational experience or on-site hardware resources.

Empowerment. Diagnostic and informatics technologies are futile unless they are used to guide clinical care and public-health interventions. To be effective, SENTINEL must empower every stakeholder, whether governments and public-health officials, front-line health-care workers, scientists or communities and individual patients.

Training people in genomic sequencing, diagnostics, bioinformatics and advanced genomic surveillance (metagenomics sequencing) will be crucial for SENTINEL to be effective in the long term. So far, researchers at the Broad and at ACEGID have trained more than 1,300 local health-care workers and scientists to use SENTINEL's diagnostic and surveillance tools. Ultimately, tens of thousands will need to be trained.

Nature **601**, 22–25 (2022)

doi: <https://doi.org/10.1038/d41586-021-03821-8>

References

1. 1.

World Health Organization Regional Office for Africa. *Emergency Operations Annual Report. Saving Lives and Reducing Suffering: WHO's Work in Emergency Response Operations in the WHO African Region in 2018* (WHO, 2020).

2. 2.

Wilkinson, E. *et al. Science* **374**, 423–431 (2021).

3. 3.

Walker, P. G. *et al. Report 12: The Global Impact of COVID-19 and Strategies for Mitigation and Suppression* (Imperial College London, 2020).

4. 4.

Nkengasong, J. *Nature* **580**, 565 (2020).

5. 5.

Maeda, J. M. & Nkengasong, J. N. *Science* **371**, 27–28 (2021).

6. 6.

Botti-Lodovico, Y. *et al. Viruses* **13**, 1605 (2021).

7. 7.

Ackerman, C. M. *et al. Nature* **582**, 277–282 (2020).

8. 8.

Myhrvold, C. *et al. Science* **360**, 444–448 (2018).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03821-8>

| [Section menu](#) | [Main menu](#) |

- CORRESPONDENCE
- 04 January 2022

Birth equity: US Supreme Court must support Black people

- [Joia A. Crear-Perry](#) ✉

As president of the National Birth Equity Collaborative, I agree that undermining the 1973 *Roe v. Wade* abortion-viability standard would detrimentally affect people with the capacity for pregnancy, particularly those who are Black (see [Nature 599, 187–189; 2021](#)). The collaborative led the submission of an *amicus* brief on safe and legal access to abortion for maternal health and well-being (see [go.nature.com/33jwjsn](#)) as evidence against Mississippi state's *Dobbs v. Jackson Women's Health Organization*.

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Nature **601**, 26 (2022)

doi: <https://doi.org/10.1038/d41586-021-03836-1>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03836-1>

| [Section menu](#) | [Main menu](#) |

- CORRESPONDENCE
- 04 January 2022

Push for ethical practices in geoscience fieldwork

- [Giuseppe Di Capua](#) ORCID: <http://orcid.org/0000-0002-1254-3200>⁰,
- [Martin Bohle](#) ORCID: <http://orcid.org/0000-0002-8794-5810>¹,
- [Dominic Hildebrandt](#) ORCID: <http://orcid.org/0000-0002-9891-0664>²,
- [Eduardo Marone](#) ORCID: <http://orcid.org/0000-0003-4521-8604>³,
- [Silvia Peppoloni](#) ORCID: <http://orcid.org/0000-0002-2667-6506>⁴ &
- [Simon Schneider](#) ORCID: <http://orcid.org/0000-0002-8561-0371>⁵

Destructive geoscientific sampling can have ethical repercussions. It can disrupt archives of Earth's history (see [R. Butler *Nature Geosci.* 8, 817–818; 2015](#)) and the management of landscapes and geological formations that have local cultural significance (see [Nature Geosci. 14, 537; 2021](#)). As members of the International Association for Promoting Geoethics, we suggest ways to ensure that geoscientific sampling and fieldwork are carried out ethically.

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

Subscribe

All prices are NET prices.
VAT will be added later in the checkout.
Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

Rent or Buy

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Nature **601**, 26 (2022)

doi: <https://doi.org/10.1038/d41586-021-03837-0>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03837-0>

- CORRESPONDENCE
- 04 January 2022

COVID-19: boost mental-health resilience

- [Lola Kola](#) 0

The social and economic effects of the COVID-19 pandemic are harming mental health and well-being. News of fast-spreading variants of the virus SARS-CoV-2 underscores the urgency of building population resilience to further damage.

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

Rent or Buy

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Nature **601**, 26 (2022)

doi: <https://doi.org/10.1038/d41586-021-03838-z>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03838-z>

| [Section menu](#) | [Main menu](#) |

- CORRESPONDENCE
- 04 January 2022

Engaging with the science of fetal origins

- [Sarah S. Richardson](#) [ORCID: http://orcid.org/0000-0002-9462-3051](http://orcid.org/0000-0002-9462-3051)

Anna Nowogrodzki states that my book, *The Maternal Imprint*, “doesn’t engage” with programmes such as the Avon Longitudinal Study of Parents and Children ([Nature 599, 199–200; 2021](#)). On the contrary, I introduce the study as exemplary of data-rich science on fetal origins in the opening pages, highlight the work of its researchers in testing causal hypotheses concerning fetal epigenetic programming, and feature my own collaborative scholarship with affiliated researchers in the book’s closing chapter. My collaborations and exchanges with researchers on developmental origins underpinned the book.

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.
VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Nature **601**, 26 (2022)

doi: <https://doi.org/10.1038/d41586-021-03839-y>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03839-y>.

| [Section menu](#) | [Main menu](#) |

Work

- **[Exiting gracefully: how to leave a job behind](#)** [20 December 2021]
Career Feature • Although burning bridges can be tempting, a better approach is to leave goodwill and cheer in your wake.
- **[Terra takes the pain out of ‘omics’ computing in the cloud](#)**
[04 January 2022]
Technology Feature • The web-based tool allows scalable, user-friendly computation across multiple data sets
- **[Close-up with a parasite that can blind](#)** [04 January 2022]
Where I Work • At the Pasteur Institute of Montevideo, María Eugenia Francia tackles a cause of vision problems in babies.

- CAREER FEATURE
- 20 December 2021

Exiting gracefully: how to leave a job behind

Although burning bridges can be tempting, a better approach is to leave goodwill and cheer in your wake.

- [Bianca Nogrady](#) ✉

[Find a new job](#)



Leaving your job on a good note will reflect well on you and could lead to future opportunities.Credit: Getty

The typical movie-style job departures made popular in films such as *Bridget Jones's Diary*, *Jerry Maguire* and *American Beauty* see the soon-to-be-ex-employee exiting spectacularly, burning every bridge on the way out, or being escorted off the premises so fast that their feet barely touch the ground.

But the reality of departing a job can be a much gentler experience. Maybe you've been head-hunted and received an offer that's too tempting to refuse, perhaps your current job has become repetitive and it's time to make a move, or your contract might simply have finished. But it's possible that you've been miserable in your position, you've been made redundant, or you've been sacked.

For medical biologist Leigh Coulter, at the Walter and Eliza Hall Institute of Medical Research in Melbourne, Australia, the decision to leave his position as a laboratory head was one he had contemplated for a while. Over the course of his academic research career, he says, he had gradually come to the view that he wanted the products of his science to be "useful for people — not just knowledge for the sake of knowledge, but knowledge for the sake of improving people's lives". So when an opportunity arose in the business-development section at the institute in February 2020 — which coincided with the end of his research project and a trainee's PhD programme — he took it.

Career changes

The pandemic seems to have changed many people's career trajectories, whether they wanted the change or not. A record number of people have quit their jobs in the United States this year, in what the world's media has dubbed 'The Great Resignation'. A survey of 30,000 people in 31 countries, commissioned by technology firm Microsoft, based in Redmond, Washington, earlier this year, found that more than 40% of respondents were thinking of leaving their employer within the next year. Another survey, of 2,000 people in the United Kingdom and Ireland, found that 38% were looking to change their role or job in the next 12 months.

Whatever the circumstances of a departure from a position, there is generally a right way and a wrong way to leave a job (see ‘How to leave a job gracefully’). The wrong way is taught by Hollywood movies. The right way ensures that you maintain goodwill with your former colleagues and managers — who might very well end up being your colleagues or managers again in the future — and delivers a smooth transition to whoever replaces you without undoing the work that you have done while in the job.

As a military spouse, clinical genomics scientist Adrienne Nugent, who now works remotely in Guilford, Connecticut, for genetic-information company Invitae, based in San Francisco, California, has experienced more than her fair share of job departures and new beginnings, because she has moved around the country with her husband’s postings. The first such move was unexpected, coming one year into what was originally supposed to be a 2.5-year position and just after Nugent had started as a postdoc with the US National Institutes of Health. “As is typical for someone a year into a postdoc, at that point in time I had multiple long-term experiments ongoing in the lab, mice breeding, collaborations forming, et cetera, and the news of our impending move was very difficult,” she says.

She was fortunate to be working with a supportive mentor, who was keen to help Nugent continue her postdoc through a combination of remote and part-time work and a 180-mile, once-weekly commute. The experience — and others she’s had since then — have emphasized to Nugent the importance and value of being open and transparent with employers throughout job transitions.

Leaving in the right way

Once you decide to leave a job, the first thing you should do is check your employment contract or any agreements that you signed when you were hired, says Deniece Maston, human-resources knowledge adviser at the Society for Human Resource Management in Brandywine, Maryland.

“You want to make sure you check such [a] document when you are resigning from a job, just to make sure you’re in compliance with something

you signed,” says Maston. For example, an employment contract might stipulate the amount of notice you must give before leaving a job.

But, particularly in senior roles, there could be clauses to ensure that departing employees maintain confidentiality around their previous role, or noncompete clauses that prohibit employees, for a certain period of time, from working in the same industry or for a potential competitor.



Rosemary Guyatt (right) recommends providing honest, constructive feedback in exit interviews.Credit: Rosemary Guyatt

This can be an issue in the sciences: researchers and investigators can take with them a wealth of knowledge when they leave, says Rosemary Guyatt, general manager of people and culture at the Australian Human Resources Institute in Melbourne.

“That definitely is a consideration when people are moving [into a] similar role, similar industry, [or] similar sector, and are potentially bringing intellectual property from one organization to another,” Guyatt says.

The next step in leaving a job is telling your manager. Although a written resignation letter is standard practice, Maston says it's important to first meet with your manager or supervisor to tell them the news in person. But before doing this, you need to work out the reason you are going to give for leaving — one that you're comfortable with telling people — and you need to stick with that reason. “Don’t give different reasons to different people in the same organization,” Maston says. “Stick to one story of why you’re leaving.” Gossip happens, and if you give people different reasons for your departure, word will inevitably get around. “Don’t be dishonest about your next move; often, your boss and perhaps former colleagues will find out where you have landed soon enough.”

This meeting is also the opportunity to work through the implications of your departure with your manager. As a lab head who was responsible for staff and PhD students, Coulter had some big loose ends to tie up. “There’s the experiments that might be ongoing, which you need to wind down; there’s the staff who work for you,” Coulter says. “Things like your funding that you’ve got going at the time, the projects that you’ve got going on, and wanting to see that all of those things come to a positive end.”

In his case, Coulter’s departure was timed to take advantage of his PhD student completing their research, his funding agreements all coming to an end, and his major projects being in their final stage of completion or publication. Although some of that was serendipity, Coulter did consider the effects of his departure on other staff members as much as was practical and tried to reduce those effects — which everyone should do when leaving a job. “You work with people for years at a time, students in particular; you invest a lot in them, and they invest a lot in their project and they’ve got a long career and a bright future in front of them,” he says. “You want to make sure that that isn’t impacted by your decision.”

This is also the time to set a date for your departure. Even if that date is predetermined by your contract, it might be possible to negotiate an earlier leaving date if that works for your employer.



Leigh Coulter suggests making arrangements for a smooth handover of duties. Credit: Walter and Eliza Hall Institute of Medical Research

Potential challenges

Neuroscientist Mark Stahl's departure from a position as head of a laboratory and clinician at the Pennsylvania State University in Philadelphia took a little longer to work through. It was more than five months from the time he decided to take a position at Neurocrine Biosciences in San Diego, California, to when he actually moved into the new job.

Sorting out what to do with his grants, which he had worked hard to gain during his tenure, was the easy part. "For better or for worse, most of them were near the end of their cycle anyway," Stahl says. Of the remaining ones, Stahl was able to transfer one to a graduate student, who was continuing the research, and the other was shared with his co-investigator.

But a larger challenge was closing off the clinical side of his work, because he had patient appointments scheduled up to six months in advance. “For a specialty clinic, it’s not like you can just flip a switch,” Stahl says.

Another challenge in the sciences can be managing research publications in the pipeline. Leaving a research project midway might mean that your name is less prominent in the author list — or it could even be left off altogether. Nugent’s advice is to be clear about your expectations regarding authorship and have those conversations before leaving your position, no matter how uncomfortable that discussion might feel. “Inevitably, experiments or reviewer comments will dictate a change in course, so while these conversations can be uncomfortable and difficult at the time, laying the groundwork and summarizing the results by writing in a follow-up e-mail will save a lot of headache and heartache in the future,” she says.

Maston says many companies will have a resignation checklist that will guide departing employees through the many steps needed to close off their time as an employee. This can cover everything from working through benefits to handing over passwords, computers and e-mail accounts.

As your time draws to a close, you might be invited — or requested — to participate in an exit interview. Guyatt says not all employers do exit interviews, but they are more likely to be standard practice at larger organizations that want to better understand why employees leave. “It can be a really important, helpful source of data and insight for an employer, whether it’s public or private or not-for-profit,” she says.

Doing a good exit interview is a careful balance between honesty and restraint, particularly if there are questions about sensitive topics, such as interpersonal conflicts. “During such times, an employee should be calm and constructive,” Maston says. “If there is going to be something that’s revealed, stick to the facts while being open and direct with those responses, because you don’t know where that’s going to go.”

But being frank can be important, particularly if there’s something that the employer can do to improve a situation, for example unpleasant or unsafe working conditions, job insecurity or low pay rates. “If the same things are repeated in negative feedback, an employer has to follow up and act upon

some of that information,” Maston says. Exit interviews are also often not anonymous, so what is said in them could make its way back to colleagues or managers.

Guyatt agrees that providing honest feedback during an exit interview is important, particularly if there is a sensitive or serious matter that can’t necessarily be given as your public reason for departure. “People also can make the choice of whether they talk that through with the manager who they’re resigning from and saying, ‘I’m letting you know that this, this and this hasn’t been right, but this is going to be my public reason’,” she says. However, she advises balance between positive and negative feedback, “so it does doesn’t look like a massive dump of emotion at the end: ‘I’ve saved it all up. And here it is: I hate you all’”.

In the case of more serious issues in a workplace, such as bullying or harassment, Guyatt says it is important to raise them before leaving, typically with someone from human resources and preferably someone senior. She advises being factual, not just relying on hearsay, and to consider putting your perspective in writing. “Note if your concerns were raised previously and to whom, and what action was taken,” she says, and, if this is your reason for leaving, tell them that.

Despite having yearned to become a professor since her teens, neuroscience doctoral student Elizabeth Aulino at Kent State University in Ohio soon discovered that academia wasn’t quite what she had expected. The long hours, weekends spent working, the teaching requirements and the isolation quickly took their toll, and in mid-2021 she came to the decision to leave academia.



Elizabeth Aulino made sure to be open with her supervisor about her plans. Credit: A.J. Giorgi

“When I realized that I was not going to stay, I just spent a lot of time thinking about what my values are, and what I actually care about, and which of those things need to be part of a job,” says Aulino. She has also been open and honest with her adviser about her reasons for departing. “The conversation I had with my boss was, ‘Hey, I know I just got here, but I’m not really feeling it’,” she says. She decided to let her adviser know when she was looking for jobs, rather than waiting until she got one and would soon depart. “I was trying to be courteous and make it easier for my adviser to find another postdoc,” she says.

Changing jobs is not an easy transition, particularly if the circumstances in which you’re leaving are less than ideal, such as having been made redundant, says Guyatt. That departure can be accompanied by shock, anger and distress, because the decision to depart has been taken away from you. “In handling that scenario, I think it’s important to avail yourself of some of

the support that might be offered because you've just got to quickly adjust to that change." That support might come in the form of assistance with a career transition — for example, employers might offer career coaching, guidance on updating a CV or LinkedIn profile, or counselling and psychological support — and could even include the option of an earlier departure.

Even if your departure is planned, and you're excited about the new job that awaits you once you leave the old one, Guyatt recommends timing your departure to allow for a week or two off between jobs. "Often, people sign up for a new job straight after the notice period, and then wish that they had a week off," she says. You might have been job-hunting for some time, which can often mean discreetly preparing applications, preparing for and attending interviews, and doing research, all the while doing your own job.

Taking time off between jobs is even more important now, when 'onboarding' processes for a new position can begin well before you actually start work. "With onboarding happening virtually or digitally, you're probably already reviewing policies and procedures, signing documents, reading some background — they're probably going to send you some reading material," Guyatt says. Taking a short break between jobs is the opportunity to rest and prepare for the new challenge. "You've really got to be on your game when you're starting a new job: it's going to require a lot more of you concentration-wise." Finishing an old job and starting a new one during the COVID-19 pandemic poses its own unique challenges. For many, it rules out a farewell party because most employees are working from home and, in some places, there are public-health restrictions preventing a typical restaurant or pub gathering. It could also mean starting a new job virtually, without having the opportunity to meet your co-workers in person, have a welcome morning tea, or get a feel for your employment situation. "It requires really a lot more of the individual, in terms of their effort and concentration," Guyatt says.

And finally, Guyatt recommends resisting any temptation to bad-mouth a past employer once you've left. Criticisms of a job or an employer should be said to the appropriate people before you leave, but then "wrap it up", she says. "It's good to leave on a good note, and then to leave it behind."

How to leave a job gracefully

Nature asked the US National Institutes of Health's human-resources department for advice. Here's what their spokesperson told us.

- Use your departure to build bridges, not burn them.
- Give your supervisor plenty of notice of your departure.
- Make sure your colleagues can get in touch with you after you leave.
- Ask for an exit interview if you have specific concerns to raise.
- Have a conversation about any unpublished papers with your supervisor and try to secure your authorship position.
- Try to put in place collaboration agreements if you plan to collaborate with your lab and colleagues.
- Leave your research notes but ask to take photocopies of them, and if you want to take any research materials, documents or equipment with you, make sure you get approval.
- Make sure any reagents you generated during your time in the job are catalogued and stored so they can be retrieved if needed.
- Make sure you are removed from any protocols regarding human participants or animal subjects in your research.
- Get an update on the status of any of your intellectual property (IP) that is owned by your institution, and if you want to continue to use that IP in your next or future job, get clear guidance on any restrictions or limits on you doing that.
- Do what you can to help your students or trainees find and transition into new positions.

Nature **601**, 151-153 (2022)

doi: <https://doi.org/10.1038/d41586-021-03788-6>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03788-6>

| [Section menu](#) | [Main menu](#) |

- TECHNOLOGY FEATURE
- 04 January 2022
- Correction [05 January 2022](#)

Terra takes the pain out of ‘omics’ computing in the cloud

The web-based tool allows scalable, user-friendly computation across multiple data sets

- [Jeffrey M. Perkel](#)

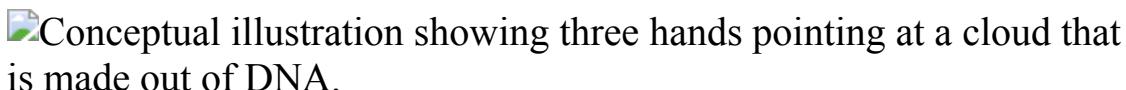


Illustration by the Project Twins.

Biomedical research produces massive amounts of data. But actually sharing them can be a challenge.

“It doesn’t matter how enthusiastic you are about sharing data, actually sharing data is logically a pain,” says Elinor Karlsson, a geneticist at the University of Massachusetts Chan Medical School in Worcester.

Karlsson is a co-lead of the genetics core for the Dog Aging Project, a longitudinal study funded by the US National Institutes of Health (NIH), that aims to understand the genetics of healthy ageing by analysing several thousand genome sequences and health data from 100,000 pets.

The project involves researchers from multiple institutions, all of whom need to be able to access and analyse the data. Conventional data-sharing strategies — such as shared servers, data downloads and even shipping physical disk drives — simply weren’t up to the task, Karlsson says.

Advances in DNA sequencing mean it's much easier to collect genetic data than it is to analyse them. As data sets have ballooned, Karlsson says, "We went from it being inconvenient to copy data to my server, to being cost-prohibitive. Institutions didn't have the resources to keep up."

So the project turned to a system that could: Terra.

Terra began life as a service called FireCloud. It was developed by the Data Sciences Platform (DSP) team at the Broad Institute of MIT and Harvard in Cambridge, Massachusetts, in collaboration with Microsoft and Verily Life Sciences — a subsidiary of Google's parent company, Alphabet. In 2019 it was renamed and folded into the National Human Genome Research Institute's Analysis, Visualization and Informatics Lab-Space (AnVIL). Terra's web-based interface provides user-friendly access to scalable workflows, collaboration and analysis tools. It allows researchers to quickly integrate and computationally analyse vast 'omics' data sets on the Google Cloud Platform. Rather than forcing researchers to find and download data to analyse them locally, Terra lets them work on the data *in situ*, using as much or as little computing power as they need.

"It's really about increasing access and breaking down [data kept in separate] silos so that you can do more interesting science with the data that's being generated," says Geraldine Van der Auwera, communications director for the DSP at the Broad Institute. Terra provides access to data sets such as the Cancer Genome Atlas, the Genome Aggregation Database and the All of Us Research Program, the last of which comprises some 3.7 petabytes alone. Users can also upload their own data; access to 'sensitive' (personally identifiable) information is limited to people with proper authorization.

Terra can be used in 'batch' mode, running scripts written in the Workflow Description Language (a tool for specifying data-processing workflows) to drive anything from one to tens of thousands of virtual computers. Alternatively researchers can explore data interactively using tools such as Jupyter Notebook, RStudio and the graphical Galaxy workflow engine.

Michael Schatz, a computational biologist at Johns Hopkins University in Baltimore, Maryland, was part of a team that used Terra and other AnVIL

components to process data for the Telomere-to-Telomere consortium, which in May 2021 documented the first gap-free assembly of the human genome. That assembly adds some 200 million base pairs of DNA to the human genome, and the Telomere-to-Telomere consortium used it to search for variants in the 3,202 genomes sequenced as part of the 1000 Genomes Project. A computing cluster within the Maryland Advanced Research Computing Center (MARCC) in Baltimore, a multimillion-dollar high-performance facility, would have needed up to a year to perform those analyses, Schatz estimates. But using Terra it took just weeks, including optimization time. “If we had to do it again, we could probably get it all done in about a week,” he says.

For Karlsson, Terra eases data access and collaboration problems. Her PhD student, Kathleen Morrill, has created scripts to automate the uploading and analysis of new data sets, for instance to determine the sex of the sequenced animals. Researchers who wish to explore those data further can simply log in, copy her shared workspace and get to work. “In the physical laboratory, you’d make your data reproducible through your lab notebook,” Morrill says. “By having the analysis up on the cloud in these Terra workspaces it becomes something that’s easier to track and share and reproduce.”

C. Titus Brown, a bioinformatician at the University of California, Davis, and his team are creating tutorials for Terra as part of their work for the NIH Common Fund Data Ecosystem, a data-reuse initiative. Although he finds the platform promising, Brown cautions that cloud computing is still new, and says that it will be some time before it’s the best option for everybody. “There’s a technological adoption curve,” he explains, “and we’re still [in the] very early stages.” Few scientists are comfortable working in the cloud, even with Terra’s relatively user-friendly interface. And researchers and institutions are more familiar with buying computers than renting time on them.

And then there’s the fact that the cost of cloud computing can be opaque. Alexander Bick, a human geneticist at Vanderbilt University Medical Center in Nashville, Tennessee, recalls one student who inadvertently over-requested computing resources on Terra, racking up a US\$500 tab for a job that should have cost just \$5. “It was among the most expensive runs that

we had,” he says. But it was also comparable to wasting a vial of antibodies, he notes. “It’s all relative.”

Schatz advises starting small, for instance by analysing a few samples to assess the per-sample cost, then scaling up. Establish firewalls so intensive workflows don’t burn through your budget. And optimize your workflow, adds Ryan Collins, a computational geneticist and PhD student at Harvard Medical School in Boston. “Every additional gigabyte of RAM that you request is going to translate directly to a cost that will get billed back to you,” he says.

Terra’s website lists workspaces that bundle different workflows together with model data, including pricing details to help users estimate the cost for their own data. Some 56 such featured workspaces have been created, Van der Auwera says, including two from the Human Cell Atlas project.

For Brown, Terra shines when working with data that are either sensitive or extremely large. Privacy rules preclude working with sensitive health data on insecure systems, he notes. And downloading data from the cloud can be expensive, with some data sets costing tens of thousands of dollars to download because of their size. “If instead you can spend \$3,000 on Terra to do your computing, that’s clearly a win.”

Besides, in many cases there is no other practical option, Karlsson notes, even for institutions with high-performance local computing resources. Powerful on-premises clusters, such as MARCC, can still struggle with modern genomics data sets. “Whether Terra or something else, we need something to get to this huge data-scale environment,” she says.

Nature **601**, 154-155 (2022)

doi: <https://doi.org/10.1038/d41586-021-03822-7>

Updates & Corrections

- **Correction 05 January 2022:** An earlier version of this Technology feature neglected to mention that Terra was developed in collaboration

with Microsoft and Verily Life Sciences. It also misnamed the group involved.

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03822-7>

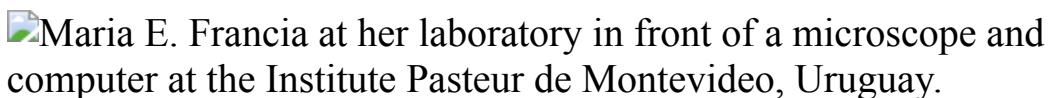
| [Section menu](#) | [Main menu](#) |

- WHERE I WORK
- 04 January 2022

Close-up with a parasite that can blind

At the Pasteur Institute of Montevideo, María Eugenia Francia tackles a cause of vision problems in babies.

- [Chris Woolston](#) ⁰



María Eugenia Francia is a group leader in the Apicomplexan Biology Laboratory at the Pasteur Institute of Montevideo, Uruguay. Credit: Pablo Albarenga for *Nature*

In this picture taken in September, I'm using a confocal microscope at the Pasteur Institute of Montevideo, one of the Pasteur Network's three South American centres. My team focuses on understanding and combating infectious agents, specifically single-celled protozoan parasites belonging to a diverse group called Apicomplexa.

Here, I'm looking at a colour-coded -3D image of the cytoskeleton of *Toxoplasma gondii*, a parasite that infects people all over the world, with infection rates above 60% in some hot, humid places. The organism can cause vision problems in babies born to infected mothers. It's also a major cause of miscarriages in sheep. Uruguay has more sheep than people, so that's a huge concern.

As a younger scientist, I spent my days at a microscope. Now I'm mostly a manager, running a group of ten people, but I appreciate any chance I get to

see these organisms up close. As damaging as they are, these parasites are also beautiful.

This microscope, maintained by the centre's bioimaging unit, isn't powerful enough to zoom in on the parasite's smallest structures. We get round that by enlarging the parasite itself with an expansion gel.

Using samples from local hospitals and sheep ranches, my laboratory studies *T. gondii* and its close relatives from many angles: genetics, virulence and basic biology. Its method for cell division is very different from ours. Anything that separates these organisms from mammals can be a potential target for drug therapy.

I've made several appearances on Uruguayan television talking about my research and the work of the centre. I'm the only female group leader here, so I also try to encourage and support other women in science. I can talk openly about our lack of representation in science without worrying about my position. I need to capitalize on that privilege for the sake of other women.

Nature **601**, 158 (2022)

doi: <https://doi.org/10.1038/d41586-021-03823-6>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03823-6>

Research

- **[Earth's eccentric orbit paced the evolution of marine phytoplankton](#)** [01 December 2021]
News & Views • Analysis of plankton fossils has revealed pulses of size diversity that are inextricably linked to the degree of circularity of Earth's orbits. Could this orbital variability provide a beat that dictates the rhythm of evolution?
- **[Aggregates of TDP-43 protein spiral into view](#)** [08 December 2021]
News & Views • In some neurodegenerative diseases, a protein called TDP-43 forms aggregates in the brain, resulting in neuronal cell death. The structure of these aggregates and their properties have been unveiled.
- **[From the archive](#)** [04 January 2022]
News & Views • Efforts to prove Fermat's last theorem, and a bewildered translator tackles a book about birds.
- **[A heritable, non-genetic road to cancer evolution](#)** [08 December 2021]
News & Views • Treatment for leukaemia can fail for reasons that are not fully clear. Tracking the progress of individual cellular lineages for this type of cancer offers a way to investigate this phenomenon.
- **[Mirror symmetry validated for proton and its antimatter twin](#)** [05 January 2022]
News & Views • A comparison of the charge-to-mass ratio for the proton and the antiproton has been performed with record-breaking precision — the best such test yet for a mirror-like symmetry that relates matter and antimatter.
- **[Iron pnictides and chalcogenides: a new paradigm for superconductivity](#)** [05 January 2022]
Review Article • The progress and the outstanding issues in understanding the correlated phases in the unconventional iron-based superconductors is reviewed.
- **[A stellar stream remnant of a globular cluster below the metallicity floor](#)** [05 January 2022]
Article • Observations of a stellar stream below the metallicity floor for a disrupted globular cluster are described.

- [An early transition to magnetic supercriticality in star formation](#) [05 January 2022]
Article • An analysis of Zeeman measurements reveals that the reduction of magnetic flux relative to mass, which is necessary for star formation, seems to have occurred earlier than previously thought.
- [A 16-parts-per-trillion measurement of the antiproton-to-proton charge–mass ratio](#) [05 January 2022]
Article • Multiple high-precision measurement campaigns at CERN of the antiproton-to-proton charge-to-mass ratio—to a precision of 16 parts per trillion—in a cryogenic multi-Penning trap offer no evidence of charge–parity–time violation, and set stringent limits on the clock-weak-equivalence principle.
- [Crystallization of bosonic quantum Hall states in a rotating quantum gas](#) [05 January 2022]
Article • Spontaneous crystallization of atoms occurs in a rotating ultracold Bose–Einstein condensate occupying the lowest Landau level, behaviour that is related to a quantum hydrodynamic instability driven by shear forces.
- [Drop-in fuels from sunlight and air](#) [03 November 2021]
Article • Carbon-neutral hydrocarbon fuels can be produced using sunlight and air via a thermochemical solar fuel production chain, thus representing a pathway towards the long-term decarbonization of the aviation sector.
- [Depressed 660-km discontinuity caused by akimotoite–bridgmanite transition](#) [05 January 2022]
Article • X-ray diffraction experiments indicate that the depression of the Earth’s 660-kilometre seismic discontinuity beneath cold subduction zones is caused by a phase transition from akimotoite to bridgmanite, leading to slab stagnation.
- [Experimental evidence for recovery of mercury-contaminated fish populations](#) [15 December 2021]
Article • In a 15-year whole-ecosystem, single-factor experiment, stopping experimental mercury loading results in rapid decreases in methylmercury contamination of fish populations and almost complete recovery within the timeframe of the study.
- [Cyclic evolution of phytoplankton forced by changes in tropical seasonality](#) [01 December 2021]
Article • Morphometric analysis of coccolith assemblages spanning the last 2,800,000 years suggests that the evolution of coccolithophores is linked to seasonality changes, paced by Earth’s orbital eccentricity with implications for the carbon cycle.
- [Spatial genomics enables multi-modal study of clonal heterogeneity in tissues](#) [15 December 2021]

Article • A technique using barcoded beads for DNA sequencing within tissue sections enables spatial resolution of tumour clonal heterogeneity and can be multiplexed with other analytical techniques for analysis of complex cellular phenotypes.

- **Building an allocentric travelling direction signal via vector computation** [15 December 2021]

Article • A neural circuit for implementing a coordinate transformation and 2D vector computation is described in Drosophila.

- **Transforming representations of movement from body- to world-centric space** [15 December 2021]

Article • Specific neurons in the fan-shaped body of the Drosophila brain convert translational information in relation to the fly's body to externally referenced coordinates for navigation.

- **Local circuit amplification of spatial selectivity in the hippocampus** [01 December 2021]

Article • Single-cell tracing and optogenetics manipulation in mice are used to show how spatial tuning of individual pyramidal cells in CA1 can propagate to and be amplified by their local subnetwork of neurons.

- **Pre-existing polymerase-specific T cells expand in abortive seronegative SARS-CoV-2** [10 November 2021]

Article • Seronegative healthcare workers with an innate signature of infection preferentially expand pre-existing T cells targeting the conserved replication transcription complex of SARS-CoV-2 in abortive infection.

- **Anatomically distinct fibroblast subsets determine skin autoimmune patterns** [15 December 2021]

Article • Single-cell analyses of skin samples from patients with vitiligo and functional genetic experiments in vitiligo mouse models show that distinct fibroblast subsets drive the organ level lesion patterns in this autoimmune disease.

- **Non-genetic determinants of malignant clonal fitness at single-cell resolution** [08 December 2021]

Article • Non-genetic malignant clonal dominance is a cell-intrinsic and heritable property that underpins clonal output and response to therapy in cancer.

- **ER proteins decipher the tubulin code to regulate organelle distribution** [15 December 2021]

Article • The endoplasmic reticulum proteins CLIMP63, kinecin and p180 bind preferentially to subsets of microtubules with different post-translational modifications, thereby linking the 'tubulin code' to the intracellular distribution of membrane organelles.

- **Structure of pathological TDP-43 filaments from ALS with FTLD** [08 December 2021]

Article • Cryo-electron microscopy of aggregated TDP-43 from postmortem brain tissue of individuals who had ALS with FTLD reveals a filament structure with distinct features to other neuropathological protein filaments, such as those of tau and α -synuclein.

- **Time-resolved structural analysis of an RNA-cleaving DNA catalyst** [23 December 2021]

Article • Using high-resolution NMR characterization, the kinetics and dynamics of the catalytic function of a DNAzyme are shown.

- NEWS AND VIEWS
- 01 December 2021

Earth's eccentric orbit paced the evolution of marine phytoplankton

Analysis of plankton fossils has revealed pulses of size diversity that are inextricably linked to the degree of circularity of Earth's orbits. Could this orbital variability provide a beat that dictates the rhythm of evolution?

- [Rosalind E. M. Rickaby](#) 

When considering the factors that shape evolution over time, attention normally focuses on biological aspects. However, [writing in *Nature*](#), Beaufort *et al.*¹ present the case for variations in Earth's orbit as being implicated in determining the outcome of evolution.

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Nature **601**, 27–28 (2022)

doi: <https://doi.org/10.1038/d41586-021-03549-5>

References

1. 1.

Beaufort, L. *et al.* *Nature* **601**, 79–84 (2021).

2. 2.

Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121–1132 (1976).

3. 3.

Imbrie, J. *et al.* *Paleoceanogr. Paleoclimatol.* **8**, 699–735 (1993).

4. 4.

Beaufort, L. & Dollfus, D. *Mar. Micropaleont.* **51**, 57–73 (2004).

5. 5.

Matsuoka, H. & Okada, H. *Proc. Ocean Drill. Prog.* **115**, 255–270 (1990).

6. 6.

Bendif, E. M. *et al. Nature Commun.* **10**, 4234 (2019).

7. 7.

McClelland, H. L. O., Bruggeman, J., Hermoso, M. & Rickaby, R. E. M. *Nature Commun.* **8**, 13511 (2017).

8. 8.

Rickaby, R. E. M. *et al. Earth Planet. Sci. Lett.* **253**, 83–95 (2007).

9. 9.

van Dam, J. A. *et al. Nature* **443**, 687–691 (2006).

10. 10.

Filatov, D. A., Bendif, E. M., Archontikis, O. A., Hagino, K. & Rickaby, R. E. M. *Curr. Biol.* <https://doi.org/10.1016/j.cub.2021.09.073> (2021).

11. 11.

Herbert, T. D. *Proc. Natl Acad. Sci. USA* **94**, 8362–8369 (1997).

12. 12.

Suchéras-Marx, B., Mattioli, E., Pittet, B., Escarguel, G. & Suan, G. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **295**, 281–292 (2010).

13. 13.

Crampton, J. S. *et al.* *Proc. Natl Acad. Sci. USA* **115**, 5686–5691 (2018).

14. 14.

Crampton, J. S. *et al.* *Proc. Natl Acad. Sci. USA* **113**, 6868–6873 (2016).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03549-5>

| [Section menu](#) | [Main menu](#) |

- NEWS AND VIEWS
- 08 December 2021

Aggregates of TDP-43 protein spiral into view

In some neurodegenerative diseases, a protein called TDP-43 forms aggregates in the brain, resulting in neuronal cell death. The structure of these aggregates and their properties have been unveiled.

- [Hana M. Odeh](#)⁰ &
- [James Shorter](#)¹

Amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD) are two devastating neurodegenerative diseases that share disease mechanisms and underlying genetics, and can co-occur in the same individual as ALS–FTLD. ALS is commonly characterized by the deterioration of the motor neurons that control movement, whereas FTLD is associated with the deterioration of neurons in a part of the brain called the cortex, leading to behavioural changes and memory loss. Unfortunately, there are no effective therapies for ALS and FTLD, and early diagnosis is challenging. A common hallmark of both conditions is the abnormal accumulation of a protein called TDP-43 inside degenerating neurons^{1–3}. [Writing in Nature](#), Arseni *et al.*⁴ report the structure of TDP-43 aggregates isolated from the brains of two individuals who had died from ALS–FTLD.

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Nature **601**, 29-30 (2022)

doi: <https://doi.org/10.1038/d41586-021-03605-0>

References

1. 1.

Neumann, M. *et al.* *Science* **314**, 130–133 (2006).

2. 2.

Arai, T. *et al.* *Biochem. Biophys. Res. Commun.* **351**, 602–611 (2006).

3. 3.

Portz, B., Lee, B. L. & Shorter, J. *Trends Biochem. Sci.* **46**, 550–563 (2021).

4. 4.

Arseni, D. *et al.* *Nature* **601**, 139–143 (2022).

5. 5.

Hallegger, M. *et al.* *Cell* **184**, 4680–4696 (2021).

6. 6.

Johnson, B. S. *et al.* *J. Biol. Chem.* **284**, 20329–20339 (2009).

7. 7.

Schweighauser, M. *et al.* *Nature* **585**, 464–469 (2020).

8. 8.

Shi, Y. *et al.* *Nature* **598**, 359–363 (2021).

9. 9.

Cao, Q., Boyer, D. R., Sawaya, M. R., Ge, P. & Eisenberg, D. S. *Nature Struct. Mol. Biol.* **26**, 619–627 (2019).

10. 10.

Li, Q., Babinchak, W. M. & Surewicz, W. K. *Nature Commun.* **12**, 1620 (2021).

11. 11.

Vogler, T. O. *et al.* *Nature* **563**, 508–513 (2018).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03605-0>

| [Section menu](#) | [Main menu](#) |

- NEWS AND VIEWS
- 04 January 2022

From the archive

Efforts to prove Fermat's last theorem, and a bewildered translator tackles a book about birds.

Efforts to prove Fermat's last theorem, and a bewildered translator tackles a book about birds.

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Nature **601**, 30 (2022)

doi: <https://doi.org/10.1038/d41586-021-03740-8>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03740-8>

| [Section menu](#) | [Main menu](#) |

- NEWS AND VIEWS
- 08 December 2021

A heritable, non-genetic road to cancer evolution

Treatment for leukaemia can fail for reasons that are not fully clear. Tracking the progress of individual cellular lineages for this type of cancer offers a way to investigate this phenomenon.

- [Tamara Prieto](#)⁰ &
- [Dan A. Landau](#)¹

The complexity of a tumour's evolving cellular population poses an obstacle to successful therapy. [Writing in Nature](#), Fennell *et al.*¹ present data that pinpoint some of the mechanisms underlying the ability of cancers to evolve to a more aggressive and treatment-resistant form.

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Nature **601**, 31–32 (2022)

doi: <https://doi.org/10.1038/d41586-021-03606-z>

References

1. 1.

Fennell, K. A. *et al.* *Nature* **601**, 125–131 (2022).

2. 2.

Duy, C. *et al.* *Cancer Discov.* **11**, 1542–1561 (2021).

3. 3.

Dhimolea, E. *et al.* *Cancer Cell* **39**, 240–256 (2021).

4. 4.

Rehman, S. K. *et al.* *Cell* **184**, 226–242 (2021).

5. 5.

Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L. & Leibler, S. *Science* **305**, 1622–1625 (2004).

6. 6.

Shlush, L. I. *et al. Nature* **547**, 104–108 (2017).

7. 7.

Hata, A. N. *et al. Nature Med.* **22**, 262–269 (2016).

8. 8.

Shaffer, S. M. *et al. Cell* **182**, 947–959 (2020).

9. 9.

Mikubo, M., Inoue, Y., Liu, G. & Tsao, M.-S. *J. Thorac. Oncol.* **16**, 1798–1809 (2021).

10. 10.

Berger, A. J. *et al. Nature Cancer* **2**, 1055–1070 (2021).

11. 11.

Bell, C. C. *et al. Nature Commun.* **10**, 2723 (2019).

12. 12.

Sharma, S. V. *et al. Cell* **141**, 69–80 (2010).

13. 13.

Pan, H. *et al. Cancer Discov.* **11**, 2266–2281 (2021).

14. 14.

Nam, A. S., Chaligne, R. & Landau, D. A. *Nature Rev. Genet.* **22**, 3–18 (2021).

15. 15.

Chaligne, R. *et al.* *Nature Genet.* **53**, 1469–1479 (2021).

16. 16.

Ludwig, L. S. *et al.* *Cell* **176**, 1325–1339 (2019).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03606-z>

| [Section menu](#) | [Main menu](#) |

- NEWS AND VIEWS
- 05 January 2022

Mirror symmetry validated for proton and its antimatter twin

A comparison of the charge-to-mass ratio for the proton and the antiproton has been performed with record-breaking precision — the best such test yet for a mirror-like symmetry that relates matter and antimatter.

- [Ralf Lehnert](#) 

Merging the theory of relativity and the microscopic world of quantum physics into a single framework has yielded some profound insights into the nature of the Universe. Among them is the concept of antiparticles, the elusive twins of ordinary particles such as the electron, the proton and the neutron. But determining the properties of these particle twins is a difficult task, because they don't mix well — to the extent that, on contact, a particle–antiparticle pair annihilates, leaving only a flash of intense light. However, the theories of relativity and quantum physics also imply that a mirror-like symmetry exists, called CPT invariance, that ensures that a particle and its antiparticle have the same mass, but opposite charge.

[Writing in Nature](#), Borchert *et al.*¹ report the most stringent test of this aspect of mirror symmetry performed so far, by studying the motion of single negatively charged hydrogen ions and antiprotons in an apparatus known as a Penning trap.

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Nature **601**, 32-33 (2022)

doi: <https://doi.org/10.1038/d41586-021-03798-4>

References

1. 1.

Borchert, M. J. *et al.* *Nature* **601**, 53–57 (2022).

2. 2.

Colladay, D. & Kostelecký, V. A. *Phys. Rev. D* **58**, 116002 (1998).

3. 3.

Ulmer, S. *et al.* *Nature* **524**, 196–199 (2015).

4. 4.

Smorra, C. *et al.* *Tech. Design Rep. BASE-STEP* [CERN-SPSC-2021-007, SPSC-TDR-007](#) (2021).

5. 5.

Nitzschke, D. *et al.* *Adv. Quant. Tech.* **3**, 1900133 (2020).

6. 6.

Kostelecký, V. A. & Li, Z. *Phys. Rev. D* **104**, 044054 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03798-4>

| [Section menu](#) | [Main menu](#) |

- Review Article
- [Published: 05 January 2022](#)

Iron pnictides and chalcogenides: a new paradigm for superconductivity

- [Rafael M. Fernandes](#) ORCID: [orcid.org/0000-0002-3584-5180¹](https://orcid.org/0000-0002-3584-5180),
- [Amalia I. Coldea](#) ORCID: [orcid.org/0000-0002-6732-5964²](https://orcid.org/0000-0002-6732-5964),
- [Hong Ding](#) ORCID: [orcid.org/0000-0003-4422-9248^{3,4}](https://orcid.org/0000-0003-4422-9248),
- [Ian R. Fisher](#) ORCID: [orcid.org/0000-0002-1278-7862^{5,6}](https://orcid.org/0000-0002-1278-7862),
- [P. J. Hirschfeld⁷](#) &
- [Gabriel Kotliar^{8,9}](#)

[Nature](#) volume **601**, pages 35–44 (2022)

- 1978 Accesses
- 20 Altmetric
- [Metrics details](#)

Subjects

- [Superconducting properties and materials](#)

Abstract

Superconductivity is a remarkably widespread phenomenon that is observed in most metals cooled to very low temperatures. The ubiquity of such conventional superconductors, and the wide range of associated critical

temperatures, is readily understood in terms of the well-known Bardeen–Cooper–Schrieffer theory. Occasionally, however, unconventional superconductors are found, such as the iron-based materials, which extend and defy this understanding in unexpected ways. In the case of the iron-based superconductors, this includes the different ways in which the presence of multiple atomic orbitals can manifest in unconventional superconductivity, giving rise to a rich landscape of gap structures that share the same dominant pairing mechanism. In addition, these materials have also led to insights into the unusual metallic state governed by the Hund’s interaction, the control and mechanisms of electronic nematicity, the impact of magnetic fluctuations and quantum criticality, and the importance of topology in correlated states. Over the fourteen years since their discovery, iron-based superconductors have proven to be a testing ground for the development of novel experimental tools and theoretical approaches, both of which have extensively influenced the wider field of quantum materials.

This is a preview of subscription content

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Fig. 1: General structural and electronic properties.

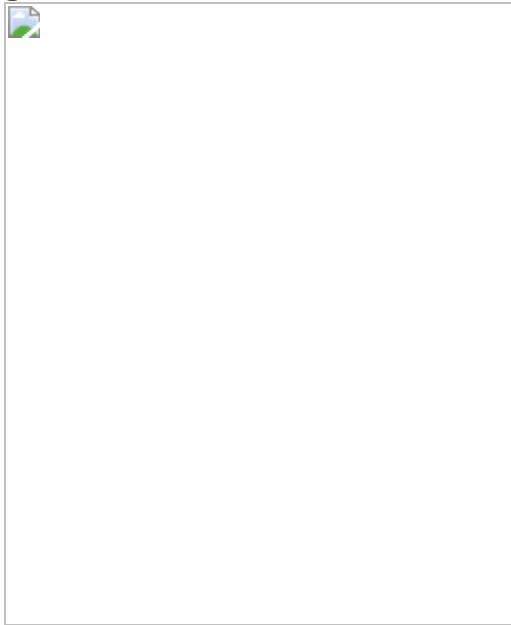


Fig. 2: Electronic correlations and orbital differentiation.

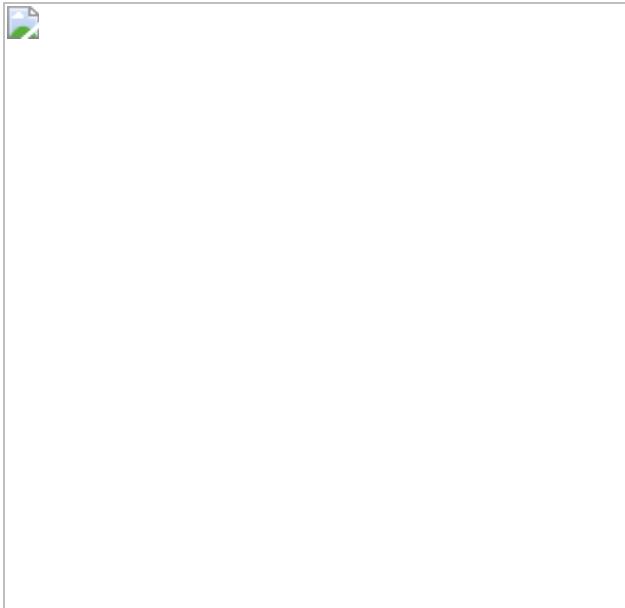


Fig. 3: Dual local–itinerant nature of magnetism.

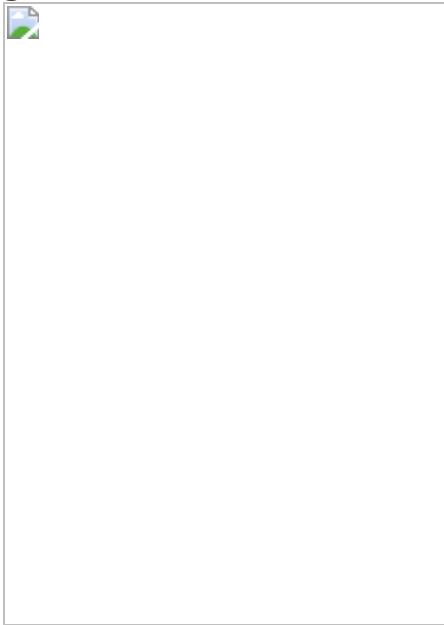


Fig. 4: Electronic nematic order and its coupling to the lattice.



Fig. 5: Superconducting gap structures and gap symmetries.

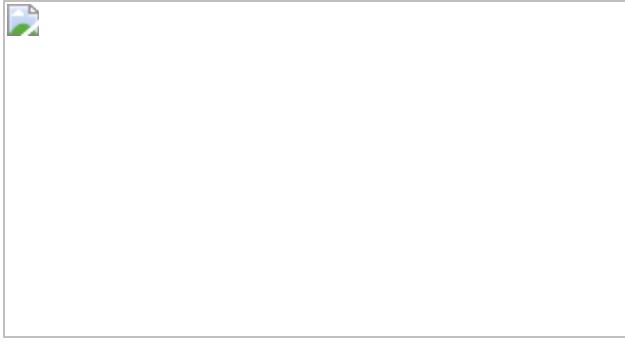
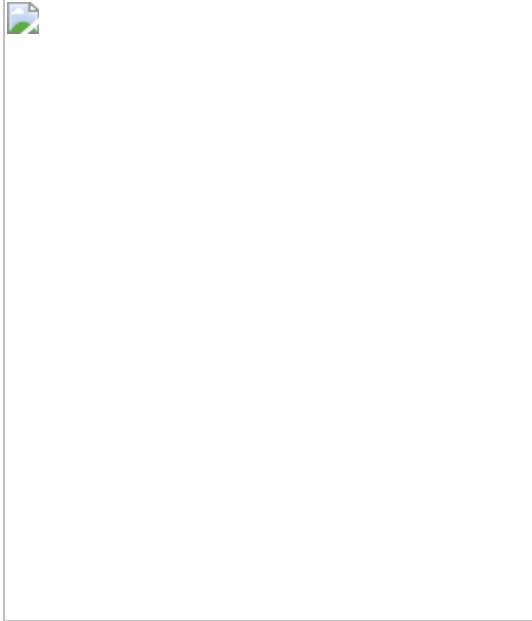


Fig. 6: Band inversion and topological phenomena.



References

1. 1.

Keimer, B., Kivelson, S. A., Norman, M. R., Uchida, S. & Zaanen, J. From quantum matter to high-temperature superconductivity in copper oxides. *Nature* **518**, 179–186 (2015).

2. 2.

Scalapino, D. J. A common thread: the pairing interaction for unconventional superconductors. *Rev. Mod. Phys.* **84**, 1383–1417 (2012).

3. 3.

Kamihara, Y., Watanabe, T., Hirano, M. & Hosono, H. Iron-based layered superconductor $\text{LaO}_{1-x}\text{F}_x\text{FeAs}$ ($x = 0.05\text{--}0.12$) with $T_c = 26\text{ K}$. *J. Am. Chem. Soc.* **130**, 3296–3297 (2008). **The seminal observation of superconductivity in an iron-arsenide compound.**

4. 4.

Mazin, I. I., Singh, D. J., Johannes, M. D. & Du, M. H. Unconventional superconductivity with a sign reversal in the order parameter of $\text{LaFeAsO}_{1-x}\text{F}_x$. *Phys. Rev. Lett.* **101**, 057003 (2008). **Theoretical proposal that the s^{+-} superconducting state in FeSCs is mediated by spin fluctuations.**

5. 5.

Kuroki, K., Usui, H., Onari, S., Arita, R. & Aoki, H. Pnictogen height as a possible switch between high- T_c nodeless and low- T_c nodal pairings in the iron-based superconductors. *Phys. Rev. B* **79**, 224511 (2009). **RPA calculation that shows the impact of the pnictogen height on the superconducting state.**

6. 6.

Hirschfeld, P. J., Korshunov, M. M. & Mazin, I. I. Gap symmetry and structure of Fe-based superconductors. *Rep. Prog. Phys.* **74**, 124508 (2011).

7. 7.

Chubukov, A. V. Pairing mechanism in Fe-based superconductors. *Annu. Rev. Condens. Matter Phys.* **3**, 57–92 (2012). **A pedagogical review that compares the RPA and renormalization group approaches to describe superconductivity in FeSCs.**

8. 8.

Wang, F. & Lee, D.-H. The electron-pairing mechanism of iron-based superconductors. *Science* **332**, 200–204 (2011).

9. 9.

Haule, K. & Kotliar, G. Coherence–incoherence crossover in the normal state of iron oxypnictides and importance of Hund’s rule coupling. *New J. Phys.* **11**, 025021 (2009). **This theoretical work predicted the coherence–incoherence crossover caused by the Hund’s coupling, which later led to the concept of a Hund metal.**

10. 10.

Yin, Z., Haule, K. & Kotliar, G. Kinetic frustration and the nature of the magnetic and paramagnetic states in iron pnictides and iron chalcogenides. *Nat. Mater.* **10**, 932–935 (2011). **This study provides principles for organizing the families of FeSCs by their correlation strength and differentiation of the d_{xy} orbitals.**

11. 11.

Stadler, K. M., Yin, Z. P., von Delft, J., Kotliar, G. & Weichselbaum, A. Dynamical mean-field theory plus numerical renormalization-group study of spin-orbital separation in a three-band Hund metal. *Phys. Rev. Lett.* **115**, 136401 (2015).

12. 12.

de' Medici, L., Hassan, S. R., Capone, M. & Dai, X. Orbital-selective Mott transition out of band degeneracy lifting. *Phys. Rev. Lett.* **102**, 126401 (2009).

13. 13.

Bascones, E., Valenzuela, B. & Calderón, M. J. Orbital differentiation and the role of orbital ordering in the magnetic state of Fe superconductors. *Phys. Rev. B* **86**, 174508 (2012).

14. 14.

Yu, R. & Si, Q. Orbital-selective Mott phase in multiorbital models for alkaline iron selenides $K_{1-x}Fe_{2-y}Se_2$. *Phys. Rev. Lett.* **110**, 146402 (2013).

15. 15.

de' Medici, L., Giovannetti, G. & Capone, M. Selective Mott physics as a key to iron superconductors. *Phys. Rev. Lett.* **112**, 177001 (2014).

16. 16.

Georges, A., Medici, L. D. & Mravlje, J. Strong correlations from Hund's coupling. *Annu. Rev. Condens. Matter Phys.* **4**, 137–178 (2013).

17. 17.

Dai, P. Antiferromagnetic order and spin dynamics in iron-based superconductors. *Rev. Mod. Phys.* **87**, 855–896 (2015).

18. 18.

Lumsden, M. D. & Christianson, A. D. Magnetism in Fe-based superconductors. *J. Phys. Condens. Matter* **22**, 203203 (2010). **A topical review that surveys early neutron scattering data on**

FeSCs, including the observation of spin-resonance modes in the superconducting state.

19. 19.

Inosov, D. et al. Normal-state spin dynamics and temperature-dependent spin-resonance energy in optimally doped BaFe_{1.85}Co_{0.15}As₂. *Nat. Phys.* **6**, 178–181 (2010).

20. 20.

Fernandes, R. M., Chubukov, A. V. & Schmalian, J. What drives nematic order in iron-based superconductors? *Nat. Phys.* **10**, 97–104 (2014).

21. 21.

Fradkin, E., Kivelson, S. A., Lawler, M. J., Eisenstein, J. P. & Mackenzie, A. P. Nematic Fermi fluids in condensed matter physics. *Annu. Rev. Condens. Matter Phys.* **1**, 153–178 (2010).

22. 22.

Chu, J.-H., Kuo, H.-H., Analytis, J. G. & Fisher, I. R. Divergent nematic susceptibility in an iron arsenide superconductor. *Science* **337**, 710–712 (2012). **Elastoresistivity measurements reveal the presence of nematic fluctuations across the phase diagram of an FeSC compound.**

23. 23.

Böhmer, A. E. et al. Nematic susceptibility of hole-doped and electron-doped BaFe₂As₂ iron-based superconductors from shear modulus measurements. *Phys. Rev. Lett.* **112**, 047001 (2014).

24. 24.

Gallais, Y. et al. Observation of incipient charge nematicity in $\text{Ba}(\text{Fe}_{1-x}\text{Co}_x)2\text{As}_2$. *Phys. Rev. Lett.* **111**, 267001 (2013).

25. 25.

Zhang, P. et al. Observation of topological superconductivity on the surface of an iron-based superconductor. *Science* **360**, 182–186 (2018).
ARPES measurements reveal surface topological spin-helical states in $\text{FeTe}_{1-x}\text{Se}_x$

26. 26.

Singh, D. J. & Du, M.-H. Density functional study of $\text{LaFeAsO}_{1-x}\text{F}_x$: a low carrier density superconductor near itinerant magnetism. *Phys. Rev. Lett.* **100**, 237003 (2008).

27. 27.

Eschrig, H. & Koepernik, K. Tight-binding models for the iron-based superconductors. *Phys. Rev. B* **80**, 104503 (2009).

28. 28.

Cvetkovic, V. & Vafek, O. Space group symmetry, spin-orbit coupling, and the low-energy effective Hamiltonian for iron-based superconductors. *Phys. Rev. B* **88**, 134510 (2013).

29. 29.

Borisenko, S. et al. Direct observation of spin-orbit coupling in iron-based superconductors. *Nat. Phys.* **12**, 311–317 (2016).

30. 30.

Wang, Z. et al. Topological nature of the $\text{FeSe}_{0.5}\text{Te}_{0.5}$ superconductor. *Phys. Rev. B* **92**, 115119 (2015).

31. 31.

Yang, W. L. et al. Evidence for weak electronic correlations in iron pnictides. *Phys. Rev. B* **80**, 014508 (2009).

32. 32.

Coldea, A. I. Electronic nematic states tuned by isoelectronic substitution in bulk $\text{FeSe}_{1-x}\text{S}_x$. *Front. Phys.* **8**, 594500 (2021).

33. 33.

Richard, P., Qian, T. & Ding, H. ARPES measurements of the superconducting gap of Fe-based superconductors and their implications to the pairing mechanism. *J. Phys. Condens. Matter* **27**, 293203 (2015).

34. 34.

Yi, M., Zhang, Y., Shen, Z.-X. & Lu, D. Role of the orbital degree of freedom in iron-based superconductors. *npj Quantum Mater.* **2**, 57 (2017).

35. 35.

Carrington, A. Quantum oscillation studies of the Fermi surface of iron-pnictide superconductors. *Rep. Prog. Phys.* **74**, 124507 (2011).

36. 36.

Coldea, A. I. et al. Fermi surface of superconducting LaFePO determined from quantum oscillations. *Phys. Rev. Lett.* **101**, 216402 (2008).

37. 37.

Qazilbash, M. et al. Electronic correlations in the iron pnictides. *Nat. Phys.* **5**, 647–650 (2009).

38. 38.

Haule, K., Shim, J. H. & Kotliar, G. Correlated electronic structure of $\text{LaO}_{1-x}\text{F}_x\text{FeAs}$. *Phys. Rev. Lett.* **100**, 226402 (2008).

39. 39.

Skornyakov, S. L. et al. Classification of the electronic correlation strength in the iron pnictides: the case of the parent compound BaFe_2As_2 . *Phys. Rev. B* **80**, 092501 (2009).

40. 40.

Werner, P. et al. Satellites and large doping and temperature dependence of electronic properties in hole-doped BaFe_2As_2 . *Nat. Phys.* **8**, 331–337 (2012).

41. 41.

Ferber, J., Foyevtsova, K., Valentí, R. & Jeschke, H. O. LDA + DMFT study of the effects of correlation in LiFeAs . *Phys. Rev. B* **85**, 094505 (2012).

42. 42.

Lee, G. et al. Orbital selective Fermi surface shifts and mechanism of high T_c superconductivity in correlated AFeAs (A = Li, Na). *Phys. Rev. Lett.* **109**, 177001 (2012).

43. 43.

Borisenko, S. V. et al. Superconductivity without nesting in LiFeAs . *Phys. Rev. Lett.* **105**, 067002 (2010).

44. 44.

Fanfarillo, L. et al. Orbital-dependent Fermi surface shrinking as a fingerprint of nematicity in FeSe. *Phys. Rev. B* **94**, 155138 (2016).

45. 45.

Ortenzi, L., Cappelluti, E., Benfatto, L. & Pietronero, L. Fermi-surface shrinking and interband coupling in iron-based pnictides. *Phys. Rev. Lett.* **103**, 046404 (2009).

46. 46.

Zantout, K., Backes, S. & Valentí, R. Effect of nonlocal correlations on the electronic structure of LiFeAs. *Phys. Rev. Lett.* **123**, 256401 (2019).

47. 47.

Tomczak, J. M., van Schilfgaarde, M. & Kotliar, G. Many-body effects in iron pnictides and chalcogenides: nonlocal versus dynamic origin of effective masses. *Phys. Rev. Lett.* **109**, 237010 (2012).

48. 48.

van der Marel, D. & Sawatzky, G. A. Electron–electron interaction and localization in *d* and *f* transition metals. *Phys. Rev. B* **37**, 10674 (1988).

49. 49.

Hardy, F. et al. Evidence of strong correlations and coherence–incoherence crossover in the iron pnictide superconductor KFe₂As₂. *Phys. Rev. Lett.* **111**, 027002 (2013).

50. 50.

Yin, Z. P., Haule, K. & Kotliar, G. Fractional power-law behavior and its origin in iron-chalcogenide and ruthenate superconductors: insights from first-principles calculations. *Phys. Rev. B* **86**, 195141 (2012).

51. 51.

Kreisel, A., Hirschfeld, P. J. & Andersen, B. M. On the remarkable superconductivity of FeSe and its close cousins. *Symmetry* **12**, 1402 (2020).

52. 52.

Yu, R., Zhu, J.-X. & Si, Q. Orbital-selective superconductivity, gap anisotropy, and spin resonance excitations in a multiorbital t - J_1 - J_2 model for iron pnictides. *Phys. Rev. B* **89**, 024509 (2014).

53. 53.

Fanfarillo, L., Valli, A. & Capone, M. Synergy between Hund-driven correlations and boson-mediated superconductivity. *Phys. Rev. Lett.* **125**, 177001 (2020).

54. 54.

Sprau, P. O. et al. Discovery of orbital-selective Cooper pairing in FeSe. *Science* **357**, 75–80 (2017). **STM observation of a strong gap anisotropy in FeSe and proposal of orbital differentiation inside the superconducting state.**

55. 55.

Rhodes, L. C. et al. Scaling of the superconducting gap with orbital character in FeSe. *Phys. Rev. B* **98**, 180503 (2018).

56. 56.

Liu, D. et al. Orbital origin of extremely anisotropic superconducting gap in nematic phase of FeSe superconductor. *Phys. Rev. X* **8**, 031033 (2018).

57. 57.

Yin, Z., Haule, K. & Kotliar, G. Spin dynamics and orbital-antiphase pairing symmetry in iron-based superconductors. *Nat. Phys.* **10**, 845–850 (2014).

58. 58.

Pelliciari, J. et al. Magnetic moment evolution and spin freezing in doped BaFe₂As₂. *Sci. Rep.* **7**, 8003 (2017).

59. 59.

Wang, M. et al. Doping dependence of spin excitations and its correlations with high-temperature superconductivity in iron pnictides. *Nat. Commun.* **4**, 2874 (2013).

60. 60.

Christensen, M. H., Kang, J., Andersen, B. M., Eremin, I. & Fernandes, R. M. Spin reorientation driven by the interplay between spin-orbit coupling and Hund's rule coupling in iron pnictides. *Phys. Rev. B* **92**, 214509 (2015).

61. 61.

Qureshi, N. et al. Inelastic neutron-scattering measurements of incommensurate magnetic excitations on superconducting LiFeAs single crystals. *Phys. Rev. Lett.* **108**, 117001 (2012).

62. 62.

Wang, Q. et al. Magnetic ground state of FeSe. *Nat. Commun.* **7**, 12182 (2016).

63. 63.

Lumsden, M. D. et al. Evolution of spin excitations into the superconducting state in FeTe_{1-x}Se_x. *Nat. Phys.* **6**, 182–186 (2010).

64. 64.

Liu, T. et al. From $(\pi, 0)$ magnetic order to superconductivity with (π, π) magnetic resonance in Fe_{1.02}Te_{1-x}Sex. *Nat. Mater.* **9**, 718–720 (2010).

65. 65.

Gastiasoro, M. N. & Andersen, B. M. Enhancement of magnetic stripe order in iron-pnictide superconductors from the interaction between conduction electrons and magnetic impurities. *Phys. Rev. Lett.* **113**, 067002 (2014).

66. 66.

Pratt, D. K. et al. Incommensurate spin-density wave order in electron-doped BaFe₂As₂ superconductors. *Phys. Rev. Lett.* **106**, 257001 (2011).

67. 67.

Allred, J. M. et al. Double- Q spin-density wave in iron arsenide superconductors. *Nat. Phys.* **12**, 493–498 (2016).

68. 68.

Lorenzana, J., Seibold, G., Ortix, C. & Grilli, M. Competing orders in FeAs layers. *Phys. Rev. Lett.* **101**, 186402 (2008).

69. 69.

Fernandes, R. M., Kivelson, S. A. & Berg, E. Vestigial chiral and charge orders from bidirectional spin-density waves: application to the iron-based superconductors. *Phys. Rev. B* **93**, 014511 (2016).

70. 70.

Meier, W. R. et al. Hedgehog spin-vortex crystal stabilized in a hole-doped iron-based superconductor. *npj Quantum Mater.* **3**, 5 (2018).

71. 71.

Si, Q. & Abrahams, E. Strong correlations and magnetic frustration in the high T_c iron pnictides. *Phys. Rev. Lett.* **101**, 076401 (2008).

72. 72.

Seo, K., Bernevig, B. A. & Hu, J. Pairing symmetry in a two-orbital exchange coupling model of oxypnictides. *Phys. Rev. Lett.* **101**, 206404 (2008).

73. 73.

Dai, P., Hu, J. & Dagotto, E. Magnetism and its microscopic origin in iron-based high-temperature superconductors. *Nat. Phys.* **8**, 709–718 (2012).

74. 74.

Eremin, I. & Chubukov, A. V. Magnetic degeneracy and hidden metallicity of the spin-density-wave state in ferropnictides. *Phys. Rev. B* **81**, 024511 (2010).

75. 75.

Fernandes, R. M. & Chubukov, A. V. Low-energy microscopic models for iron-based superconductors: a review. *Rep. Prog. Phys.* **80**, 014503 (2016).

76. 76.

Yildirim, T. Origin of the 150-K anomaly in LaFeAsO: competing antiferromagnetic interactions, frustration, and a structural phase transition. *Phys. Rev. Lett.* **101**, 057010 (2008).

77. 77.

Glasbrenner, J. et al. Effect of magnetic frustration on nematicity and superconductivity in iron chalcogenides. *Nat. Phys.* **11**, 953–958 (2015).

78. 78.

Hirayama, M., Misawa, T., Miyake, T. & Imada, M. Ab initio studies of magnetism in the iron chalcogenides FeTe and FeSe. *J. Phys. Soc. Jpn* **84**, 093703 (2015).

79. 79.

Abrahams, E. & Si, Q. Quantum criticality in the iron pnictides and chalcogenides. *J. Phys. Condens. Matter* **23**, 223201 (2011).

80. 80.

Shibauchi, T., Carrington, A. & Matsuda, Y. A quantum critical point lying beneath the superconducting dome in iron pnictides. *Annu. Rev. Condens. Matter Phys.* **5**, 113–135 (2014). **A review of the evidence of quantum critical behaviour in FeSCs, including the observation of a sharp peak in the doping-dependent penetration depth.**

81. 81.

Hayes, I. M. et al. Scaling between magnetic field and temperature in the high-temperature superconductor $\text{BaFe}_2(\text{As}_{1-x}\text{P}_x)_2$. *Nat. Phys.* **12**, 916–919 (2016).

82. 82.

Chowdhury, D., Swingle, B., Berg, E. & Sachdev, S. Singularity of the London penetration depth at quantum critical points in superconductors. *Phys. Rev. Lett.* **111**, 157004 (2013).

83. 83.

Levchenko, A., Vavilov, M. G., Khodas, M. & Chubukov, A. V. Enhancement of the London penetration depth in pnictides at the onset of spin-density-wave order under superconducting dome. *Phys. Rev. Lett.* **110**, 177003 (2013).

84. 84.

Lu, X. et al. Nematic spin correlations in the tetragonal state of uniaxial-strained $\text{BaFe}_{2-x}\text{Ni}_x\text{As}_2$. *Science* **345**, 657–600 (2014).

Inelastic neutron scattering experiments in a detwinned FeSC

compound reveal the intertwining between nematic order and spin fluctuations.

85. 85.

Chu, J.-H. et al. In-plane resistivity anisotropy in an underdoped iron arsenide superconductor. *Science* **329**, 824–826 (2010).

86. 86.

Mirri, C. et al. Origin of the resistive anisotropy in the electronic nematic phase of BaFe_2As_2 revealed by optical spectroscopy. *Phys. Rev. Lett.* **115**, 107001 (2015).

87. 87.

Chuang, T.-M. et al. Nematic electronic structure in the “parent” state of the iron-based superconductor $\text{Ca}(\text{Fe}_{1-x}\text{Co}_x)_2\text{As}_2$. *Science* **327**, 181–184 (2010).

88. 88.

Liang, S., Moreo, A. & Dagotto, E. Nematic state of pnictides stabilized by interplay between spin, orbital, and lattice degrees of freedom. *Phys. Rev. Lett.* **111**, 047004 (2013).

89. 89.

Lee, C.-C., Yin, W.-G. & Ku, W. Ferro-orbital order and strong magnetic anisotropy in the parent compounds of iron-pnictide superconductors. *Phys. Rev. Lett.* **103**, 267001 (2009).

90. 90.

Lv, W., Krüger, F. & Phillips, P. Orbital ordering and unfrustrated $(\pi, 0)$ magnetism from degenerate double exchange in the iron pnictides. *Phys. Rev. B* **82**, 045125 (2010).

91. 91.

Fang, C., Yao, H., W.-F. Tsai, J. Hu, & S. A. Kivelson, Theory of electron nematic order in LaFeAsO. *Phys. Rev. B* **77**, 224509 (2008).

92. 92.

Xu, C., Müller, M. & Sachdev, S. Ising and spin orders in the iron-based superconductors. *Phys. Rev. B* **78**, 020501 (2008).

93. 93.

Fernandes, R. M., Orth, P. P. & Schmalian, J. Intertwined vestigial order in quantum materials: nematicity and beyond. *Annu. Rev. Condens. Matter Phys.* **10**, 133–154 (2019).

94. 94.

Wang, F., Kivelson, S. A. & LeeD.-H. Nematicity and quantum paramagnetism in FeSe. *Nat. Phys.* **11**, 959–963 (2015).

95. 95.

Fernandes, R. M., Chubukov, A. V., Knolle, J., Eremin, I. & Schmalian, J. Preemptive nematic order, pseudogap, and orbital order in the iron pnictides. *Phys. Rev. B* **85**, 024534 (2012).

96. 96.

Gati, E., Xiang, L., Bud'ko, S. L. & Canfield, P. C. Role of the Fermi surface for the pressure-tuned nematic transition in the BaFe₂As₂ family. *Phys. Rev. B* **100**, 064512 (2019).

97. 97.

Fernandes, R. M., Böhmer, A. E., Meingast, C. & Schmalian, J. Scaling between magnetic and lattice fluctuations in iron pnictide superconductors. *Phys. Rev. Lett.* **111**, 137001 (2013).

98. 98.

Baek, S. et al. Orbital-driven nematicity in FeSe. *Nat. Mater.* **14**, 210–214 (2015).

99. 99.

Böhmer, A. E. et al. Distinct pressure evolution of coupled nematic and magnetic orders in FeSe. *Phys. Rev. B* **100**, 064515 (2019).

100. 100.

Suzuki, Y. et al. Momentum-dependent sign inversion of orbital order in superconducting FeSe. *Phys. Rev. B* **92**, 205117 (2015).

101. 101.

Lederer, S., Schattner, Y., Berg, E. & Kivelson, S. A. Superconductivity and non-Fermi liquid behavior near a nematic quantum critical point. *Proc. Natl Acad. Sci. USA* **114**, 4905–4910 (2017).

102. 102.

Klein, A. & Chubukov, A. V. Superconductivity near a nematic quantum critical point: interplay between hot and lukewarm regions. *Phys. Rev. B* **98**, 220501 (2018).

103. 103.

Worasaran, T. et al. Nematic quantum criticality in an Fe-based superconductor revealed by strain-tuning. *Science* **372**, 973–977 (2021).

104. 104.

Shibauchi, T., Hanaguri, T. & Matsuda, Y. Exotic superconducting states in FeSe-based materials. *J. Phys. Soc. Jpn* **89**, 102002 (2020).

105. 105.

Reiss, P. et al. Quenched nematic criticality and two superconducting domes in an iron-based superconductor. *Nat. Phys.* **16**, 89–94 (2020).

106. 106.

Huang, D. & Hoffman, J. E. Monolayer FeSe on SrTiO₃. *Annu. Rev. Condens. Matter Phys.* **8**, 311–336 (2017).

107. 107.

Hosono, H., Yamamoto, A., Hiramatsu, H. & Ma, Y. Recent advances in iron-based superconductors toward applications. *Mater. Today* **21**, 278–302 (2018).

108. 108.

Boeri, L., Dolgov, O. V. & Golubov, A. A. Is LaFeAsO_{1-x}F_x an electron–phonon superconductor? *Phys. Rev. Lett.* **101**, 026403 (2008).

109. 109.

Mandal, S., Cohen, R. E. & Haule, K. Strong pressure-dependent electron–phonon coupling in FeSe. *Phys. Rev. B* **89**, 220502 (2014).

110. 110.

Lee, J. et al. Interfacial mode coupling as the origin of the enhancement of T_c in FeSe films on SrTiO₃. *Nature* **515**, 245–248 (2014). **The observation of a connection between a substrate phonon mode and the enhancement of superconductivity in monolayer FeSe grown on SrTiO₃.**

111. 111.

Thomale, R., Platt, C., Hanke, W., Hu, J. & Bernevig, B. A. Exotic d -wave superconducting state of strongly hole-doped K_xBa_{1-x}Fe₂As₂.

Phys. Rev. Lett. **107**, 117001 (2011).

112. 112.

Paul, I. & Garst, M. Lattice effects on nematic quantum criticality in metals. *Phys. Rev. Lett.* **118**, 227601 (2017).

113. 113.

Kontani, H. & Onari, S. Orbital-fluctuation-mediated superconductivity in iron pnictides: analysis of the five-orbital Hubbard–Holstein model. *Phys. Rev. Lett.* **104**, 157001 (2010).

114. 114.

Chen, C.-T., Tsuei, C., Ketchen, M., Ren, Z.-A. & Zhao, Z. Integer and half-integer flux-quantum transitions in a niobium-iron pnictide loop. *Nat. Phys.* **6**, 260–264 (2010).

115. 115.

Hanaguri, T., Niitaka, S., Kuroki, K. & Takagi, H. Unconventional *s*-wave superconductivity in Fe(Se,Te). *Science* **328**, 474–476 (2010).

116. 116.

Cho, K., Kończykowski, M., Teknowijoyo, S., Tanatar, M. A. & Prozorov, R. Using electron irradiation to probe iron-based superconductors. *Supercond. Sci. Technol.* **31**, 064002 (2018).

117. 117.

Yang, H. et al. In-gap quasiparticle excitations induced by non-magnetic Cu impurities in Na(Fe_{0.96}Co_{0.03}Cu_{0.01})As revealed by scanning tunnelling spectroscopy. *Nat. Commun.* **4**, 2749 (2013).

118. 118.

Okazaki, K. et al. Octet-line node structure of superconducting order parameter in KFe_2As_2 . *Science* **337**, 1314–1317 (2012). **Direct observation of accidental nodes in a hole-doped FeSC compound via laser ARPES measurements.**

119. 119.

Lee, T.-H., Chubukov, A. V., Miao, H. & Kotliar, G. Pairing mechanism in Hund's metal superconductors and the universality of the superconducting gap to critical temperature ratio. *Phys. Rev. Lett.* **121**, 187003 (2018).

120. 120.

Stanev, V. & Tešanović, Z. Three-band superconductivity and the order parameter that breaks time-reversal symmetry. *Phys. Rev. B* **81**, 134522 (2010).

121. 121.

Lee, W.-C., Zhang, S.-C. & Wu, C. Pairing state with a time-reversal symmetry breaking in FeAs-based superconductors. *Phys. Rev. Lett.* **102**, 217002 (2009).

122. 122.

Grinenko, V. et al. Superconductivity with broken time-reversal symmetry inside a superconducting s-wave state. *Nat. Phys.* **16**, 789–794 (2020).

123. 123.

Kretzschmar, F. et al. Raman-scattering detection of nearly degenerate s-wave and d-wave pairing channels in iron-based $\text{Ba}_{0.6}\text{K}_{0.4}\text{Fe}_2\text{As}_2$ and $\text{Rb}_{0.8}\text{Fe}_{1.6}\text{Se}_2$ superconductors. *Phys. Rev. Lett.* **110**, 187002 (2013).

124. 124.

Thorsmølle, V. K. et al. Critical quadrupole fluctuations and collective modes in iron pnictide superconductors. *Phys. Rev. B* **93**, 054515 (2016).

125. 125.

Gallais, Y., Paul, I., Chauvière, L. & Schmalian, J. Nematic resonance in the Raman response of iron-based superconductors. *Phys. Rev. Lett.* **116**, 017001 (2016).

126. 126.

Tafti, F. et al. Sudden reversal in the pressure dependence of T_c in the iron-based superconductor KFe_2As_2 . *Nat. Phys.* **9**, 349–352 (2013).

127. 127.

Rinott, S. et al. Tuning across the BCS–BEC crossover in the multiband superconductor $\text{Fe}_{1+y}\text{Se}_x\text{Te}_{1-x}$: an angle-resolved photoemission study. *Sci. Adv.* **3**, e1602372 (2017).

128. 128.

Lohani, H. et al. Band inversion and topology of the bulk electronic structure in $\text{FeSe}_{0.45}\text{Te}_{0.55}$. *Phys. Rev. B* **101**, 245146 (2020).

129. 129.

Zhang, P. et al. Multiple topological states in iron-based superconductors. *Nat. Phys.* **15**, 41–47 (2019).

130. 130.

König, E. J. & Coleman, P. Crystalline-symmetry-protected helical Majorana modes in the iron pnictides. *Phys. Rev. Lett.* **122**, 207001 (2019).

131. 131.

Kong, L. et al. Half-integer level shift of vortex bound states in an iron-based superconductor. *Nat. Phys.* **15**, 1181–1187 (2019).

132. 132.

Wang, D. et al. Evidence for Majorana bound states in an iron-based superconductor. *Science* **362**, 333–335 (2018). **STM measurements reveal a zero-bias peak inside vortices of superconducting FeTe_{1-x}Se_x suggestive of Majorana zero modes.**

133. 133.

Machida, T. et al. Zero-energy vortex bound state in the superconducting topological surface state of Fe(Se,Te). *Nat. Mater.* **18**, 811–815 (2019).

134. 134.

Yin, J.-X. et al. Observation of a robust zero-energy bound state in iron-based superconductor Fe(Te,Se). *Nat. Phys.* **11**, 543–546 (2015).

135. 135.

Chen, C. et al. Atomic line defects and zero-energy end states in monolayer Fe(Te,Se) high-temperature superconductors. *Nat. Phys.* **16**, 536–540 (2020).

136. 136.

Wang, Z. et al. Evidence for dispersing 1D Majorana channels in an iron-based superconductor. *Science* **367**, 104–108 (2020).

137. 137.

Zhang, R.-X., Cole, W. S. & Das Sarma, S. Helical hinge Majorana modes in iron-based superconductors. *Phys. Rev. Lett.* **122**, 187001 (2019).

138. 138.

Misawa, T., Nakamura, K. & Imada, M. Ab initio evidence for strong correlation associated with Mott proximity in iron-based superconductors. *Phys. Rev. Lett.* **108**, 177007 (2012).

139. 139.

Aichhorn, M., Biermann, S., Miyake, T., Georges, A. & Imada, M. Theoretical evidence for strong correlations and incoherent metallic state in FeSe. *Phys. Rev. B* **82**, 064504 (2010).

140. 140.

Miyake, T., Nakamura, K., Arita, R. & Imada, M. Comparison of ab initio low-energy models for LaFePO, LaFeAsO, BaFe₂As₂, LiFeAs, FeSe, and FeTe: electron correlation and covalency. *J. Phys. Soc. Jpn* **79**, 044705 (2010).

141. 141.

Zaki, N., Gu, G., Tsvelik, A., Wu, C. & Johnson, P. D. Time-reversal symmetry breaking in the Fe-chalcogenide superconductors. *Proc. Natl Acad. Sci. USA* **118**, e2007241118 (2021).

142. 142.

Kong, L. et al. Majorana zero modes in impurity-assisted vortex of LiFeAs superconductor. *Nat. Commun.* **12**, 4146 (2021).

143. 143.

Karahasanovic, U. & Schmalian, J. Elastic coupling and spin-driven nematicity in iron-based superconductors. *Phys. Rev. B* **93**, 064520 (2016).

144. 144.

Dioguardi, A. P. et al. NMR evidence for inhomogeneous glassy behavior driven by nematic fluctuations in iron arsenide

superconductors. *Phys. Rev. B* **92**, 165116 (2015).

145. 145.

Frandsen, B. A., Wang, Q., Wu, S., Zhao, J. & Birgeneau, R. J. Quantitative characterization of short-range orthorhombic fluctuations in FeSe through pair distribution function analysis. *Phys. Rev. B* **100**, 020504 (2019).

146. 146.

Kuo, H.-H., Chu, J.-H., Palmstrom, J. C., Kivelson, S. A. & Fisher, I. R. Ubiquitous signatures of nematic quantum criticality in optimally doped Fe-based superconductors. *Science* **352**, 958–962 (2016).

147. 147.

Vafek, O. & Chubukov, A. V. Hund interaction, spin–orbit coupling, and the mechanism of superconductivity in strongly hole-doped iron pnictides. *Phys. Rev. Lett.* **118**, 087003 (2017).

148. 148.

Katayama, N. et al. Superconductivity in $\text{Ca}_{1-x}\text{La}_x\text{FeAs}_2$: a novel 112-type iron pnictide with arsenic zigzag bonds. *J. Phys. Soc. Jpn* **82**, 123702 (2013).

149. 149.

Dagotto, E. Colloquium: The unexpected properties of alkali metal iron selenide superconductors. *Rev. Mod. Phys.* **85**, 849–867 (2013).

150. 150.

Wu, S., Frandsen, B. A., Wang, M., Yi, M. & Birgeneau, R. Iron-based chalcogenide spin ladder BaFe_2X_3 ($\text{X} = \text{Se}, \text{S}$). *J. Supercond. Nov. Magn.* **33**, 143–158 (2020).

151. 151.

Momma, K. & Izumi, F. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. *J. Appl. Crystallogr.* **44**, 1272–1276 (2011).

152. 152.

Kong, L. & Ding, H. Emergent vortex Majorana zero mode in iron-based superconductors. *Acta Phys. Sin.* **69**, 110301 (2020).

Acknowledgements

We thank all our co-authors and collaborators with whom we have had many discussions since the discovery of the iron-based superconductors. In particular, we thank H. Miao and T. H. Lee (Figs. 2 and 5), M. Christensen (Fig. 4) and L.-Y. Kong (Fig. 6) for their assistance in making some of the figures panels. R.M.F. was supported by the US Department of Energy, Office of Science, Basic Energy Sciences, Materials Science and Engineering Division, under award number DE-SC0020045. A.I.C. acknowledges an EPSRC Career Acceleration Fellowship (EP/I004475/1) and the Oxford Centre for Applied Superconductivity (CFAS) for financial support. A.I.C. is grateful to the KITP programme ‘correlated20’, which was supported in part by the National Science Foundation under grant number NSF PHY-1748958. H.D. is supported by the National Natural Science Foundation of China (grant numbers 11888101 and 11674371), the Strategic Priority Research Program of the Chinese Academy of Sciences, China (grant numbers XDB28000000 and XDB07000000) and the Beijing Municipal Science and Technology Commission, China (grant number Z191100007219012). I.R.F. was supported by the US Department of Energy, Office of Basic Energy Sciences, under contract DE-AC02-76SF00515. P.J.H. was supported by the US Department of Energy, Office of Basic Sciences under grant number DE-FG02-05ER46236. G.K. was supported by NSF DMR-1733071.

Author information

Affiliations

1. School of Physics and Astronomy, University of Minnesota,
Minneapolis, MN, USA

Rafael M. Fernandes

2. Clarendon Laboratory, Department of Physics, University of Oxford,
Oxford, UK

Amalia I. Coldea

3. Beijing National Laboratory for Condensed Matter Physics and
Institute of Physics, Chinese Academy of Sciences, Beijing, China

Hong Ding

4. CAS Center for Excellence in Topological Quantum Computation,
University of Chinese Academy of Sciences, Beijing, China

Hong Ding

5. Geballe Laboratory for Advanced Materials and Department of
Applied Physics, Stanford University, Stanford, CA, USA

Ian R. Fisher

6. Stanford Institute for Materials and Energy Science, SLAC National
Accelerator Laboratory, Menlo Park, CA, USA

Ian R. Fisher

7. Department of Physics, University of Florida, Gainesville, FL, USA

P. J. Hirschfeld

8. Physics and Astronomy Department, Rutgers University, Piscataway,
NJ, USA

Gabriel Kotliar

9. Condensed Matter Physics and Materials Science Department,
Brookhaven National Laboratory, Upton, NY, USA

Gabriel Kotliar

Contributions

All authors contributed to the text.

Corresponding author

Correspondence to [Rafael M. Fernandes](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review information

Nature thanks Lara Benfatto and Masatoshi Imada and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Fernandes, R.M., Coldea, A.I., Ding, H. *et al.* Iron pnictides and chalcogenides: a new paradigm for superconductivity. *Nature* **601**, 35–44 (2022). <https://doi.org/10.1038/s41586-021-04073-2>

- Received: 25 March 2021
- Accepted: 29 September 2021
- Published: 05 January 2022
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04073-2>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04073-2>

- Article
- [Published: 05 January 2022](#)

A stellar stream remnant of a globular cluster below the metallicity floor

- [Nicolas F. Martin](#) [ORCID: orcid.org/0000-0002-1349-202X](#)^{1,2},
- [Kim A. Venn](#) [ORCID: orcid.org/0000-0003-4134-2042](#)³,
- [David S. Aguado](#) [ORCID: orcid.org/0000-0001-5200-3973](#)^{4,5,6},
- [Else Starkenburg](#)⁷,
- [Jonay I. González Hernández](#)^{5,8},
- [Rodrigo A. Ibata](#) [ORCID: orcid.org/0000-0002-3292-9709](#)¹,
- [Piercarlo Bonifacio](#)⁹,
- [Elisabetta Caffau](#)⁹,
- [Federico Sestito](#) [ORCID: orcid.org/0000-0002-3182-3574](#)³,
- [Anke Arentsen](#)¹,
- [Carlos Allende Prieto](#) [ORCID: orcid.org/0000-0002-0084-572X](#)^{5,8},
- [Raymond G. Carlberg](#) [ORCID: orcid.org/0000-0002-7667-0081](#)¹⁰,
- [Sébastien Fabbro](#) [ORCID: orcid.org/0000-0003-2239-7988](#)^{3,11},
- [Morgan Fouesneau](#)²,
- [Vanessa Hill](#) [ORCID: orcid.org/0000-0002-3795-0005](#)¹²,
- [Pascale Jablonka](#)^{9,13},
- [Georges Kordopatis](#)¹²,
- [Carmela Lardo](#)¹⁴,
- [Khyati Malhan](#)¹⁵,
- [Lyudmila I. Mashonkina](#)¹⁶,
- [Alan W. McConnachie](#)¹¹,
- [Julio F. Navarro](#)³,
- [Rubén Sánchez-Janssen](#)¹⁷,

- [Guillaume F. Thomas^{5,8}](#),
- [Zhen Yuan¹](#) &
- [Alessio Mucciarelli ORCID: orcid.org/0000-0001-9158-8580^{14,18}](#)

[*Nature*](#) volume **601**, pages 45–48 (2022)

- 720 Accesses
- 231 Altmetric
- [Metrics details](#)

Subjects

- [Galaxies and clusters](#)
- [Stars](#)

Abstract

Stellar ejecta gradually enrich the gas out of which subsequent stars form, making the least chemically enriched stellar systems direct fossils of structures formed in the early Universe¹. Although a few hundred stars with metal content below 1,000th of the solar iron content are known in the Galaxy^{2,3,4}, none of them inhabit globular clusters, some of the oldest known stellar structures. These show metal content of at least approximately 0.2% of the solar metallicity $([\rm{Fe}/\rm{H}] \gtrsim -2.7)$. This metallicity floor appears universal^{5,6}, and it has been proposed that protogalaxies that merged into the galaxies we observe today were simply not massive enough to form clusters that survived to the present day⁷. Here we report observations of a stellar stream, C-19, whose metallicity is less than 0.05% of the solar metallicity $([\rm{Fe}/\rm{H}] = -3.38 \pm 0.06, (\rm{s} \pm \rm{t}) \approx 0.20)$. The low metallicity dispersion and the chemical abundances of the C-19 stars show

that this stream is the tidal remnant of the most metal-poor globular cluster ever discovered, and is significantly below the purported metallicity floor: clusters with significantly lower metallicities than observed today existed in the past and contributed their stars to the Milky Way halo.

This is a preview of subscription content

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Properties of the C-19 member stars.



Fig. 2: Metallicity properties of C-19 and its stars observed by spectroscopy.

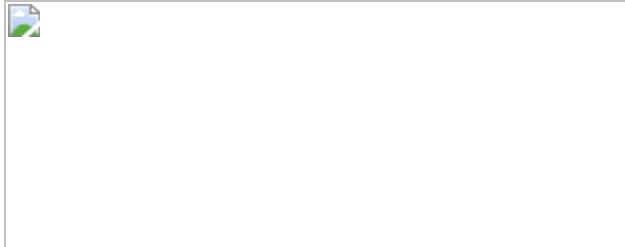
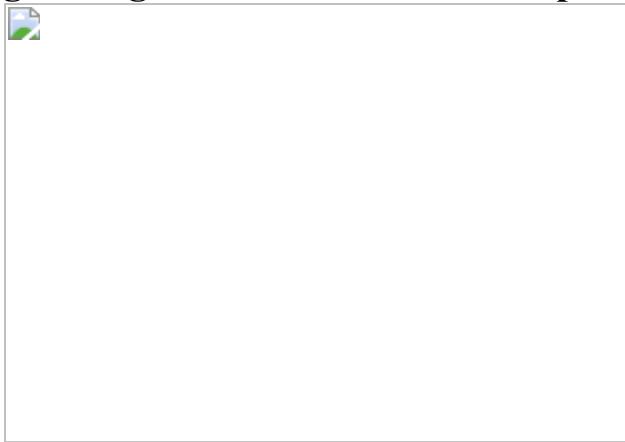


Fig. 3: High-resolution GRACES spectra for three members of C-19.



Data availability

The data used in this article are listed in Extended Data Tables [1–5](#).

Code availability

The codes used for the analysis were not designed to be made public but can be requested from the corresponding author.

References

1. 1.
Frebel, A. & Norris, J. E. Near-field cosmology with extremely metal-poor stars. *Ann. Rev. Astron. Astrophys.* **53**, 631–688 (2015).
2. 2.
Yong, D. et al. The most metal-poor stars. III. The metallicity distribution function and carbon-enhanced metal-poor fraction. *Astrophys. J.* **762**, 27 (2013).
3. 3.
Li, H., Tan, K. & Zhao, G. A catalog of 10,000 very metal-poor stars from LAMOST DR3. *Astrophys. J. Supp.* **238**, 16 (2018).
4. 4.
Aguado, D. S. et al. The Pristine survey - VI. The first three years of medium-resolution follow-up spectroscopy of Pristine EMP star candidates. *Mon. Not. R. Astron. Soc.* **490**, 2241–2253 (2019).
5. 5.
Beasley, M. A. et al. An old, metal-poor globular cluster in Sextans A and the metallicity floor of globular cluster systems. *Mon. Not. R. Astron. Soc.* **487**, 1986–1993 (2019).
6. 6.
Wan, Z. et al. The tidal remnant of an unusually metal-poor globular cluster. *Nature* **583**, 768–770 (2020).

7. 7.

Kruijssen, J. M. D. The minimum metallicity of globular clusters and its physical origin - implications for the galaxy mass-metallicity relation and observations of proto-globular clusters at high redshift. *Mon. Not. R. Astron. Soc.* **486**, L20–L25 (2019).

8. 8.

Ibata, R. et al. Charting the Galactic acceleration field. I. A search for stellar streams with Gaia DR2 and EDR3 with follow-up from ESPaDOnS and UVES. *Astrophys. J.* **914**, 123 (2021).

9. 9.

Gaia Collaboration et al. Gaia Early Data Release 3. Summary of the contents and survey properties. *Astron. Astrophys.* **649**, A1 (2021).

10. 10.

Starkenburg, E. et al. The Pristine survey - I. Mining the Galaxy for the most metal-poor stars. *Mon. Not. R. Astron. Soc.* **471**, 2587–2604 (2017).

11. 11.

Harris, W. E. A catalog of parameters for globular clusters in the Milky Way. *Astron. J.* **112**, 1487 (1996).

12. 12.

Willman, B. & Strader, J. "Galaxy," defined. *Astron. J.* **144**, 76 (2012).

13. 13.

Leaman, R. Insights into pre-enrichment of star clusters and self-enrichment of dwarf galaxies from their intrinsic metallicity dispersions. *Astron. J.* **144**, 183 (2012).

14. 14.

Kirby, E. N. et al. The universal stellar mass-stellar metallicity relation for dwarf galaxies. *Astrophys. J.* **779**, 102 (2013).

15. 15.

Gratton, R. G., Carretta, E. & Bragaglia, A. Multiple populations in globular clusters. Lessons learned from the Milky Way globular clusters. *Astron. Astrophys. R.* **20**, 50 (2012).

16. 16.

Bastian, N. & Lardo, C. Multiple stellar populations in globular clusters. *Ann. Rev. Astron. Astrophys.* **56**, 83–136 (2018).

17. 17.

Ji, A. P. et al. The Southern Stellar Stream Spectroscopic Survey (S⁵): chemical abundances of seven stellar streams. *Astron. J.* **160**, 181 (2020).

18. 18.

Roederer, I. U. Are there any stars lacking neutron-capture elements? Evidence from strontium and barium. *Astron. J.* **145**, 26 (2013).

19. 19.

Côté, B. et al. Neutron star mergers might not be the only source of r-process elements in the Milky Way. *Astrophys. J.* **875**, 106 (2019).

20. 20.

Ji, A. P., Frebel, A., Simon, J. D. & Chiti, A. Complete element abundances of nine stars in the r-process galaxy Reticulum II. *Astrophys. J.* **830**, 93 (2016).

21. 21.

Hansen, T. T. et al. An r-process enhanced star in the dwarf galaxy Tucana III. *Astrophys. J.* **838**, 44 (2017).

22. 22.

Roederer, I. U. Primordial r-process dispersion in metal-poor globular clusters. *Astrophys. J. Lett.* **732**, L17 (2011).

23. 23.

Yoon, J. et al. Galactic archeology with the AEGIS survey: the evolution of carbon and iron in the Galactic halo. *Astrophys. J.* **861**, 146 (2018).

24. 24.

Norris, J. E. et al. The most metal-poor stars. IV. The two populations with $[\text{Fe}/\text{H}] < -3.0$. *Astrophys. J.* **762**, 28 (2013).

25. 25.

Youakim, K. et al. The Pristine survey - VIII. The metallicity distribution function of the Milky Way halo down to the extremely metal-poor regime. *Mon. Not. R. Astron. Soc.* **492**, 4986–5002 (2020).

26. 26.

Roederer, I. U. & Gnedin, O. Y. High-resolution optical spectroscopy of stars in the Sylgr stellar stream. *Astrophys. J.* **883**, 84 (2019).

27. 27.

Larsen, S. S., Romanowsky, A. J., Brodie, J. P. & Wasserman, A. An extremely metal-deficient globular cluster in the Andromeda galaxy. *Science* **370**, 970–973 (2020).

28. 28.

Bonaca, A., Hogg, D. W., Price-Whelan, A. M. & Conroy, C. The spur and the gap in GD-1: dynamical evidence for a dark substructure in the Milky Way halo. *Astrophys. J.* **880**, 38 (2019).

29. 29.

Ibata, R. A., Lewis, G. F. & Martin, N. F. Feeling the pull: a study of natural galactic accelerometers. I. Photometry of the delicate stellar stream of the Palomar 5 globular cluster. *Astrophys. J.* **819**, 1 (2016).

30. 30.

Erkal, D., Koposov, S. E. & Belokurov, V. A sharper view of Pal 5's tails: discovery of stream perturbations with a novel non-parametric technique. *Mon. Not. R. Astron. Soc.* **470**, 60–84 (2017).

31. 31.

Ma, X. et al. The origin and evolution of the galaxy mass-metallicity relation. *Mon. Not. R. Astron. Soc.* **456**, 2140–2156 (2016).

32. 32.

Kielty, C. L. et al. The Pristine survey - XII. Gemini-GRACES chemo-dynamical study of newly discovered extremely metal-poor stars in the Galaxy. *Mon. Not. R. Astron. Soc.* **506**, 1438–1461 (2021).

33. 33.

Malhan, K., Ibata, R. A. & Martin, N. F. Ghostly tributaries to the Milky Way: charting the halo's stellar streams with the Gaia DR2 catalogue. *Mon. Not. R. Astron. Soc.* **481**, 3442–3455 (2018).

34. 34.

Carlberg, R. G. Globular clusters in a cosmological N-body simulation. *Astrophys. J.* **861**, 69 (2018).

35. 35.

Malhan, K., Ibata, R. A., Carlberg, R. G., Valluri, M. & Freese, K. Butterfly in a cocoon, understanding the origin and morphology of globular cluster streams: the case of GD-1. *Astrophys. J.* **881**, 106 (2019).

36. 36.

Martin, N. F. et al. The Pristine survey – XVI. The metallicity of 21 stellar streams around the Milky Way detected with the STREAMFINDER in Gaia EDR3. *Mon. Not. R. Astron. Soc submitted* (2021).

37. 37.

Bressan, A. et al. PARSEC: stellar tracks and isochrones with the PAdova and TRieste stellar evolution code. *Mon. Not. R. Astron. Soc.* **427**, 127–145 (2012).

38. 38.

Maraston, C. Evolutionary population synthesis: models, analysis of the ingredients and application to high-z galaxies. *Mon. Not. R. Astron. Soc.* **362**, 799–825 (2005).

39. 39.

Deason, A. J., Belokurov, V. & Evans, N. W. The Milky Way stellar halo out to 40 kpc: squashed, broken but smooth. *Mon. Not. R. Astron. Soc.* **416**, 2903–2915 (2011).

40. 40.

Lindgren, L. et al. Gaia Early Data Release 3. Parallax bias versus magnitude, colour, and position. *Astron. Astrophys.* **649**, A4 (2021).

41. 41.

Ibata, R. A., Malhan, K., Martin, N. F. & Starkenburg, E. Phlegethon, a nearby 75-degree-long retrograde Stellar Stream. *Astrophys. J.* **865**,

85 (2018).

42. 42.

Chene, A.-N. et al. in *Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation* Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9151 (eds. Navarro, R. et al.) 915147 (SPIE, 2014).

43. 43.

Pazder, J., Fournier, P., Pawluczyk, R. & van Kooten, M. in *Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation* Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9151 (eds. Navarro, R. et al.) 915124 (SPIE, 2014).

44. 44.

Martioli, E. et al. in *Software and Cyberinfrastructure for Astronomy II* Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8451 (eds. Radziwill, N. M. & Chiozzi, G.) 84512B (SPIE, 2012).

45. 45.

Schlegel, D. J., Finkbeiner, D. P. & Davis, M. Maps of dust infrared emission for use in estimation of reddening and cosmic microwave background radiation foregrounds. *Astrophys. J.* **500**, 525 (1998).

46. 46.

Schlafly, E. F. & Finkbeiner, D. P. Measuring reddening with Sloan Digital Sky Survey stellar spectra and recalibrating SFD. *Astrophys. J.* **737**, 103 (2011).

47. 47.

González Hernández, J. I. & Bonifacio, P. A new implementation of the infrared flux method using the 2MASS catalogue. *Astron. Astrophys.* **497**, 497–509 (2009).

48. 48.

Mucciarelli, A., Bellazzini, M. & Massari, D. Exploiting the Gaia EDR3 photometry to derive stellar temperatures. *Astron. Astrophys.* (in the press).

49. 49.

Mashonkina, L., Jablonka, P., Pakhomov, Y., Sitnova, T. & North, P. The formation of the Milky Way halo and its dwarf satellites; a NLTE-1D abundance analysis. I. Homogeneous set of atmospheric parameters. *Astron. Astrophys.* **604**, A129 (2017).

50. 50.

Karovicova, I. et al. Fundamental stellar parameters of benchmark stars from CHARA interferometry. I. Metal-poor stars. *Astron. Astrophys.* **640**, A25 (2020).

51. 51.

Giribaldi, R. E., da Silva, A. R., Smiljanic, R. & Cornejo Espinoza, D. TITANS metal-poor reference stars. I. Accurate effective temperatures and surface gravities for dwarfs and subgiants from 3D non-LTE H α profiles and Gaia parallaxes. *Astron. Astrophys.* **650**, A194 (2021).

52. 52.

Kurucz, R. L. ATLAS12, SYNTHE, ATLAS9, WIDTH9, et cetera. *Mem. Soc. Astron. Ital. Suppl.* **8**, 14 (2005).

53. 53.

Sneden, C. A. *Carbon and Nitrogen Abundances in Metal-Poor Stars*. PhD thesis, Univ. of Texas at Austin (1973).

54. 54.

Sobeck, J. S. et al. The abundances of neutron-capture species in the very metal-poor globular cluster M15: a uniform analysis of red giant branch and red horizontal branch stars. *Astron. J.* **141**, 175 (2011).

55. 55.

Placco, V. M. et al. Linemake: an atomic and molecular line list generator. *Res. Notes AAS* **5**, 92 (2021).

56. 56.

<http://inspect-stars.net>.

57. 57.

<http://nlte.mpiag.de>.

58. 58.

Lind, K., Bergemann, M. & Asplund, M. Non-LTE line formation of Fe in late-type stars - II. 1D spectroscopic stellar parameters. *Mon. Not. R. Astron. Soc.* **427**, 50–60 (2012).

59. 59.

Bergemann, M. & Cescutti, G. Chromium: NLTE abundances in metal-poor stars and nucleosynthesis in the Galaxy (2010).

60. 60.

Bergemann, M., Lind, K., Collet, R., Magic, Z. & Asplund, M. Non-LTE line formation of Fe in late-type stars - I. Standard stars with 1D and <3D> model atmospheres. *Mon. Not. R. Astron. Soc.* **427**, 27–49 (2012).

61. 61.

Mashonkina, L., Korn, A. J. & Przybilla, N. A non-LTE study of neutral and singly-ionized calcium in late-type stars. *Astron. Astrophys.* **461**, 261–275 (2007).

62. 62.

Asplund, M., Grevesse, N., Sauval, A. J. & Scott, P. The chemical composition of the Sun. *Ann. Rev. Astron. Astrophys.* **47**, 481–522 (2009).

63. 63.

Tody, D. in *Astronomical Data Analysis Software and Systems II* Astronomical Society of the Pacific Conference Series, Vol. 52 (eds. Hanisch, R. J. et al.) 173 (1993).

64. 64.

Aguado, D. S., Allende Prieto, C., González Hernández, J. I., Rebolo, R. & Caffau, E. New ultra metal-poor stars from SDSS: follow-up GTC medium-resolution spectroscopy. *Astron. Astrophys.* **604**, A9 (2017).

65. 65.

Aguado, D. S., González Hernández, J. I., Allende Prieto, C. & Rebolo, R. J0815+4729: a chemically primitive dwarf star in the galactic halo observed with Gran Telescopio Canarias. *Astrophys. J. Lett.* **852**, L20 (2018).

66. 66.

Allende Prieto, C. et al. Deep SDSS optical spectroscopy of distant halo stars. I. Atmospheric parameters and stellar metallicity distribution. *Astron. Astrophys.* **568**, A7 (2014).

67. 67.

Aguado, D. S., González Hernández, J. I., Allende Prieto, C. & Rebolo, R. WHT follow-up observations of extremely metal-poor stars identified from SDSS and LAMOST. *Astron. Astrophys.* **605**, A40 (2017).

68. 68.

Koesterke, L., Allende Prieto, C. & Lambert, D. L. Center-to-limb variation of solar three-dimensional hydrodynamical simulations. *Astrophys. J.* **680**, 764–773 (2008).

69. 69.

Boender, C. G. E., Rinnoy Kan, A. H. G., Timmer, G. T. & Stougie, L. A stochastic method for global optimization. *Math. Program.* **22**, 125 (1982).

70. 70.

Wenger, M. et al. The SIMBAD astronomical database. The CDS reference database for astronomical objects. *Astron. Astrophys. Suppl. Ser.* **143**, 9–22 (2000).

71. 71.

Ochsenbein, F., Bauer, P. & Marcout, J. The VizieR database of astronomical catalogues. *Astron. Astrophys. Suppl. Ser.* **143**, 23–32 (2000).

72. 72.

<http://www.inasan.rssi.ru/~lima/pristine>.

Acknowledgements

N.F.M., R.A.I., A.A. and Z.Y. gratefully acknowledge support from the French National Research Agency (ANR)-funded project Pristine (ANR-18-CE31-0017) along with funding from Centre National de la Recherche

Scientifique (CNRS)–INSU through the Programme National de Cosmologie et Galaxies and through CNRS grant PICS07708 and from the European Research Council under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 834148). K.A.V. is grateful for funding through the National Science and Engineering Research Council of Canada (NSERC) Discovery Grants and CREATE programmes. E.S. acknowledges funding through VIDI grant ‘Pushing Galactic Archaeology to its Limits’ (project number VI.Vidi.193.093), which is funded by the Dutch Research Council (NWO). J.I.G.H. acknowledges financial support from Spanish Ministry of Science and Innovation (MICINN) project AYA2017-86389-P, and also from the Spanish MICINN under the 2013 Ramón y Cajal programme (RYC-2013-14875). G.F.T. acknowledges support from the Agencia Estatal de Investigación of the Spanish MCINN (grant number FJC2018-037323-I). We thank the Canada–France–Hawaii Telescope (CFHT) staff for performing the Pristine observations in queue mode, for their reactivity in adapting the schedule and for answering questions during the data-reduction process. We are also grateful to the High Performance Computing Centre of the University of Strasbourg and its staff for very generous time allocation and for their support during the development of the STREAMFINDER project. This research used the SIMBAD database⁷⁰, managed and run at CDS, Strasbourg, France. This research used the VizieR catalogue access tool⁷¹, CDS, Strasbourg, France. This work is based on observations obtained with MegaPrime–MegaCam, a joint project of the CFHT and CEA–DAPNIA, at the CFHT, which is operated by the National Research Council of Canada, the Institut National des Science de l’Univers of the French CNRS and the University of Hawaii. ESPaDOnS is a collaborative project funded by France (CNRS, MENESR, OMP, and LATT), Canada (NSERC), the CFHT and the European Space Agency. Data were reduced with use of the CFHT-developed OPERA data-reduction pipeline. We recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the Indigenous Hawaiian community. We are very fortunate to have had the opportunity to conduct observations from this mountain. This work is based on observations obtained with Gemini Remote Access to CFHT ESPaDOnS Spectrograph (GRACES), as part of the Gemini Large and Long Program, GN-2020B-LP-102. The international Gemini Observatory, a program of

NSF's NOIRLab, is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation, on behalf of the Gemini Observatory partnership: the National Science Foundation (United States), National Research Council (Canada), Agencia Nacional de Investigación y Desarrollo (Chile), Ministerio de Ciencia, Tecnología e Innovación (Argentina), Ministério da Ciência, Tecnologia, Inovações e Comunicações (Brazil) and Korea Astronomy and Space Science Institute (Republic of Korea). This work is based on observations made with the GTC telescope, at the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias, under Director's Discretionary Time. This work was partly based on data obtained with the instrument OSIRIS, built by a consortium led by the Instituto de Astrofísica de Canarias in collaboration with the Instituto de Astronomía of the Universidad Autónoma de México. OSIRIS was funded by GRANTECAN and the National Plan of Astronomy and Astrophysics of the Spanish Government. This work used data from the European Space Agency 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 the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement. Guo Shoujing Telescope (LAMOST) is a national major scientific project of the Chinese Academy of Sciences. Funding for the project was provided by the National Development and Reform Commission. LAMOST is operated and managed by the National Astronomical Observatories, Chinese Academy of Sciences.

Author information

Affiliations

1. Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, Strasbourg, France

Nicolas F. Martin, Rodrigo A. Ibata, Anke Arentsen & Zhen Yuan

2. Max-Planck-Institut für Astronomie, Heidelberg, Germany

Nicolas F. Martin & Morgan Fouesneau

3. Department of Physics and Astronomy, University of Victoria,
Victoria, British Columbia, Canada

Kim A. Venn, Federico Sestito, Sébastien Fabbro & Julio F. Navarro

4. Institute of Astronomy, University of Cambridge, Cambridge, UK

David S. Aguado

5. Instituto de Astrofísica de Canarias, La Laguna, Tenerife, Spain

David S. Aguado, Jonay I. González Hernández, Carlos Allende
Prieto & Guillaume F. Thomas

6. Dipartimento di Fisica e Astronomia, Università degli Studi di Firenze,
Sesto Fiorentino, Italy

David S. Aguado

7. Kapteyn Astronomical Institute, University of Groningen, Groningen,
the Netherlands

Else Starkenburg

8. Departamento de Astrofísica, Universidad de La Laguna, La Laguna,
Tenerife, Spain

Jonay I. González Hernández, Carlos Allende Prieto & Guillaume F.
Thomas

9. GEPI, Observatoire de Paris, Université PSL, CNRS, Meudon, France

Piercarlo Bonifacio, Elisabetta Caffau & Pascale Jablonka

10. Department of Astronomy & Astrophysics, University of Toronto,
Toronto, Ontario, Canada

Raymond G. Carlberg

11. NRC Herzberg Astronomy & Astrophysics, Victoria, British Columbia, Canada

Sébastien Fabbro & Alan W. McConnachie

12. Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Nice, France

Vanessa Hill & Georges Kordopatis

13. Laboratoire d'astrophysique, École Polytechnique Fédérale de Lausanne, Observatoire, Versoix, Switzerland

Pascale Jablonka

14. Dipartimento di Fisica & Astronomia “Augusto Righi”, Alma Mater Studiorum, Università di Bologna, Bologna, Italy

Carmela Lardo & Alessio Mucciarelli

15. The Oskar Klein Centre, Department of Physics, Stockholm University, Stockholm, Sweden

Khyati Malhan

16. Institute of Astronomy, Russian Academy of Sciences, Moscow, Russia

Lyudmila I. Mashonkina

17. UK Astronomy Technology Centre, Royal Observatory, Edinburgh, UK

Rubén Sánchez-Janssen

18. INAF – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Bologna, Italy

Alessio Mucciarelli

Contributions

N.F.M. is a co-leader of the Pristine survey, led the discovery of the C-19 stream, coordinated the spectroscopic observations and co-led the writing of the manuscript. K.A.V. led the GRACES spectroscopic follow-up and the analysis of the resulting spectra, and co-led the writing of the manuscript. D.S.A. led the analysis of the OSIRIS spectra and the writing of the corresponding section of the manuscript. E.S. is a co-leader of the Pristine survey and derived the Pristine photometric metallicities. J.I.G.H. coordinated the OSIRIS follow-up, performed the radial velocity analysis of these spectra and was greatly involved in writing this part of the manuscript. R.A.I. led the STREAMFINDER analysis and derived the orbit of the stream. P.B., E.C. and F.S. derived stellar and orbital parameters for the stars with spectroscopic follow-up. All other authors helped in the development of the Pristine survey and all authors assisted in the development and writing of the paper. A.M. derived the relations used to infer the stellar parameters of the spectroscopically observed stars.

Corresponding author

Correspondence to [Nicolas F. Martin](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Pristine photometric information for all stars of the C-19 stream selected by STREAMFINDER and present in the Pristine survey.

Large symbols represent stars with $G_0 < 19.0$, for which the STREAMFINDER selection is very reliable, and small symbols represent stars fainter than this limit, which are more likely contaminated in the STREAMFINDER catalogue (see also Fig. 1). The lines represent model expectations as determined from the spectral libraries and filter response curves, without an assumption on whether the star is a dwarf or a giant¹⁰. Both model lines and data points are color-coded by their $([\{\mathrm{Fe}\}]/\{\mathrm{H}\})$ metallicities. Most C-19 candidate members are located in the region that corresponds to metallicities below $([\{\mathrm{Fe}\}]/\{\mathrm{H}\}) = -3.0$ (above the blue line). For the large data points, we specifically used a photometric metallicity model tailored to giant stars. Near the tip of the red giant branch, it deviates significantly from the generic model represented by the coloured lines and explains the higher metallicity of the right-most point compared with the models.

Extended Data Fig. 2 Favourite orbital solutions for the C-19 stream.

The dashed line shows the orbit of C-19 constrained using the proper motions (red symbols in **a**, **b**) of C-19 members identified by STREAMFINDER, their Gaia parallaxes (**c**), their radial velocities when available (**d**) and their location on the sky compared with the distribution of extinction⁴⁵ (**e**). The orbit determined without using the Gaia parallax information but instead anchoring the distance at 18 kpc is represented by the dotted line. **f**, **g** The two orbits, integrated for ± 1 Gyr, projected on the

Galactic plane, and in the R – z plane. The thick red lines correspond to the part of the orbits that overlaps the observed C-19 member stars.

[Extended Data Fig. 3 Spectra of the C-19 member stars observed with OSIRIS, normalized using a running mean filter after removing the velocity signal in the rest frame \(black lines\), together with the best fit \(blue lines\) derived by adopting a fitting procedure.](#)

The metallicity, [Fe/H], computed from [M/H] and [Ca/H] is also indicated for each target (see the text for more details).

Extended Data Table 1 List of potential C-19 members from the STREAMFINDER sample

Extended Data Table 2 Summary of observations for the C-19 candidate stars

Extended Data Table 3 Spectroscopic parameters and one-dimensional LTE chemical abundances for the Gemini–GRACES spectra

Extended Data Table 4 Spectral lines and atomic data used for the chemical abundances for the Gemini–GRACES spectra

Extended Data Table 5 Stellar parameters and abundances of C-19 stars observed with GTC–OSIRIS

Supplementary information

[Peer Review File](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

[Cite this article](#)

Martin, N.F., Venn, K.A., Aguado, D.S. *et al.* A stellar stream remnant of a globular cluster below the metallicity floor. *Nature* **601**, 45–48 (2022). <https://doi.org/10.1038/s41586-021-04162-2>

- Received: 14 June 2021
- Accepted: 19 October 2021
- Published: 05 January 2022
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04162-2>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04162-2>

- Article
- Open Access
- [Published: 05 January 2022](#)

An early transition to magnetic supercriticality in star formation

- [T.-C. Ching](#) ORCID: orcid.org/0000-0001-8516-2532¹,
- [D. Li](#) ORCID: orcid.org/0000-0003-3010-7661^{1,2,3},
- [C. Heiles](#)⁴,
- [Z.-Y. Li](#)⁵,
- [L. Qian](#) ORCID: orcid.org/0000-0003-0597-0957¹,
- [Y. L. Yue](#) ORCID: orcid.org/0000-0003-4415-2148¹,
- [J. Tang](#)¹ &
- [S. H. Jiao](#)^{1,2}

Nature volume 601, pages 49–52 (2022)

- 2418 Accesses
- 155 Altmetric
- [Metrics details](#)

Subjects

- [Astrophysical magnetic fields](#)
- [Interstellar medium](#)

Abstract

Magnetic fields have an important role in the evolution of interstellar medium and star formation^{1,2}. As the only direct probe of interstellar field strength, credible Zeeman measurements remain sparse owing to the lack of suitable Zeeman probes, particularly for cold, molecular gas³. Here we report the detection of a magnetic field of $+3.8 \pm 0.3$ microgauss through the H I narrow self-absorption (HINSA)^{4,5} towards L1544^{6,7}—a well-studied prototypical prestellar core in an early transition between

starless and protostellar phases^{8,9,10} characterized by a high central number density¹¹ and a low central temperature¹². A combined analysis of the Zeeman measurements of quasar H I absorption, H I emission, OH emission and HINSA reveals a coherent magnetic field from the atomic cold neutral medium (CNM) to the molecular envelope. The molecular envelope traced by the HINSA is found to be magnetically supercritical, with a field strength comparable to that of the surrounding diffuse, magnetically subcritical CNM despite a large increase in density. The reduction of the magnetic flux relative to the mass, which is necessary for star formation, thus seems to have already happened during the transition from the diffuse CNM to the molecular gas traced by the HINSA. This is earlier than envisioned in the classical picture where magnetically supercritical cores capable of collapsing into stars form out of magnetically subcritical envelopes^{13,14}.

[Download PDF](#)

Main

In non-masing interstellar medium, only H I, OH and CN have successfully produced systematic Zeeman measurements. The comprehensive Zeeman surveys¹⁵ indicate that the magnetic fields in the diffuse cold neutral medium (CNM) probed by H I do not scale significantly with density, whereas above a critical break-point density of approximatey 300 cm^{-3} , the magnetic fields in dense cores probed by OH tend to increase with density. However, owing to the gap in densities between the H I (about 40 cm^{-3}) and OH ($\sim 10^3 \text{ cm}^{-3}$) Zeeman measurements, the field transition around the critical density of about 300 cm^{-3} (where the dependence of the field strength on the density changes behaviour) remains a controversial topic and could have crucial implications for star formation^{16,17,18,19}. Recently, a CCS Zeeman detection²⁰ shed light on regions denser than that probed by OH. A Zeeman probe that is sensitive for a wide range of densities, particularly the low-density molecular envelope, is highly desirable and could help to distinguish different core-formation scenarios.

We developed the so-called H I narrow self-absorption (HINSA) technique to provide a probe of the transition from H I to H₂ (refs. ^{4,5}). HINSA traces cold atomic hydrogen well mixed with H₂, which provides the necessary cooling, not available in the CNM, of H I through collision. Close to the steady state between H₂ formation and destruction, the HINSA strength is independent of the gas density⁵ and thus capable of probing the transition around the critical density. Although the Zeeman effect of the H I self-absorption feature has been reported^{21,22}, the broad line widths of the absorption components are mostly associated with diffuse atomic gas rather than dense

molecular gas. Considering that HINSA typically has much higher brightness temperatures than most molecular lines, is impervious to depletion²³ and can be detected in a wide range of H₂ densities, HINSA is a promising Zeeman probe for molecular gas.

The HINSA feature in L1544 has a strong absorption dip and a nearly thermalized narrow line width at a temperature lower than 15 K (ref. ⁴). The non-thermal line width and centroid velocity of the HINSA are very close to those of the emission lines of OH, ¹³CO and C¹⁸O molecules, and their column densities are well correlated, suggesting that a significant fraction of the atomic hydrogen is located in the cold, well-shielded portions of L1544⁵. We thus assume that the column density sampled by the HINSA can be approximated by that obtained from dust, despite the substantially larger apparent area covered by HINSA (Fig. ^{1a}). The previous OH Zeeman detection with Arecibo²⁴ towards the L1544 centre resulted in a field strength of $B_{\text{los}} = +10.8 \pm 1.7 \mu\text{G}$, where B_{los} is the magnetic field component along the line of sight, with a positive sign representing the field pointing away from the observer. In contrast, the OH Zeeman observations of the Green Bank Telescope (GBT) towards four envelope locations 6.0' (0.24 pc) from the centre yielded a marginal detection of $B_{\text{los}} = +2 \pm 3 \mu\text{G}$ (ref. ¹⁶), leaving the structure of the envelope field undetermined.

Fig. 1: L1544 core and illustration of the structure of interstellar medium from the CNM to the core.

 [figure1](#)

a, A composite of Digitized Sky Survey 2 (DSS2) images of L1544 with the i band in red, the r band in green and the b band in blue overlaid with the HINSA and H₂ column density maps. The white dashed contours are 30%, 50%, 70% and 90% of the peak HINSA column density, and the orange contours are $2 \times 10^{21} \text{ cm}^{-2}$, $4 \times 10^{21} \text{ cm}^{-2}$, $6 \times 10^{21} \text{ cm}^{-2}$, $8 \times 10^{21} \text{ cm}^{-2}$ and $10 \times 10^{21} \text{ cm}^{-2}$ for the H₂ column density. The red, green and cyan circles mark the locations and beam sizes of the FAST, Arecibo and GBT Zeeman observations, respectively. **b**, A composite of three 0.5 km s⁻¹ velocity slices of the Arecibo GALFA-H I images at 6.2 km s⁻¹, 6.7 km s⁻¹

and 7.3 km s^{-1} local standard of rest (LSR) velocities in blue, green and red. The dashed rectangle shows the region of **a**. The two absorption dots represent the locations of quasars 3C132 and 3C133. **c**, Schematic view of the CNMs, the molecular envelope and the L1544 core.

With the Five-hundred-meter Aperture Spherical radio Telescope (FAST)^{[25](#)}, we detected Zeeman splittings in a $2.9'$ beam (0.12 pc) towards the HINSA column density peak, $3.6'$ (0.15 pc) away from the L1544 centre (Fig. [1](#)). The spectra of the Stokes $I(v)$ and $V(v)$ parameters (where v denotes velocity) are shown in Fig. [2](#). The $I(v)$ spectrum contains the H I emission of the CNM and the warm neutral medium (WNM) clouds in the direction towards the Taurus complex and a HINSA feature at the centroid velocity of L1544. Figure [2a](#) shows our decomposition of $I(v)$ into a foreground HINSA component, a background WNM component, and three CNM components between the HINSA and WNM. Our fitted parameters of the HINSA component are in good agreement with the previous HINSA observations^{[4,5](#)}, and our parameters of the CNM and WNM components are similar to the Arecibo results towards quasars around L1544^{[26](#)}.

Fig. 2: The Stokes $I(v)$ and $V(v)$ spectra at 21-cm wavelength towards the HINSA column density peak.

 **figure2**

a, The black profile represents the $I(v)$ spectrum. The red profile represents the absorption from the foreground HINSA component. The blue dashed and dotted profiles represent the emission of the CNM and WNM components, respectively. The CNM and WNM profiles include the absorption from the CNM components that lie in front but do not include the absorption from the HINSA. The black dashed profile

represents the sum of the absorption and emission profiles. **b**, The black profile represents the $V(v)$ spectrum. The black dashed profile represents the sum of the Zeeman splitting profiles of the five components. The red profile represents the Zeeman splitting profile with $B_{\text{los}} = +3.8 \mu\text{G}$ of the HINSA component.

The $V(v)$ spectrum shows features of classic ‘S curve’ patterns proportional to the first derivatives of $I(v)$ for the HINSA, CNM and WNM components, as expected for Zeeman splittings. The Zeeman splitting profile of HINSA has a maximum at high velocity and a minimum at low velocity, opposite to the Zeeman splitting profile of CNM1, the closest CNM component at a velocity similar to L1544, which shows positive V at low velocity and negative V at high velocity. From our least-squares fits to $V(v)$, Fig. 2b shows the Zeeman splitting of the HINSA and the total Zeeman profile of the five components, and Fig. 3 shows the individual Zeeman splittings and B_{los} of the components. The HINSA Zeeman effect gives $B_{\text{los}} = +3.8 \pm 0.3 \mu\text{G}$, and the H I Zeeman effect of CNM1 gives $B_{\text{los}} = +4.0 \pm 1.1 \mu\text{G}$. The magnetic field strengths of HINSA and CNM1 are consistent with the results of $B_{\text{los}} = +5.8 \pm 1.1 \mu\text{G}$ and $B_{\text{los}} = +4.2 \pm 1.0 \mu\text{G}$ obtained from the Zeeman observations towards quasars 3C133 and 3C132, probing the magnetic fields of CNM1 at distances of $17.7'$ (0.72 pc) and $174.5'$ (7.1 pc) from L1544, respectively²⁷. For the second and third CNM components (CNM2 and CNM3) along the line of sight, our results of $B_{\text{los,CNM2}} = -7.6 \pm 1.0 \mu\text{G}$ and $B_{\text{los,CNM3}} = +2.9 \pm 0.4 \mu\text{G}$ are also consistent with the results of $B_{\text{los,CNM2}} = -9.6 \pm 6.3 \mu\text{G}$ and $B_{\text{los,CNM3}} = -0.3 \pm 1.7 \mu\text{G}$ towards quasar 3C133²⁷.

Fig. 3: The individual $V(v)$ profiles for the HINSA, CNM and WNM components.

 **figure3**

In each panel, the red profile represents the fitted Zeeman profile of the component and the black profile represents the observed $V(v)$ subtracted by the fitted Zeeman profiles of the other four components. The CNM and WNM Zeeman profiles include the absorption from the CNM components that lie in front but do not include the

absorption from the HINSA. The sum of the red profiles of the five components is the black dashed profile in Fig. 2b.

Comparing the Zeeman observations of HINSA, OH and H I tracing the CNM1 and the molecular envelope of L1544, it is clear that the magnetic fields at distances of 0.15 pc, 0.24 pc, 0.72 pc and 7.1 pc from the centre all have the same direction of B_{los} and consistent strengths roughly within the 1σ . This finding is in agreement with the conclusion of a median value of 6 μG in absolute total strength in H I clouds inferred from comprehensive Zeeman surveys¹⁵. The HINSA Zeeman effect thus provides a connection between the magnetic fields from H I clouds to molecular clouds. The H I emission components (CNM1, CNM2 and CNM3) and the H I absorption towards 3C133 and 3C132 trace the CNM with a kinematic temperature of about 100 K (ref. 27) and a number density of about 40 cm^{-3} (ref. 15), whereas the HINSA and OH observations trace the envelope of about 10–15 K (ref. 4) and about 10^3 cm^{-3} (refs. 5,15). Despite the one to two orders of magnitude change in both temperature and density in the phase transition from the atomic CNM to the molecular envelope, the Zeeman observations reveal a magnetic field that is coherent in both direction and strength across multi-scales and multi-phases of the interstellar medium. To constrain the uniformity of the coherent magnetic field, our likelihood analysis of the HINSA, OH and H I Zeeman measurements suggests a Gaussian distribution of the B_{los} with a mean strength of $B_0 = +4.1 \pm 1.6 \mu\text{G}$ and an intrinsic spread of $\langle(\sigma)_0\rangle = \{1.2\}_{-0.6}^{+1.2} \langle(\rm{\mu})\rangle \langle(\rm{G})\rangle$, a significantly better constraint than the previous estimation of $\langle(B)_0\rangle = +4_{-8}^{+10} \langle(\rm{\mu})\rangle \langle(\rm{G})\rangle$ based on only the OH results¹⁸.

It is well known that the progenitor of molecular gas, the atomic CNM, is strongly magnetized, as measured by the dimensionless mass-to-magnetic-flux ratio λ in units of the critical value of $2\pi G^{1/2}$ ($\lambda = 7.6 \times 10^{-21} [N_{\text{H}_2} (\text{cm}^{-2})] [B_{\text{tot}} (\mu\text{G})]^{-1}$, where N_{H_2} is the column density of H_2 gas and B_{tot} is the total magnetic field strength¹), which is well below unity (that is, magnetically subcritical)²⁸. However, the immediate progenitors of stars, the prestellar cores of molecular clouds such as the L1554 core, are observed to be magnetically supercritical ($\lambda > 1$)²⁴, which is required for the self-gravity to overwhelm the magnetic support and form stars through gravitational collapse. When and how the transition from the magnetically subcritical CNM that is incapable of forming stars through direct gravitational collapse to the supercritical star-forming cores occurs is a central unresolved question in star formation.

Our HINSA Zeeman observations can be used to address this question. Using the physical parameters of the clouds (Table 1) and the statistically most probable value of $B_{\text{tot}} = 2B_{\text{los}}$, the λ of CNM1 is about 0.10–0.18, consistent with previous results²⁸. The λ of the envelope and core of L1544 core is 2.5–3.5, which is well above unity,

indicating that the transition to magnetic supercriticality has already occurred. We further consider the relative values of λ between CNM1 and L1544 to avoid the geometrical correction from B_{los} to B_{tot} (ref. 16), assuming that the inclination angles of the magnetic fields in the L1544 core and envelope are similar. Therefore, the molecular envelope of the L1544 core traced by HINSA is at least 13 times less magnetized relative to its mass compared with its ambient CNM. This is different from the ‘classic’ theory of low-mass star formation, which envisions the transition from magnetic subcriticality to supercriticality occurring as the supercritical core forms out of the magnetically supported (subcritical) envelope^{13,14}. Our results suggest that the transition from magnetic subcriticality to supercriticality occurs earlier, during the formation of the molecular envelope, favouring the more rapidly evolving scenario of core formation and evolution for L1544⁸ over the slower, magnetically retarded scenario⁹. In other words, by the time that the molecular envelope is formed, the problem of excessive magnetic flux as a fundamental obstacle to gravitational collapse and star formation is already resolved. This early reduction of flux relative to mass is unlikely owing to the ‘classical’ scenario where gravity drives neutrals through ions (and the magnetic field tied to them) in a process called ‘ambipolar diffusion’ because the CNM is not self-gravitating. The coherent magnetic fields reviewed here provide a new specific question on how to create supercritical dense cores such as L1544 from subcritical clouds. Plausible scenarios include mass accumulation along field lines²⁹ and (turbulence enhanced) magnetic reconnection³⁰, although whether such scenarios can reproduce the distributions of gas and magnetic field observed in the L1544 region remains to be seen. In any case, the already magnetically supercritical envelope can in principle go on to form dense cores and stars without having to further reduce its magnetic flux relative to the mass.

Table 1 Physical parameters of the clouds

Methods

Data reduction

The FAST Zeeman observations towards the HINSA column density peak in L1544 were carried out on five days between August and November 2019 with a total integration time of 7.6 h. The HINSA spectra were obtained with the central beam of the L-band 19-beam receiver³¹. The central beam has an average system temperature of 24 K, a main beam efficiency of 0.63 and a main beam diameter at the half-power point of 2.9' with a pointing accuracy of 7.9''. The 19-beam receiver had orthogonal linear polarization feeds followed by a temperature-stabilized noise injection system and low noise amplifiers to produce the X and Y signals of the two polarization paths. The XX, YY, XY and YX correlations of the signals then were simultaneously

recorded using the ROACH backend with 65,536 spectral channels in each polarization. The spectral bandwidth was 32.75 MHz centred at the frequency of the H I 21-cm line for a channel spacing of 500 Hz, and the $V(v)$ spectrum presented in this work was Hanning smoothed, which produced a spectral resolution of 0.21 km s^{-1} .

The data reduction, including gain and phase calibrations of the two polarization paths, bandpass calibrations of the four correlated spectra and polarization calibrations to generate the Stokes I , Q , U and V spectra, was carried out using the IDL RHSTK package written by C. Heiles and T. Robishaw, which is widely used for Arecibo and GBT polarization data. The 19-beam receiver is rotatable from -80° to $+80^\circ$ with respect to the line of equatorial latitude. The polarization calibrations used drifting scans of the continuum source 3C286 at rotation angles of -60° , -30° , 0° , 30° and 60° over 1.5 h surrounding its transit. The details of the polarization calibration procedure are provided in ref. ³². We performed polarization calibrations once per month during the observations. The calibrated polarization of 3C286 of the three epochs were $8.9\% \pm 0.1\%$, $8.7\% \pm 0.2\%$ and $9.0\% \pm 0.1\%$ for polarization degrees and $30.4^\circ \pm 0.3^\circ$, $33.8^\circ \pm 0.5^\circ$ and $29.4^\circ \pm 0.3^\circ$ for polarization angles. Considering that the ionosphere can generate a Faraday rotation of 1° – 3° in polarization angle at the L band³³, our results were consistent with the intrinsic polarization degree of 9.5% and polarization angle of 33° of 3C286 at 1,450 MHz (ref. ³⁴). In addition to the polarization observations of L1544 and 3C286, we observed the circularly polarized OH maser source IRAS02524+2046³⁵ to verify that our procedures produced consistent B_{los} , including the sign or direction of the magnetic field, as had been obtained previously.

The convolutions of the sidelobes of the Stokes V beam with the spatial gradient of the Stokes I emission may generate a false ‘S curve’ in the V spectrum²⁷. To check the credibility of our Zeeman detections, we measured the Stokes V beam of FAST and convolved the beam with the Galactic Arecibo L-band Feed Array (GALFA) Stokes I cube³⁶ of L1544. The convolved V spectrum showed a profile with a shape similar to the I spectrum and a strength less than 0.03% of the I spectrum, different from the ‘S curve’ patterns in the observed V spectrum. Meanwhile, the 19-beam receiver was rotated to -45° , 0° and 45° in the three epochs of the L1544 observations, and all of the three epochs showed ‘S curve’ patterns in the V spectra, indicating that our Zeeman results were true detections.

Although the data of the 19 beams of the FAST L-band receiver were simultaneously taken in our observations, only the polarization of the central beam was commissioned at the time of writing. The results represented in this work were made with only the central beam pointing towards the HINSA column density peak in Fig. 1. The Zeeman results of the 18 off-central beams will be published in the future.

Multiple Gaussians and radiative transfer fitting to $I(v)$ and $V(v)$

We adopt the least-squares fits of multiple Gaussians with radiative transfer²⁶ to decompose the $I(v)$ into the HINSA, CNM and WNM components. The expected profile of $I(v)$ consists of multiple CNM components providing opacity and also brightness temperature and a WNM component providing only brightness temperature:

$$\text{\$\$ } I(v) = \{I\}_{\text{CNM}}(v) + \{I\}_{\text{WNM}}(v). \text{\$\$} \quad (1)$$

The $I_{\text{CNM}}(v)$ is an assembly of N CNM components

$$\text{\$\$ } \{I\}_{\text{CNM}}(v) = \sum_{n=1}^N \{I\}_{\text{peak},n}(v) \cdot \tau_n(v) \cdot \tau_0(v) \text{\$\$} \quad (2)$$

where the subscript m with its associated optical depth profile $\tau_m(v)$ represents each of the M CNM clouds that lie in front of cloud n . The optical depth of the i th component is

$$\text{\$\$ } \tau_i(v) = \tau_0(v) \cdot \exp\left(-\frac{(v - v_{0,i})}{\sigma_{v,i}}\right)^2 \text{\$\$} \quad (3)$$

in which τ_0 represents the HINSA providing only opacity and no brightness temperature. For the WNM in the background

$$\text{\$\$ } \{I\}_{\text{WNM}}(v) = \{I\}_{\text{peak}} \cdot \{I\}_{\text{WNM}} \cdot \{I\}_{\text{e}} \cdot \left(\sum_{i=0}^N \tau_i(v)\right)^2 \text{\$\$} \quad (4)$$

The fitting of $I(v)$ thus yields values for the intrinsic peak Stokes I emission (I_{peak}), τ_0 and the Gaussian dispersion (σ_v) of the components.

We consider the radiative transfer of $V(v)$ in terms of right circular polarization (RCP) and left circular polarization (LCP). The Zeeman effect states that with the existence of B_{los} , the frequency of RCP shifts from its original frequency v_0 to $v_0 + v_z$ and the frequency of LCP shifts to $v_0 - v_z$ with $v_z = (Z/2) \times B_{\text{los}}$, where Z is the Zeeman splitting factor (2.8 Hz μG^{-1} for the H I 21-cm line). As the RCP and LCP are

orthogonal components of radiation, the radiative transfer processes of RCP and LCP are independent of each other. For RCP, equation (1) becomes

$$\text{Equation 5: } \text{\$}\{T\}_{\{\{\text{rm}\{RCP\}\}\}}=\{T\}_{\{\{\text{rm}\{RCP\}\},\{\text{rm}\{CNM\}\}\}}(v,\{\tau\}_{\{\{\text{rm}\{RCP\}\},i\}})+\{T\}_{\{\{\text{rm}\{RCP\}\},\{\text{rm}\{WNM\}\}\}}(v,\{\tau\}_{\{\{\text{rm}\{RCP\}\},i\}}), \$\$$$

(5)

where for the i th component, $\langle\{T\}_{\{\{\text{rm}\{RCP\}\}\}},i\rangle=\langle I\}_{\{\{\text{rm}\{peak\}\},i\}}/2\rangle$, $\tau_{RCP,i}$ is the optical depth in the RCP radiation to substitute the τ_i in equation (3) with $\langle\{\tau\}_{\{\{\text{rm}\{RCP\}\},i\}}=\langle\tau\}_{\{i\}}(\langle\nu\}_{\{0\}}+\langle\nu\}_{\{\{\text{rm}\{z\}\},i\}})\rangle$ for $B_{\text{los},i}$ of the component, and the parameters of v_0 and σ_v remain the same. Similarly, for LCP

$$\text{Equation 6: } \text{\$}\{T\}_{\{\{\text{rm}\{LCP\}\}\}}=\{T\}_{\{\{\text{rm}\{LCP\}\},\{\text{rm}\{CNM\}\}\}}(v,\{\tau\}_{\{\{\text{rm}\{LCP\}\},i\}})+\{T\}_{\{\{\text{rm}\{LCP\}\},\{\text{rm}\{WNM\}\}\}}(v,\{\tau\}_{\{\{\text{rm}\{LCP\}\},i\}}), \$\$$$

(6)

with $\langle\{T\}_{\{\{\text{rm}\{LCP\}\}\}},i\rangle=\langle I\}_{\{\{\text{rm}\{peak\}\},i\}}/2\rangle$ and $\langle\{\tau\}_{\{\{\text{rm}\{LCP\}\}\}},i\rangle=\langle\tau\}_{\{i\}}(\langle\nu\}_{\{0\}}-\langle\nu\}_{\{\{\text{rm}\{z\}\},i\}})\rangle$. The fitting of $\langle V(v)=\{T\}_{\{\{\text{rm}\{RCP\}\}\}}-\langle\{T\}_{\{\{\text{rm}\{LCP\}\}\}}\rangle+cI(v)\rangle$, which includes a c term accounting for leakage of $I(v)$ into $V(v)$, thus yields values for B_{los} of the components. In Extended Data Table 1, we list the parameters of the components obtained from least-squares fits to $I(v)$ and $V(v)$. The leakage of our HINSA Zeeman observations is $c = 0.034\%$.

HINSA and H₂ column density maps

L1544 is a low-mass prestellar core in the Taurus molecular cloud complex at a distance of about 140 pc. The core has a size of about 0.1 pc (ref. 37), presumably formed out of a parsec-long elongated molecular ridge⁷, which, for simplicity, we refer as the molecular envelope. We show the HINSA and H₂ column density maps of L1544 in Fig. 1a, and we use the H₂ column density map to calculate the N_{H_2} of the envelope and core at the beams of FAST and Arecibo observations in Table 1. The HINSA column density map is a revision of Fig. 8 in ref. 4. To derive the H₂ column density map, we retrieved the level-2.5 processed, archival Herschel images that were taken at 250 μm/350 μm/500 μm using the SPIRE instrument³⁸ (observation ID 1342204842). We smoothed the Herschel images to a common angular resolution of the 36" beam at 500 μm and regridded the images to the same pixel size of 6". We performed least-squares fits of the 250 μm/350 μm/500 μm spectral energy distributions weighted by the squares of the measured noise levels to derive the pixel-

to-pixel distributions of dust temperature T_d and dust optical depth τ_v using $\langle S \rangle_{\nu} = \langle \Omega_m B_\mu(T_d) (1 - \langle \rm{e} \rangle)^{-\langle \tau \rangle_\nu} \rangle$, where S_ν is the flux density at frequency ν , Ω_m is the solid angle of the pixel, $B_\mu(T_d)$ is the Planck function at T_d and $\langle \tau \rangle_\nu = \langle \tau \rangle_{230} (\nu / (\nu_{\rm GHz}/230))^\beta$ with a dust opacity index β of 1.8. Next, we obtained the H₂ column density with $\langle N \rangle_{\rm H_2} = g \langle \tau \rangle_{230} / (\kappa_{230})$, where $g = 100$ is the gas-to-dust mass ratio, $\kappa_{230} = 0.09 \text{ cm}^2 \text{ g}^{-2}$ (ref. 39) is the dust opacity at 230 GHz, $\mu_m = 2.8$ is the mean molecular weight and m_H is the atomic mass of hydrogen. To estimate the uncertainties in the H₂ column density, we used a Monte Carlo technique. For each pixel, we created artificial 250 μm/350 μm/500 μm flux densities by adding the original flux densities with normal-distributed errors taking account the uncertainty in the measured flux and a 10% correlation for the calibration uncertainty in SPIRE⁴⁰. We then estimated the uncertainty in each pixel with 1,000 fittings of the H₂ column density. The $N_{\rm H_2}$ and its uncertainty in Table 1 were obtained from the convolutions of the H₂ column density map and uncertainty map with the FAST and Arecibo beams.

Note that the equivalent H₂ column density $N_{\rm H_2}$ of the CNM1 are derived from H I data towards 3C132 and 3C133, a method different to the $N_{\rm H_2}$ of the L1544 envelope and core that are derived from dust emission. Therefore, in addition to the statistical errors listed in Table 1, there is a systematic difference between the $N_{\rm H_2}$ derived from the two methods. Considering that the regime traced by dust emission can be different from those traced by HINSA or OH, which is particularly noticeable from the different spatial extents of dust and HINSA in Fig. 1a, we expect that the systematic difference could be as large as a factor of a few. As the values of λ between CNM1 and L1544 are different at least by a factor of 13, the systematic difference between the two methods should not change the qualitative conclusion of this work.

In Fig. 1a, the peak of HINSA column density appears to be shifted from the centre of L1544 by 0.15 pc and the 70% and 90% contours of the peak HINSA column density do not enclose L1544. We note that such offset has also been seen for other dense gas tracers in prestellar cores⁴¹. The core geometry may not be as simple as envisioned in idealized theories, where the dense core sits near the centre of a lower-density molecular envelope. In particular, the L1544 core appears to sit near one end of an elongated molecular (and dust) ridge, which roughly coincides with the region traced by the HINSA. Such an offset can result from complexities in chemistry and formation history, but does not affect the main science result of this work, namely, the HINSA Zeeman probes the magnetic fields of the current molecular ridge that is the progenitor of the dense core.

Maximum likelihood

We adopt the analysis of maximum likelihood¹⁸ to study the uniformity of magnetic fields in the envelope of L1544. Assuming that the true B_{los} follows a Gaussian distribution with mean B_0 and intrinsic spread σ_0 , the likelihood l_j for a single observation in a set of N measurements ($j = 1, \dots, N$) to measure B_j with Gaussian error σ_j is proportional to the convolution of the probability $\exp[-(B - B_0)^2/2\sigma_0^2]$ for the magnetic field to have a true value of B with the probability $\exp[-(B_j - B_0)^2/2\sigma_j^2]$ of observing a value B_j of the field. Therefore, l_j is the integral over all possible true values of the magnetic field

$$\$ \$ l_j = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dB \frac{\exp[-(B - B_0)^2/2\sigma_0^2]}{\sqrt{2\pi}\sigma_0} \frac{\exp[-(B_j - B)^2/2\sigma_j^2]}{\sqrt{2\pi}\sigma_j} dB. \quad (7)$$

Although the overall likelihood \mathcal{L} for a set of observations is the product of individual likelihoods of the observations $\mathcal{L} = \prod_{j=1}^N l_j$, the B_0 and σ_0 can be estimated by maximizing the likelihood \mathcal{L} . After performing the integration in equation (7) and some algebraic manipulations

$$\$ \$ \mathcal{L}(B_0, \sigma_0) = \left(\prod_{j=1}^N \frac{1}{\sqrt{2\pi}\sigma_j} \exp \left[-\frac{(B_j - B_0)^2}{2\sigma_0^2} \right] \right) \exp \left[-\frac{1}{2\sigma_0^2} \sum_{j=1}^N (B_j - B_0)^2 \right]. \quad (8)$$

Extended Data Fig. 1 shows the distribution of \mathcal{L} as functions of B_0 and σ_0 and the probability distributions of B_0 and σ_0 by integrating \mathcal{L} along the B_0 axis and the σ_0 axis, respectively. The probability distribution of B_0 is similar to a normal distribution with a mean value of $+4.1 \mu\text{G}$ and a standard deviation of $1.6 \mu\text{G}$. The probability distribution of σ_0 is highly asymmetric as the values of σ_0 cannot be negative. The first, second and third quartiles of the σ_0 distribution are $0.6 \mu\text{G}$, $1.2 \mu\text{G}$ and $2.4 \mu\text{G}$. We therefore suspect that the Zeeman measurements in the L1544 envelope can be explained by a magnetic field with $B_0 = +4.1 \pm 1.6 \mu\text{G}$ and $\langle s_0 \rangle = 1.2 \times (-0.6)^{+1.2} \langle \mu \rangle \langle G \rangle$.

Inclination angle of magnetic field

Given the uniformity of magnetic fields in the envelope of L1544 and CNM1 is well constrained by the maximum likelihood analysis, the coherent B_{los} suggests that the inclination angles of magnetic fields in the CNM1 and L1544 envelope are likely to be similar, or a special geometry of magnetic field structure across multi-scales and multi-phases of the interstellar medium is needed. In contrast, the B_{los} of the L1544 envelope and core differ by a factor of 2.6. There are two physical explanations for the 2.6-times difference between the HINSA and OH Zeeman measurements. First, the OH measurement probably samples a denser gas than the HINSA measurement, as the column density along the OH sightline is twice that along the HINSA sightline (Table 1). As the magnetic field strength in molecular clouds tends to increase with number density¹⁵, the stronger field is naturally expected in the denser core. Alternatively, the inclination angle of the L1544 core magnetic field could differ substantially from that of the coherent field. As we cannot rule out the second possibility, an assumption of similar inclination angles of the magnetic fields in the L1544 envelope and core thus is required to calculate the relative values of λ .

We note that dust polarization observation may give some clues as dust polarization traces the position angle of the plane-of-sky component of the magnetic field. The near-infrared polarization observations of L1544⁴² indicate that the mean position angle of the magnetic field towards the core location of the Arecibo beam is 29.0° – 36.9° , and the mean position angles of magnetic fields towards the four envelope locations of the GBT beams are 30.5° – 55.8° . The difference in the position angles between the core and envelope thus may be about 10° – 20° .

We perform Monte Carlo simulations⁴³ to study whether the 2.6-times difference between the B_{los} of the L1544 envelope and core can be explained by different inclination angles. The simulations randomly generate two unit vectors in three dimensions, and then measure the difference between the inclination angles and the difference between the position angles of the two vectors. The probability of the cases that the line-of-sight length of one vector is 2.6-times larger than that of the other is roughly 0.19. For those cases, the mean difference between the inclination angles and the mean difference between the position angles of the two vectors are about 38° and 45° , respectively. As the probability of 0.19 is small and the difference of about 45° between the simulated position angles is about a factor of two to four times larger than the difference of about 10° – 20° between the observed position angles, it is less likely that the 2.6-times difference between the B_{los} of the L1544 envelope and core can be solely explained by different inclination angles.

CCS Zeeman measurements

Ref. [20](#) reported a CCS Zeeman detection of $117 \pm 21 \mu\text{G}$ in the dense core of TMC-1 that has an estimated H_2 column density of $3 \times 10^{22} \text{ cm}^{-2}$, which is four-times higher than that probed by OH Zeeman measurements in L1544 and nearly one order of magnitude higher than that probed by our HINSA measurements. It appears to provide further support to the evolutionary scenario suggested by our HINSA measurements: namely, once the gas loses its magnetic support during the transition from the CNM to the molecular envelope (or ridge) and becomes magnetically supercritical, there is no longer any need to lose magnetic flux further (relative to the mass) for a piece of the envelope/ridge to condense into a (magnetically supercritical) core (for example, the L1544 core probed by OH) and for the core to evolve further by increasing its column density (for example, the TMC-1 core probed by CCS).

Technically, we note that one potential source of significant uncertainty in frequency shift, namely the uncertainty of beam squint, was not included in the CCS result, which may affect the level of significance. In comparison, the HINSA measurement is robust with a greater than 10σ significance with the beam squint and velocity gradient being taken into account by convolving the FAST Stokes V beam with the Stokes I cube of L1544 (see the third paragraph in the data reduction section in [Methods](#)).

Data availability

The data that support the findings of this study are openly available in Science Data Bank at <https://www.doi.org/10.11922/sciencedb.01221>. The FAST raw data are available from the <http://fast.bao.ac.cn> site one year after data-taking, per the data policy of FAST. Owing to the large data volume of this work and the speciality of polarization calibration, interested users are encouraged to contact the corresponding authors to arrange data transfer. The reduced $I(v)$ and $V(v)$ spectra are available at <https://github.com/taochung/HINSAzeeman>.

Code availability

The codes analysing the $I(v)$ and $V(v)$ spectra reported here are available at <https://github.com/taochung/HINSAzeeman>. The IDL RHSTK package is available at <http://w.astro.berkeley.edu/heiles/>.

References

1. 1.

McKee, C. F. & Ostriker, E. C. Theory of star formation. *Annu. Rev. Astron. Astrophys.* **45**, 565–687 (2007).

2. 2.

Hennebelle, P. & Inutsuka, S.-i. The role of magnetic field in molecular cloud formation and evolution. *Front. Astron. Space Sci.* **6**, 5 (2019).

3. 3.

Crutcher, R. M. & Kemball, A. J. Review of Zeeman effect observations of regions of star formation. *Front. Astron. Space Sci.* **6**, 66 (2019).

4. 4.

Li, D. & Goldsmith, P. F. H I narrow self-absorption in dark clouds. *Astrophys. J.* **585**, 823–839 (2003).

5. 5.

Goldsmith, P. F. & Li, D. H I narrow self-absorption in dark clouds: correlations with molecular gas and implications for cloud evolution and star formation. *Astrophys. J.* **622**, 938–958 (2005).

6. 6.

Elias, J. H. A study of the Taurus dark cloud complex. *Astrophys. J.* **224**, 857–872 (1978).

7. 7.

Tafalla, M. et al. L1544: a starless dense core with extended inward motions. *Astrophys. J.* **504**, 900–914 (1998).

8. 8.

Aikawa, Y., Ohashi, N., Inutsuka, S.-i., Herbst, E. & Takakuwa, S. Molecular evolution in collapsing prestellar cores. *Astrophys. J.* **552**, 639–653 (2001).

9. 9.

Li, Z.-Y., Shematovich, V. I., Wiebe, D. S. & Shustov, B. M. A coupled dynamical and chemical model of starless cores of magnetized molecular clouds. I. Formulation and initial results. *Astrophys. J.* **569**, 792–802 (2002).

10. 10.

Keto, E., Caselli, P. & Rawlings, J. The dynamics of collapsing cores and star formation. *Mon. Not. R. Astron. Soc.* **446**, 3731–3740 (2015).

11. 11.

Caselli, P. et al. The central 1000 AU of a pre-stellar core revealed with ALMA. I. 1.3 mm continuum observations. *Astrophys. J.* **874**, 89 (2019).

12. 12.

Crapsi, A., Caselli, P., Walmsley, M. C. & Tafalla, M. Observing the gas temperature drop in the high-density nucleus of L 1544. *Astron. Astrophys.* **470**, 221–230 (2007).

13. 13.

Shu, F. H., Adams, F. C. & Lizano, S. Star formation in molecular clouds: observation and theory. *Annu. Rev. Astron. Astrophys.* **25**, 23–81 (1987).

14. 14.

Mouschovias, T. C. & Ciolek, G. E. in *The Origin of Stars and Planetary Systems* (eds Lada, C. J. & Kylafis, N. D.) 305–340 (NATO Advanced Study Institute Series C, **Vol. 540**, Springer, 1999).

15. 15.

Crutcher, R. M., Wandelt, B., Heiles, C., Falgarone, E. & Troland, T. H. Magnetic fields in interstellar clouds from Zeeman observations: inference of total field strengths by Bayesian analysis. *Astrophys. J.* **725**, 466–479 (2010).

16. 16.

Crutcher, R. M., Hakobian, N. & Troland, T. H. Testing magnetic star formation theory. *Astrophys. J.* **692**, 844–855 (2009).

17. 17.

Crutcher, R. M., Hakobian, N. & Troland, T. H. Self-consistent analysis of OH Zeeman observations. *Mon. Not. R. Astron. Soc.* **402**, L64–L66 (2010).

18. 18.

Mouschovias, T. C. & Tassis, K. Testing molecular-cloud fragmentation theories: self-consistent analysis of OH Zeeman observations. *Mon. Not. R. Astron. Soc.* **400**, L15–L19 (2009).

19. 19.

Mouschovias, T. C. & Tassis, K. Self-consistent analysis of OH-Zeeman observations: too much noise about noise. *Mon. Not. R. Astron. Soc.* **409**, 801–807 (2010).

20. 20.

Nakamura, F. et al. First clear detection of the CCS Zeeman splitting toward the pre-stellar core, Taurus molecular cloud 1. *Publ. Astron. Soc. Jpn* **71**, psz102 (2019).

21. 21.

Goodman, A. A. & Heiles, C. The magnetic field in the Ophiuchus dark cloud complex. *Astrophys. J.* **424**, 208–221 (1994).

22. 22.

Heiles, C. A holistic view of the magnetic field in the Eridanus/Orion region. *Astrophys. J. Suppl. Ser.* **111**, 245–288 (1997).

23. 23.

Goldsmith, P. F., Li, D. & Krčo, M. The transition from atomic to molecular hydrogen in interstellar clouds: 21 cm signature of the evolution of cold atomic hydrogen in dense clouds. *Astrophys. J.* **654**, 273–289 (2007).

24. 24.

Crutcher, R. M. & Troland, T. H. OH Zeeman measurement of the magnetic field in the L1544 core. *Astrophys. J.* **537**, L139–L142 (2000).

25. 25.

Li, D. et al. FAST in space: considerations for a multibeam, multipurpose survey using China’s 500-m aperture spherical radio telescope (FAST). *IEEE Microw. Mag.* **19**, 112–119 (2018).

26. 26.

Heiles, C. & Troland, T. H. The Millennium Arecibo 21 Centimeter Absorption-Line Survey. I. Techniques and Gaussian fits. *Astrophys. J. Suppl. Ser.* **145**, 329–354 (2003).

27. 27.

Heiles, C. & Troland, T. H. The Millennium Arecibo 21 Centimeter Absorption-Line Survey. III. Techniques for spectral polarization and results for Stokes V . *Astrophys. J. Suppl. Ser.* **151**, 271–297 (2004).

28. 28.

Heiles, C. & Troland, T. H. The Millennium Arecibo 21 Centimeter Absorption-Line Survey. IV. Statistics of magnetic field, column density, and turbulence. *Astrophys. J.* **624**, 773–793 (2005).

29. 29.

Vázquez-Semadeni, E. et al. Molecular cloud evolution—IV. Magnetic fields, ambipolar diffusion and the star formation efficiency. *Mon. Not. R. Astron. Soc.* **414**, 2511–2527 (2011).

30. 30.

Lazarian, A., Esquivel, A. & Crutcher, R. Magnetization of cloud cores and envelopes and other observational consequences of reconnection diffusion. *Astrophys. J.* **757**, 154 (2012).

31. 31.

Jiang, P. et al. The fundamental performance of FAST with 19-beam receiver at L band. *Res. Astron. Astrophys.* **20**, 064 (2020).

32. 32.

Heiles, C. et al. Mueller matrix parameters for radio telescopes and their observational determination. *Publ. Astron. Soc. Pac.* **113**, 1274–1288 (2001).

33. 33.

Jehle, M., Ruegg, M., Zuberbuhler, L., Small, D. & Meier, E. Measurement of ionospheric Faraday rotation in simulated and real spaceborne SAR data. *IEEE Trans. Geosci. Remote Sens.* **47**, 1512–1523 (2009).

34. 34.
Perley, R. A. & Butler, B. J. Integrated polarization properties of 3C48, 3C138, 3C147, and 3C286. *Astrophys. J. Suppl. Ser.* **206**, 16 (2013).
35. 35.
McBride, J. & Heiles, C. An Arecibo survey for Zeeman Sslitting in OH megamaser galaxies. *Astrophys. J.* **763**, 8 (2013).
36. 36.
Peek, J. E. G. et al. The GALFA-H I Survey data release 2. *Astrophys. J. Suppl. Ser.* **234**, 2 (2018).
37. 37.
Ward-Thompson, D., Motte, F. & Andre, P. The initial conditions of isolated star formation—III. Millimetre continuum mapping of pre-stellar cores. *Mon. Not. R. Astron. Soc.* **305**, 143–150 (1999).
38. 38.
Griffin, M. J. et al. The Herschel-SPIRE instrument and its in-flight performance. *Astron. Astrophys.* **518**, L3 (2010).
39. 39.
Ossenkopf, V. & Henning, T. Dust opacities for protostellar cores. *Astron. Astrophys.* **291**, 943–959 (1994).
40. 40.
Roy, A. et al. Reconstructing the density and temperature structure of prestellar cores from Herschel data: a case study for B68 and L1689B. *Astron. Astrophys.* **562**, A138 (2014).
41. 41.
Lai, S.-P., Velusamy, T., Langer, W. D. & Kuiper, T. B. H. The physical and chemical status of pre-protostellar core B68. *Astron. J.* **126**, 311–318 (2003).
42. 42.

Clemens, D. P., Tassis, K. & Goldsmith, P. F. The magnetic field of L1544. I. Near-infrared polarimetry and the non-uniform envelope. *Astrophys. J.* **833**, 176 (2016).

43. 43.

Hull, C. L. H. et al. Misalignment of magnetic fields and outflows in protostellar cores. *Astrophys. J.* **768**, 159 (2013).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (NSFC) grant numbers 11988101, U1931117 and 11725313; by the CAS International Partnership Program number 114A11KYSB20160008; by the National Key R&D Program of China number 2017YFA0402600; and by the Cultivation Project for FAST Scientific Payoff and Research Achievement of CAMS-CAS. T.-C.C. is funded by the Chinese Academy of Sciences Taiwan Young Talent Program grant number 2018TW2JB0002. T.-C.C. and J.T. were supported by Special Funding for Advanced Users, budgeted and administrated by Center for Astronomical Mega-Science (CAMS), Chinese Academy of Sciences. C.H. is funded by the Chinese Academy of Sciences President's International Fellowship Initiative grant number 2020DM0005. Z.-Y.L. is supported in part by NASA 80NSSC20K0533 and NSF AST-1716259 and 1815784. This work made use of data from FAST, a Chinese national mega-science facility built and operated by the National Astronomical Observatories, Chinese Academy of Sciences. This research has made use of the services of the DSS2 survey of the ESO Science Archive Facility.

Author information

Affiliations

1. National Astronomical Observatories, Chinese Academy of Sciences, Beijing, China
T.-C. Ching, D. Li, L. Qian, Y. L. Yue, J. Tang & S. H. Jiao
2. Department of Astronomy, University of Chinese Academy of Sciences, Beijing, China
D. Li & S. H. Jiao

3. NAOC-UKZN Computational Astrophysics Centre, University of KwaZulu-Natal, Durban, South Africa

D. Li

4. Department of Astronomy, University of California, Berkeley, Berkeley, CA, USA

C. Heiles

5. Astronomy Department, University of Virginia, Charlottesville, VA, USA

Z.-Y. Li

Contributions

T.-C.C., D.L. and C.H. launched the FAST Zeeman project; T.-C.C. processed the data and analysis in consultation with C.H.; T.-C.C., Z.-Y.L., D.L. and C.H. drafted the paper; L.Q., Y.L.Y. and J.T. made key contributions to arrange the FAST observations of L1544 and polarization calibration; S.H.J. provided the H₂ column density map.

Corresponding author

Correspondence to [D. Li](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Richard Crutcher and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

[Extended Data Fig. 1 Likelihood \$\(\mathcal{L}\)\$ for the coherent magnetic field to have mean \(\$B_0\$ \) and spread \(\$\sigma_0\$ \) values.](#)

a, Contours of (\mathcal{L}) as functions of B_0 and σ_0 plotted at 10%, 30%, 50%, 70% and 90% of the peak value. **b**, The probability distribution of B_0 while allowing all possible values of σ_0 . **c**, The probability distribution of σ_0 while allowing all possible values of B_0 .

Extended Data Table 1 Gaussian fit parameters

Supplementary information

[Peer Review File](#)

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

[Reprints and Permissions](#)

About this article

Cite this article

Ching, TC., Li, D., Heiles, C. *et al.* An early transition to magnetic supercriticality in star formation. *Nature* **601**, 49–52 (2022). <https://doi.org/10.1038/s41586-021-04159-x>

- Received: 19 March 2021
- Accepted: 18 October 2021

- Published: 05 January 2022
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04159-x>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

The magnetic field of a molecular cloud revealed

Research Briefing 05 Jan 2022

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04159-x>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 05 January 2022](#)

A 16-parts-per-trillion measurement of the antiproton-to-proton charge–mass ratio

- [M. J. Borchert](#)^{1,2,3},
- [J. A. Devlin](#)^{1,4},
- [S. R. Erlewein](#)^{1,4,5},
- [M. Fleck](#) ORCID: orcid.org/0000-0003-4114-1902^{1,6},
- [J. A. Harrington](#)^{1,5},
- [T. Higuchi](#) ORCID: orcid.org/0000-0003-3281-4669^{1,6},
- [B. M. Latacz](#) ORCID: orcid.org/0000-0003-2320-1713¹,
- [F. Voelksen](#)^{1,7},
- [E. J. Wursten](#) ORCID: orcid.org/0000-0002-2413-2214^{1,4,5},
- [F. Abbass](#)⁸,
- [M. A. Bohman](#) ORCID: orcid.org/0000-0003-1322-1489^{1,5},
- [A. H. Moosser](#)⁵,
- [D. Popper](#)⁸,
- [M. Wiesinger](#) ORCID: orcid.org/0000-0003-0111-8812^{1,5},
- [C. Will](#) ORCID: orcid.org/0000-0003-4622-7799⁵,
- [K. Blaum](#) ORCID: orcid.org/0000-0003-4468-9316⁵,
- [Y. Matsuda](#) ORCID: orcid.org/0000-0002-9847-3791⁶,
- [C. Ospelkaus](#) ORCID: orcid.org/0000-0002-4170-2936^{2,3},
- [W. Quint](#)⁷,
- [J. Walz](#)^{8,9},
- [Y. Yamazaki](#)¹,
- [C. Smorra](#) ORCID: orcid.org/0000-0001-5584-7960^{1,8} &
- [S. Ulmer](#) ORCID: orcid.org/0000-0002-4185-4147¹

[Nature](#) volume **601**, pages 53–57 (2022)

- 1083 Accesses
- 263 Altmetric
- [Metrics details](#)

Subjects

- [Exotic atoms and molecules](#)
- [Experimental particle physics](#)

Abstract

The standard model of particle physics is both incredibly successful and glaringly incomplete. Among the questions left open is the striking imbalance of matter and antimatter in the observable universe¹, which inspires experiments to compare the fundamental properties of matter/antimatter conjugates with high precision^{2,3,4,5}. Our experiments deal with direct investigations of the fundamental properties of protons and antiprotons, performing spectroscopy in advanced cryogenic Penning trap systems⁶. For instance, we previously compared the proton/antiproton magnetic moments with 1.5 parts per billion fractional precision^{7,8}, which improved upon previous best measurements⁹ by a factor of greater than 3,000. Here we report on a new comparison of the proton/antiproton charge-to-mass ratios with a fractional uncertainty of 16 parts per trillion. Our result is based on the combination of four independent long-term studies, recorded in a total time span of 1.5 years. We use different measurement methods and experimental set-ups incorporating different systematic effects. The final result, $\langle -\{q/m\}_p/\{q/m\}_{\bar{p}} \rangle = 1.00000000003(16)$, is consistent with the fundamental charge–parity–time reversal invariance, and improves the precision of our previous best measurement⁶ by a factor of 4.3. The measurement tests the standard model at an energy scale of 1.96×10^{-27} gigaelectronvolts (confidence level 0.68), and improves ten coefficients of

the standard model extension¹⁰. Our cyclotron clock study also constrains hypothetical interactions mediating violations of the clock weak equivalence principle (WEP_{cc}) for antimatter to less than 1.8×10^{-7} , and enables the first differential test of the WEP_{cc} using antiprotons¹¹. From this interpretation we constrain the differential WEP_{cc}-violating coefficient to less than 0.030.

This is a preview of subscription content

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Elements of the experiment to determine the antiproton-to-H⁻ charge-to-mass ratio.

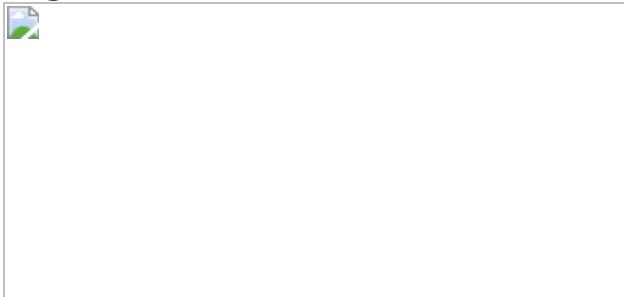


Fig. 2: Results.

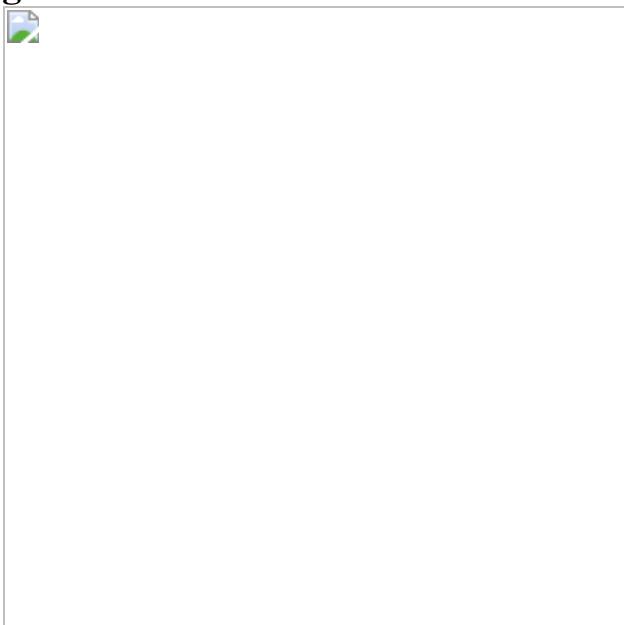


Fig. 3: Trajectory of the Earth on its orbit around the Sun.



Data and code availability

The datasets and analysis codes will be made available on reasonable request.

References

1. 1.

Dine, M. & Kusenko, A. Origin of the matter–antimatter asymmetry. *Rev. Mod. Phys.* **76**, 1–30 (2003).

2. 2.

Van Dyck Jr, R. S., Schwinberg, P. B. & Dehmelt, H. G. New high-precision comparison of electron and positron g factors. *Phys. Rev. Lett.* **59**, 26–29 (1987).

3. 3.

Ahmadi, M. et al. Characterization of the 1S–2S transition in antihydrogen. *Nature* **557**, 71–75 (2018).

4. 4.

Hori, M. et al. Buffer-gas cooling of antiprotonic helium to 1.5 K to 1.7 K, and antiproton-to-electron mass ratio. *Science* **354**, 610–614 (2016).

5. 5.

Schwingenheuer, B. et al. CPT tests in the neutral kaon system. *Phys. Rev. Lett.* **74**, 4376–4379 (1995).

6. 6.

Ulmer, S. et al. High-precision comparison of the antiproton-to-proton charge-to-mass ratio. *Nature* **524**, 196–199 (2015).

7. 7.

Smorra, C. et al. A parts-per-billion measurement of the antiproton magnetic moment. *Nature* **550**, 371–374 (2017).

8. 8.

Schneider, G. et al. Double-trap measurement of the proton magnetic moment at 0.3 parts per billion precision. *Science* **358**, 1081–1084 (2017).

9. 9.

DiSciacca, J. et al. One-particle measurement of the antiproton magnetic moment. *Phys. Rev. Lett.* **110**, 130801 (2013).

10. 10.

Ding, Y. & Rawnak, M. F. Lorentz and CPT tests with charge-to-mass ratio comparisons in Penning traps. *Phys. Rev. D* **102**, 056009 (2020).

11. 11.

Hughes, R. J. & Holzscheiter, M. H. Constraints on the gravitational properties of antiprotons and positrons from cyclotron-frequency

measurements. *Phys. Rev. Lett.* **66**, 854–857 (1991).

12. 12.

Lehnert, R. CPT symmetry and its violation. *Symmetry* **8**, 114 (2016).

13. 13.

Lüders, G. Proof of the TCP theorem. *Ann. Phys.* **2**, 1–15 (1957).

14. 14.

Edwards, B. R. & Kostelecký, V. A. Riemann–Finsler geometry and Lorentz-violating scalar fields. *Phys. Lett. B* **786**, 319–326 (2018).

15. 15.

Tsujikawa, S. Quintessence: a review. *Class. Quantum Gravity* **30**, 214003 (2013).

16. 16.

Kostelecký, V. A. & Potting, R. CPT and strings. *Nucl. Phys. B* **359**, 545–570 (1991).

17. 17.

Weinberg, S. *Cosmology* (Oxford Univ. Press, 2008).

18. 18.

Hughes, R. J. Constraints on new macroscopic forces from gravitational redshift experiments. *Phys. Rev. D* **41**, 2367–2373 (1990).

19. 19.

Smorra, C. et al. BASE – the Baryon Antibaryon Symmetry Experiment. *Eur. Phys. J. Spec. Top.* **224**, 3055–3108 (2015).

20. 20.

Sellner, S. et al. Improved limit on the directly measured antiproton lifetime. *New J. Phys.* **19**, 083023 (2017).

21. 21.

Smorra, C. et al. A reservoir trap for antiprotons. *Int. J. Mass Spectrom.* **389**, 10–13 (2015).

22. 22.

Gabrielse, G., Haarsma, L. & Rolston, S. L. Open-endcap Penning traps for high precision experiments. *Int. J. Mass Spectrom.* **88**, 319–332 (1989).

23. 23.

Brown, L. S. & Gabrielse, G. Geonium theory: physics of a single electron or ion in a Penning trap. *Rev. Mod. Phys.* **58**, 233–311 (1986).

24. 24.

Gabrielse, G. et al. Precision mass spectroscopy of the antiproton and proton using simultaneously trapped particles. *Phys. Rev. Lett.* **82**, 3198–3201 (1999).

25. 25.

Nagahama, H. et al. Highly sensitive superconducting circuits at 700 kHz with tunable quality factors for image-current detection of single trapped antiprotons. *Rev. Sci. Instrum.* **87**, 113305 (2016).

26. 26.

BASE Collaboration. *Future Program of the BASE Experiment at the Antiproton Decelerator of CERN*. Document no. CERN-SPSC-2019-047; SPSC-P-363 (CERN, 2019); available at <https://cds.cern.ch/record/2702758>.

27. 27.

Devlin, J. A. et al. Superconducting solenoid system with adjustable shielding factor for precision measurements of the properties of the antiproton. *Phys. Rev. Appl.* **12**, 044012 (2019).

28. 28.

Heiße, F. et al. High-precision measurement of the proton's atomic mass. *Phys. Rev. Lett.* **119**, 033001 (2017).

29. 29.

Cornell, E. A., Weisskoff, R. M., Boyce, K. R. & Pritchard, D. E. Mode coupling in a Penning trap: π pulses and a classical avoided crossing. *Phys. Rev. A* **41**, 312–315 (1990).

30. 30.

Ketter, J., Eronen, T., Höcker, M., Streubel, S. & Blaum, K. First-order perturbative calculation of the frequency-shifts caused by static cylindrically-symmetric electric and magnetic imperfections of a Penning trap. *Int. J. Mass Spectrom.* **358**, 1–16 (2014).

31. 31.

Hoaglin, D. C., Mosteller, F. & Tukey, J. W. *Understanding Robust and Exploratory Data Analysis* (Wiley, 2000).

32. 32.

Le Cam, L. *Asymptotic Methods in Statistical Decision Theory* (Springer, 2012).

33. 33.

Rao, C. R. Information and the accuracy attainable in the estimation of statistical parameters. In *Breakthroughs in Statistics* (eds Kotz S. & Johnson, N. L.) 235–247 (Springer, 1992).

34. 34.

Natarayan, V. *Penning Trap Mass Spectroscopy at 0.1 ppb*. PhD thesis, MIT (1993).

35. 35.

Wang, Y. & Liu, Q. Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of stock–recruitment relationships. *Fish. Res.* **77**, 220–225 (2006).

36. 36.

Charlton, M., Eriksson, S. & Shore, G. Testing fundamental physics in antihydrogen experiments. Preprint at <https://arxiv.org/abs/2002.09348> (2020).

37. 37.

Kenyon, I. A recalculation on the gravitational mass difference between the K^0 and \bar{K}^0 mesons. *Phys. Lett. B* **237**, 274–277 (1990).

38. 38.

Tchernin, C., Lau, E. T., Stapelberg, S., Hug, D. & Bartelmann, M. Characterizing galaxy clusters by their gravitational potential: systematics of cluster potential reconstruction. *Astron. Astrophys.* **644**, A126 (2020).

39. 39.

Chardin, G. & Manfredi, G. Gravity, antimatter and the Dirac–Milne universe. *Hyperfine Interact.* **239**, 45 (2018).

40. 40.

Super-Kamiokande Collaboration. Search for proton decay via $p \rightarrow e^+\pi^0$ and $p \rightarrow \mu^+\pi^0$ in 0.31 megaton years exposure of the Super-

Kamiokande water Cherenkov detector. *Phys. Rev. D* **95**, 012004 (2017).

41. 41.

Perez, P. & Sacquin, Y. The GBAR experiment: gravitational behaviour of antihydrogen at rest. *Class. Quantum Gravity* **29**, 184008 (2012).

42. 42.

Bertsche, W. A. Prospects for comparison of matter and antimatter gravitation with ALPHA-g. *Phil. Trans. R. Soc. A* **376**, 20170265 (2018).

43. 43.

Scampoli, P. & Storey, J. The AEgIS experiment at CERN for the measurement of antihydrogen gravity acceleration. *Mod. Phys. Lett. A* **29**, 1430017 (2014).

44. 44.

Bluhm, R., Kostelecký, V. A. & Russell, N. CPT and Lorentz tests in Penning traps. *Phys. Rev. D* **57**, 3932–3943 (1998).

45. 45.

Kostelecký, V. A. & Russell, N. Data tables for Lorentz and CPT violation. *Rev. Mod. Phys.* **83**, 11–31 (2011).

46. 46.

Smorra, C. et al. *Technical Design Report of BASE-STEP* (CERN, 2021); <https://cds.cern.ch/record/2756508/files/SPSC-TDR-007.pdf>

47. 47.

Thompson, J. K., Rainville, S. & Pritchard, D. E. Cyclotron frequency shifts arising from polarization forces. *Nature* **430**, 58–61 (2004).

48. 48.

Rau, S. et al. Penning trap mass measurements of the deuteron and the HD⁺ molecular ion. *Nature* **585**, 43–47 (2020).

49. 49.

Kortunov, I. et al. Proton–electron mass ratio by high-resolution optical spectroscopy of ion ensembles in the resolved carrier regime. *Nat. Phys.* **17**, 569–573 (2021).

50. 50.

Patra, S. et al. Proton–electron mass ratio from laser spectroscopy of HD⁺ at the part-per-trillion level. *Science* **369**, 1238–1241 (2020).

51. 51.

Kramida, A., Ralchenko, Yu., Reader, J. & NIST ASD Team. *NIST Atomic Spectra Database (ver. 5.8)*. (National Institute of Standards and Technology, accessed 26 November 2020);
<https://doi.org/10.18434/T4W30F>.

52. 52.

Parthey, C. G. et al. Improved measurement of the hydrogen 1S–2S transition frequency. *Phys. Rev. Lett.* **107**, 203001 (2011).

53. 53.

Jentschura, U. D., Kotochigova, S., Le Bigot, E.-O., Mohr, P. J. & Taylor, B. N. Precise calculation of transition frequencies of hydrogen and deuterium based on a least-squares analysis. *Phys. Rev. Lett.* **95**, 163003 (2005).

54. 54.

Lykke, K. R., Murray, K. K. & Lineberger, W. C. Threshold photodetachment of H⁻. *Phys. Rev. A* **43**, 6104–6107 (1991).

55. 55.

Sahoo, B. Determination of the dipole polarizability of the alkali-metal negative ions. *Phys. Rev. A* **102**, 022820 (2020).

56. 56.

Wineland, D. & Dehmelt, H. Principles of the stored ion calorimeter. *J. Appl. Phys.* **46**, 919–930 (1975).

57. 57.

D'Urso, B., Odom, B. & Gabrielse, G. Feedback cooling of a one-electron oscillator. *Phys. Rev. Lett.* **90**, 043001 (2003).

Acknowledgements

We acknowledge technical support by CERN, especially the Antiproton Decelerator operation group, CERN's cryolab team and engineering department, and all other CERN groups that provide support to Antiproton Decelerator experiments. We acknowledge Y. Ding for comments in the discussion of the updated SME limits. We acknowledge financial support by RIKEN, the RIKEN EEE pioneering project funding, the RIKEN SPDR and JRA programme, the Max Planck Society, the European Union (FunI-832848, STEP-852818), CRC 1227 ‘DQ-mat’(DFG 274200144), the Cluster of Excellence ‘Quantum Frontiers’ (DFG 390837967), AVA-721559, the CERN fellowship programme and the Helmholtz-Gemeinschaft. This work was supported by the Max Planck, RIKEN, PTB Center for Time, Constants, and Fundamental Symmetries (C-TCFS).

Author information

Affiliations

1. Ulmer Fundamental Symmetries Laboratory, RIKEN, Saitama, Japan

M. J. Borchert, J. A. Devlin, S. R. Erlewein, M. Fleck, J. A. Harrington, T. Higuchi, B. M. Latacz, F. Voelksen, E. J. Wursten, M. A. Bohman, M. Wiesinger, Y. Yamazaki, C. Smorra & S. Ulmer

2. Institut für Quantenoptik, Leibniz Universität Hannover, Hannover, Germany

M. J. Borchert & C. Ospelkaus

3. Physikalisch-Technische Bundesanstalt, Braunschweig, Germany

M. J. Borchert & C. Ospelkaus

4. CERN, Meyrin, Switzerland

J. A. Devlin, S. R. Erlewein & E. J. Wursten

5. Max-Planck-Institut für Kernphysik, Heidelberg, Germany

S. R. Erlewein, J. A. Harrington, E. J. Wursten, M. A. Bohman, A. H. Mooser, M. Wiesinger, C. Will & K. Blaum

6. Graduate School of Arts and Sciences, University of Tokyo, Tokyo, Japan

M. Fleck, T. Higuchi & Y. Matsuda

7. GSI-Helmholtzzentrum für Schwerionenforschung, Darmstadt, Germany

F. Voelksen & W. Quint

8. Institut für Physik, Johannes Gutenberg-Universität, Mainz, Germany

F. Abbass, D. Popper, J. Walz & C. Smorra

9. Helmholtz-Institut Mainz, Johannes Gutenberg-Universität, Mainz,
Germany

J. Walz

Contributions

The experiment was designed and built by S.U. and C.S.; and M.J.B., J.A.D., J.A.H., T.H. and E.J.W. developed several technical upgrades. J.A.H., S.U., T.H., J.A.D., E.J.W. and M.J.B. developed the control code. J.A.H., M.J.B., T.H., J.A.D., E.J.W. and S.U. took part in the data acquisition. M.J.B., S.U., J.A.D., J.A.H., E.J.W. and M.F. performed the systematic studies. J.A.H., M.J.B., T.H., J.A.D., E.J.W., S.R.E. and S.U. contributed to the maintenance of the experiment during the measurement campaign. The data were analysed by S.U., E.J.W. and J.A.H.; and J.A.D., M.J.B., B.M.L. and C.W. contributed to the systematic analysis. The final results were discussed with all co-authors. The manuscript was written by S.U. and discussed with E.J.W., J.A.D., B.M.L., C.S. and K.B.; all co-authors discussed and approved the content.

Corresponding author

Correspondence to [S. Ulmer](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Energy calibration.

Upper: Measured cyclotron frequency shift as a function of the measured axial frequency shift. Lower: Measured cyclotron frequency shift as a function of particle energy E_+ .

Extended Data Fig. 2 Dominant systematic uncertainty.

Upper: measured axial frequency ratio as a function of the frequency ratio of the axial detection resonators. We observe a weak linear scaling of the measured axial frequency ratio as a function of the detuning of the axial frequency with respect to the resonator centre. Green line: weighted linear fit, red and blue functions represent CL 0.68 and CL 0.95 error bands.
Lower: Residuals of upper plot.

Extended Data Fig. 3 B_2 -imposed uncertainty in peak frequency ratio.

Upper: Sensitivity of the frequency ratio R as a function of the coefficient B_2 for different particle energy differences $E_{+,p} - E_{+,H}$, expressed as $\Delta R / \Delta B_2(\Delta E_+)$. Lower: Measured particle energy differences $E_{+,p} - E_{+,H}$ throughout the peak run. The green vertical lines indicate the mean difference and the uncertainty, the vertical green lines define the frequency-ratio shift and its uncertainty caused by the uncertainties in energy similarity and B_2 .

Extended Data Table 1 Summary of lineshape corrections applied to the different datasets

Extended Data Table 2 Axial temperatures

Extended Data Table 3 Total uncertainty budget

Extended Data Table 4 Improved SME coefficients

Supplementary information

Supplementary Information

This file contains a description of the data analysis, data cleaning, and data processing. It also contains notes regarding order systematic shifts and includes Supplementary Figs 1–10 and Supplementary Tables 1–3.

Peer Review File

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Borchert, M.J., Devlin, J.A., Erlewein, S.R. *et al.* A 16-parts-per-trillion measurement of the antiproton-to-proton charge–mass ratio. *Nature* **601**, 53–57 (2022). <https://doi.org/10.1038/s41586-021-04203-w>

- Received: 25 May 2021
- Accepted: 03 November 2021
- Published: 05 January 2022
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04203-w>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

Mirror symmetry validated for proton and its antimatter twin

- Ralf Lehnert

News & Views 05 Jan 2022

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04203-w>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 05 January 2022](#)

Crystallization of bosonic quantum Hall states in a rotating quantum gas

- [Biswaroop Mukherjee](#) ORCID: [orcid.org/0000-0002-7155-5412¹](https://orcid.org/0000-0002-7155-5412),
- [Airlia Shaffer¹](#),
- [Parth B. Patel¹](#),
- [Zhenjie Yan¹](#),
- [Cedric C. Wilson](#) ORCID: [orcid.org/0000-0001-6586-7666¹](https://orcid.org/0000-0001-6586-7666),
- [Valentin Crépel¹](#),
- [Richard J. Fletcher](#) ORCID: [orcid.org/0000-0002-3650-2446¹](https://orcid.org/0000-0002-3650-2446) &
- [Martin Zwierlein](#) ¹

[Nature](#) volume **601**, pages 58–62 (2022)

- 1064 Accesses
- 122 Altmetric
- [Metrics details](#)

Subjects

- [Bose–Einstein condensates](#)
- [Quantum fluids and solids](#)
- [Quantum Hall](#)
- [Ultracold gases](#)

This article has been [updated](#)

Abstract

The dominance of interactions over kinetic energy lies at the heart of strongly correlated quantum matter, from fractional quantum Hall liquids¹, to atoms in optical lattices² and twisted bilayer graphene³. Crystalline phases often compete with correlated quantum liquids, and transitions between them occur when the energy cost of forming a density wave approaches zero. A prime example occurs for electrons in high-strength magnetic fields, where the instability of quantum Hall liquids towards a Wigner crystal^{4,5,6,7,8,9} is heralded by a roton-like softening of density modulations at the magnetic length^{7,10,11,12}. Remarkably, interacting bosons in a gauge field are also expected to form analogous liquid and crystalline states^{13,14,15,16,17,18,19,20,21}. However, combining interactions with strong synthetic magnetic fields has been a challenge for experiments on bosonic quantum gases^{18,21}. Here we study the purely interaction-driven dynamics of a Landau gauge Bose–Einstein condensate²² in and near the lowest Landau level. We observe a spontaneous crystallization driven by condensation of magneto-rotons^{7,10}, excitations visible as density modulations at the magnetic length. Increasing the cloud density smoothly connects this behaviour to a quantum version of the Kelvin–Helmholtz hydrodynamic instability, driven by the sheared internal flow profile of the rapidly rotating condensate. At long times the condensate self-organizes into a persistent array of droplets separated by vortex streets, which are stabilized by a balance of interactions and effective magnetic forces.

This is a preview of subscription content

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Spontaneous crystallization of an interacting Bose–Einstein condensate in an artificial magnetic field.

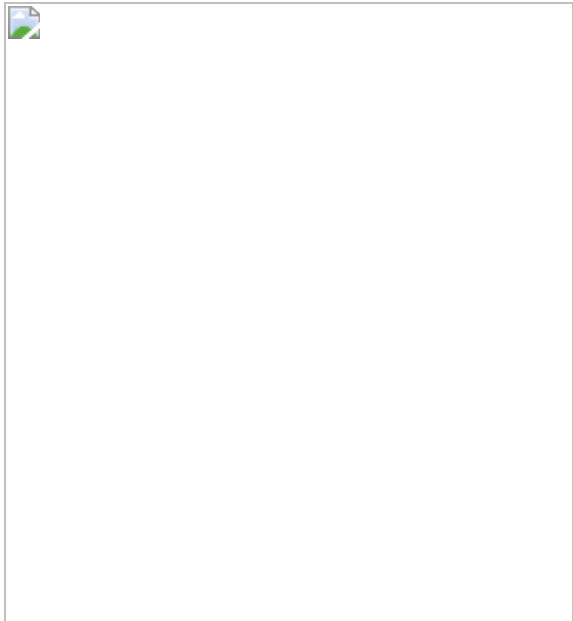


Fig. 2: Structure factor and lengthscale of the emergent crystal.

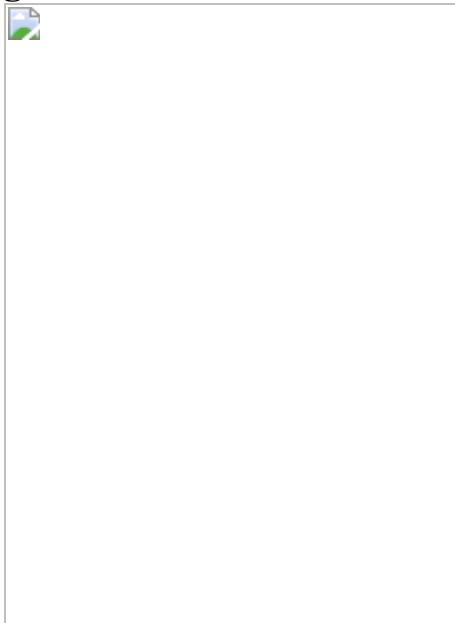


Fig. 3: Instability growth dynamics.

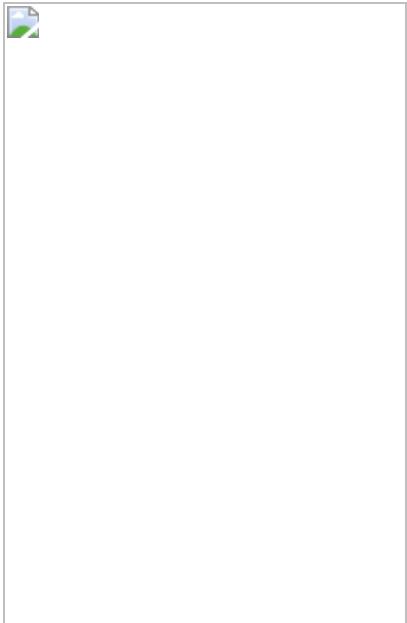
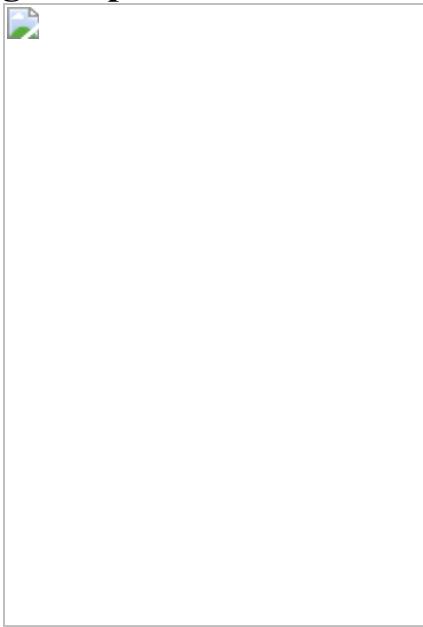


Fig. 4: Spontaneous breaking of translational symmetry.



Data availability

All data files are available from the corresponding author upon request. Accompanying data, including those for figures, are available from Zenodo (<https://doi.org/10.5281/zenodo.5533142>).

Code availability

The simulation and analysis code are available from the corresponding author upon reasonable request.

Change history

- 07 January 2022

This Article was amended to correct the Peer review information.

References

1. 1.
Stormer, H. L., Tsui, D. C. & Gossard, A. C. The fractional quantum Hall effect. *Rev. Mod. Phys.* **71**, S298–S305 (1999).
2. 2.
Bloch, I., Dalibard, J. & Zwerger, W. Many-body physics with ultracold gases. *Rev. Mod. Phys.* **80**, 885–964 (2008).
3. 3.
Cao, Y. et al. Unconventional superconductivity in magic-angle graphene superlattices. *Nature* **556**, 43–50 (2018).
4. 4.
Wigner, E. On the interaction of electrons in metals. *Phys. Rev.* **46**, 1002–1011 (1934).
5. 5.
Yoshioka, D. & Fukuyama, H. Charge density wave state of two-dimensional electrons in strong magnetic fields. *J. Phys. Soc. Jpn.* **47**, 394–402 (1979).

6. 6.

Lam, P. K. & Girvin, S. M. Liquid–solid transition and the fractional quantum-Hall effect. *Phys. Rev. B* **30**, 473–475 (1984).

7. 7.

Girvin, S. M., MacDonald, A. H. & Platzman, P. M. Magneto-roton theory of collective excitations in the fractional quantum Hall effect. *Phys. Rev. B* **33**, 2481–2494 (1986).

8. 8.

Jiang, H. W. et al. Quantum liquid versus electron solid around $\nu = 1/5$ Landau-level filling. *Phys. Rev. Lett.* **65**, 633–636 (1990).

9. 9.

Jang, J., Hunt, B. M., Pfeiffer, L. N., West, K. W. & Ashoori, R. C. Sharp tunnelling resonance from the vibrations of an electronic Wigner crystal. *Nat. Phys.* **13**, 340–344 (2017).

10. 10.

Haldane, F. D. M. & Rezayi, E. H. Finite-size studies of the incompressible state of the fractionally quantized Hall effect and its excitations. *Phys. Rev. Lett.* **54**, 237–240 (1985).

11. 11.

Pinczuk, A., Dennis, B. S., Pfeiffer, L. N. & West, K. Observation of collective excitations in the fractional quantum Hall effect. *Phys. Rev. Lett.* **70**, 3983–3986 (1993).

12. 12.

Kukushkin, I. V., Smet, J. H., Scarola, V. W., Umansky, V. & von Klitzing, K. Dispersion of the excitations of fractional quantum Hall states. *Science* **324**, 1044–1047 (2009).

13. 13.

Ho, T.-L. Bose–Einstein condensates with large number of vortices. *Phys. Rev. Lett.* **87**, 060403 (2001).

14. 14.

Oktel, M. Ö. Vortex lattice of a Bose–Einstein condensate in a rotating anisotropic trap. *Phys. Rev. A* **69**, 023618 (2004).

15. 15.

Sinha, S. & Shlyapnikov, G. V. Two-dimensional Bose–Einstein condensate under extreme rotation. *Phys. Rev. Lett.* **94**, 150401 (2005).

16. 16.

Cooper, N. R. Rapidly rotating atomic gases. *Adv. Phys.* **57**, 539–616 (2008).

17. 17.

Aftalion, A., Blanc, X. & Lerner, N. Fast rotating condensates in an asymmetric harmonic trap. *Phys. Rev. A* **79**, 011603 (2009).

18. 18.

Dalibard, J., Gerbier, F., Juzeliūnas, G. & Öhberg, P. Artificial gauge potentials for neutral atoms. *Rev. Mod. Phys.* **83**, 1523–1543 (2011).

19. 19.

Chen, X., Gu, Z.-C., Liu, Z.-X. & Wen, X.-G. Symmetry-protected topological orders in interacting bosonic systems. *Science* **338**, 1604–1606 (2012).

20. 20.

Senthil, T. & Levin, M. Integer quantum Hall effect for bosons. *Phys. Rev. Lett.* **110**, 046801 (2013).

21. 21.

Goldman, N., Juzeliūnas, G., Öhberg, P. & Spielman, I. B. Light-induced gauge fields for ultracold atoms. *Rep. Prog. Phys.* **77**, 126401 (2014).

22. 22.

Fletcher, R. J. et al. Geometric squeezing into the lowest Landau level. *Science* **372**, 1318–1322 (2021).

23. 23.

Tsui, D. C., Stormer, H. L. & Gossard, A. C. Two-dimensional magnetotransport in the extreme quantum limit. *Phys. Rev. Lett.* **48**, 1559–1562 (1982).

24. 24.

Nozières, P. Is the roton in superfluid ^4He the ghost of a Bragg spot? *J. Low Temp. Phys.* **137**, 45–67 (2004).

25. 25.

Graß, T. et al. Fractional quantum Hall phases of bosons with tunable interactions: from the Laughlin liquid to a fractional Wigner crystal. *Phys. Rev. Lett.* **121**, 253403 (2018).

26. 26.

Galitski, V. & Spielman, I. B. Spin–orbit coupling in quantum gases. *Nature* **494**, 49–54 (2013).

27. 27.

Chalopin, T. et al. Probing chiral edge dynamics and bulk topology of a synthetic Hall system. *Nat. Phys.* **16**, 1017–1021 (2020).

28. 28.

Struck, J. et al. Tunable gauge potential for neutral and spinless particles in driven optical lattices. *Phys. Rev. Lett.* **108**, 225304 (2012).

29. 29.

Jotzu, G. et al. Experimental realization of the topological Haldane model with ultracold fermions. *Nature* **515**, 237–240 (2014).

30. 30.

Aidelsburger, M. et al. Measuring the Chern number of Hofstadter bands with ultracold bosonic atoms. *Nat. Phys.* **11**, 162–166 (2015).

31. 31.

Stuhl, B. K., Lu, H.-I., Aycock, L. M., Genkina, D. & Spielman, I. B. Visualizing edge states with an atomic Bose gas in the quantum Hall regime. *Science* **349**, 1514–1518 (2015).

32. 32.

Mancini, M. et al. Observation of chiral edge states with neutral fermions in synthetic Hall ribbons. *Science* **349**, 1510–1513 (2015).

33. 33.

Schweikhard, V., Coddington, I., Engels, P., Mogendorff, V. P. & Cornell, E. A. Rapidly rotating Bose–Einstein condensates in and near the lowest Landau level. *Phys. Rev. Lett.* **92**, 040404 (2004).

34. 34.

Bretin, V., Stock, S., Seurin, Y. & Dalibard, J. Fast rotation of a Bose–Einstein condensate. *Phys. Rev. Lett.* **92**, 050403 (2004).

35. 35.

Bukov, M., D’Alessio, L. & Polkovnikov, A. Universal high-frequency behavior of periodically driven systems: from dynamical stabilization to Floquet engineering. *Adv. Phys.* **64**, 139–226 (2015).

36. 36.

Iordanskii, S. V. & Pitaevskii, L. P. Bose condensation of moving rotons. *Sov. Phys. Usp.* **23**, 317–318 (1980).

37. 37.

Pitaevskii, L. P. Layered structure of superfluid ^4He with super-critical motion. *JETP Lett.* **39**, 511–514 (1984).

38. 38.

Martone, G. I., Recati, A. & Pavloff, N. Supersolidity of cnoidal waves in an ultracold Bose gas. *Phys. Rev. Res.* **3**, 013143 (2021).

39. 39.

Mottl, R. et al. Roton-type mode softening in a quantum gas with cavity-mediated long-range interactions. *Science* **336**, 1570–1573 (2012).

40. 40.

Léonard, J., Morales, A., Zupancic, P., Esslinger, T. & Donner, T. Supersolid formation in a quantum gas breaking a continuous translational symmetry. *Nature* **543**, 87–90 (2017).

41. 41.

Ji, S.-C. et al. Softening of roton and phonon modes in a Bose–Einstein condensate with spin–orbit coupling. *Phys. Rev. Lett.* **114**, 105301 (2015).

42. 42.

Li, J.-R. et al. A stripe phase with supersolid properties in spin–orbit-coupled Bose–Einstein condensates. *Nature* **543**, 91–94 (2017).

43. 43.

Ha, L.-C., Clark, L. W., Parker, C. V., Anderson, B. M. & Chin, C. Roton-maxon excitation spectrum of Bose condensates in a shaken optical lattice. *Phys. Rev. Lett.* **114**, 055301 (2015).

44. 44.

Feng, L., Clark, L. W., Gaj, A. & Chin, C. Coherent inflationary dynamics for Bose–Einstein condensates crossing a quantum critical point. *Nat. Phys.* **14**, 269–272 (2018).

45. 45.

Zhang, Z., Yao, K.-X., Feng, L., Hu, J. & Chin, C. Pattern formation in a driven Bose–Einstein condensate. *Nat. Phys.* **16**, 652–656 (2020).

46. 46.

Petter, D. et al. Probing the roton excitation spectrum of a stable dipolar Bose gas. *Phys. Rev. Lett.* **122**, 183401 (2019).

47. 47.

Hertkorn, J. et al. Density fluctuations across the superfluid–supersolid phase transition in a dipolar quantum gas. *Phys. Rev. X* **11**, 011037 (2021).

48. 48.

Schmidt, J.-N. et al. Roton excitations in an oblate dipolar quantum gas. *Phys. Rev. Lett.* **126**, 193002 (2021).

49. 49.

Guo, M. et al. The low-energy Goldstone mode in a trapped dipolar supersolid. *Nature* **574**, 386–389 (2019).

50. 50.

Tanzi, L. et al. Supersolid symmetry breaking from compressional oscillations in a dipolar quantum gas. *Nature* **574**, 382–385 (2019).

51. 51.

Chomaz, L. et al. Long-lived and transient supersolid behaviors in dipolar quantum gases. *Phys. Rev. X* **9**, 21012 (2019).

52. 52.

Davidson, R. C., Chan, H.-W., Chen, C. & Lund, S. Equilibrium and stability properties of intense non-neutral electron flow. *Rev. Mod. Phys.* **63**, 341–374 (1991).

53. 53.

Cerfon, A. J. Vortex dynamics and shear-layer instability in high-intensity cyclotrons. *Phys. Rev. Lett.* **116**, 174801 (2016).

54. 54.

Chandrasekhar, S. C. *Hydrodynamic and Hydromagnetic Stability* (Clarendon Press, 1961).

55. 55.

Landau, L. D. & Lifshitz, E. M. *Fluid Mechanics* (Pergamon Press, 1987).

56. 56.

Finne, A. P. et al. Dynamics of vortices and interfaces in superfluid ^3He . *Rep. Prog. Phys.* **69**, 3157–3230 (2006).

57. 57.

Baggaley, A. W. & Parker, N. G. Kelvin–Helmholtz instability in a single-component atomic superfluid. *Phys. Rev. A* **97**, 053608 (2018).

58. 58.

Fetter, A. L. & Walecka, J. D. *Quantum Theory of Many-particle Systems* (McGraw-Hill, 1971).

59. 59.

Recati, A., Zambelli, F. & Stringari, S. Overcritical rotation of a trapped Bose–Einstein condensate. *Phys. Rev. Lett.* **86**, 377–380 (2001).

60. 60.

Petrich, W., Anderson, M. H., Ensher, J. R. & Cornell, E. A. Stable, tightly confining magnetic trap for evaporative cooling of neutral atoms. *Phys. Rev. Lett.* **74**, 3352–3355 (1995).

61. 61.

Sinha, S. & Castin, Y. Dynamic instability of a rotating Bose–Einstein condensate. *Phys. Rev. Lett.* **87**, 190402 (2001).

62. 62.

Ronveaux, A. (ed.) *Heun's Differential Equations* (Oxford Univ. Press, 1995).

63. 63.

Bao, W. & Wang, H. An efficient and spectrally accurate numerical method for computing dynamics of rotating Bose–Einstein condensates. *J. Comput. Phys.* **217**, 612–626 (2006).

Acknowledgements

We thank T. Pfau and his research group, J. Dunkel, A. Fetter, T. Senthil, T. Simula and W. Zwerger for discussions. This work was supported by the National Science Foundation (Center for Ultracold Atoms and grant no. PHY-2012110), Air Force Office of Scientific Research (FA9550-16-1-0324 and MURI Quantum Phases of Matter FA9550-14-1-0035), Office of Naval Research (N00014-17-1-2257), the DARPA A-Phi program through ARO grant W911NF-19-1-0511, and the Vannevar Bush Faculty Fellowship. A.S. acknowledges support from the NSF GRFP. M.Z. acknowledges funding from the Alexander von Humboldt Foundation.

Author information

Affiliations

1. MIT-Harvard Center for Ultracold Atoms, Research Laboratory of Electronics, Department of Physics, Massachusetts Institute of Technology, Cambridge, MA, USA

Biswaroop Mukherjee, Airlia Shaffer, Parth B. Patel, Zhenjie Yan, Cedric C. Wilson, Valentin Crépel, Richard J. Fletcher & Martin Zwierlein

Contributions

B.M., A.S., C.C.W., P.B.P., Z.Y. and R.J.F. constructed the apparatus. B.M., A.S., C.C.W. and R.J.F. performed the measurements. B.M. and A.S. analysed the data. B.M. contributed the GP numerical simulations. V.C., R.J.F. and M.Z. developed the theoretical description. R.J.F. and M.Z. supervised the project. All authors contributed to interpretation of the results and preparation of the manuscript.

Corresponding author

Correspondence to [Martin Zwierlein](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review information

Nature thanks Alessio Recati and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

[Extended Data Fig. 1 Numerical GP simulation of the condensate evolution in the rotating frame.](#)

a–c, Time evolution of the condensate density without the addition of noise (top), with added phase noise (middle), and in the experiment (bottom). The frames correspond to times $\omega_c t/(2\pi) = 0, 4$ and 6 . **d–e**, Evolution of the structure factor $S_k(t)$ extracted from the simulation (**d**) and the experiment (**e**) which show good agreement. **f**, The extracted instability growth rate as a function of wavevector k . The experimental measurements are shown by red points, and the Bogoliubov prediction by the black line. The blue line shows the result of the GP simulation. Here, the blue shading and the red error bars indicate 1σ standard error. This model captures the experimentally measured growth at wavevectors above the instability region provided by the linear Bogoliubov description.

[Extended Data Fig. 2 Phase profile of the crystal.](#)

a, b, The density profiles of the crystals in the experiment (**a**) and GP simulation (**b**) appear to contain vortices, which are marked in **c** and **d**. **e**, The phase of the macroscopic wavefunction can be inferred from the locations of the vortices in the experimental image. Note that additional contributions from undetected vortices may exist. **f**, The simulated phase profile from a GP simulation shows a similar structure of irrotational flow within each segment of the crystal. In both **e** and **f**, the phase shown is in the rotating frame.

Supplementary information

Supplementary Information

This Supplementary Information file contains details on our Bogoliubov stability analysis, as well as on superfluid hydrodynamics in the Landau gauge, and includes Supplementary Figures 1–3 and additional references.

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Mukherjee, B., Shaffer, A., Patel, P.B. *et al.* Crystallization of bosonic quantum Hall states in a rotating quantum gas. *Nature* **601**, 58–62 (2022). <https://doi.org/10.1038/s41586-021-04170-2>

- Received: 21 June 2021
- Accepted: 21 October 2021
- Published: 05 January 2022
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04170-2>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04170-2>

- Article
- [Published: 03 November 2021](#)

Drop-in fuels from sunlight and air

- [Remo Schäppi ORCID: orcid.org/0000-0001-8439-2141¹](#),
- [David Rutz¹](#),
- [Fabian Dähler¹](#),
- [Alexander Muroyama¹](#),
- [Philipp Haueter¹](#),
- [Johan Lilliestam ORCID: orcid.org/0000-0001-6913-5956^{2,3}](#),
- [Anthony Patt⁴](#),
- [Philipp Furler ORCID: orcid.org/0000-0001-5583-9274^{1,5}](#) &
- [Aldo Steinfeld ORCID: orcid.org/0000-0001-7797-686X¹](#)

[Nature](#) volume **601**, pages 63–68 (2022)

- 13k Accesses
- 1 Citations
- 561 Altmetric
- [Metrics details](#)

Subjects

- [Chemical engineering](#)
- [Hydrogen energy](#)
- [Mechanical engineering](#)
- [Solar fuels](#)
- [Solar thermal energy](#)

Abstract

Aviation and shipping currently contribute approximately 8% of total anthropogenic CO₂ emissions, with growth in tourism and global trade projected to increase this contribution further^{1,2,3}. Carbon-neutral transportation is feasible with electric motors powered by rechargeable batteries, but is challenging, if not impossible, for long-haul commercial travel, particularly air travel⁴. A promising solution are drop-in fuels (synthetic alternatives for petroleum-derived liquid hydrocarbon fuels such as kerosene, gasoline or diesel) made from H₂O and CO₂ by solar-driven processes^{5,6,7}. Among the many possible approaches, the thermochemical path using concentrated solar radiation as the source of high-temperature process heat offers potentially high production rates and efficiencies⁸, and can deliver truly carbon-neutral fuels if the required CO₂ is obtained directly from atmospheric air⁹. If H₂O is also extracted from air¹⁰, feedstock sourcing and fuel production can be colocated in desert regions with high solar irradiation and limited access to water resources. While individual steps of such a scheme have been implemented, here we demonstrate the operation of the entire thermochemical solar fuel production chain, from H₂O and CO₂ captured directly from ambient air to the synthesis of drop-in transportation fuels (for example, methanol and kerosene), with a modular 5 kW_{thermal} pilot-scale solar system operated under field conditions. We further identify the research and development efforts and discuss the economic viability and policies required to bring these solar fuels to market.

This is a preview of subscription content

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

Subscribe

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

Rent or Buy

All prices are NET prices.

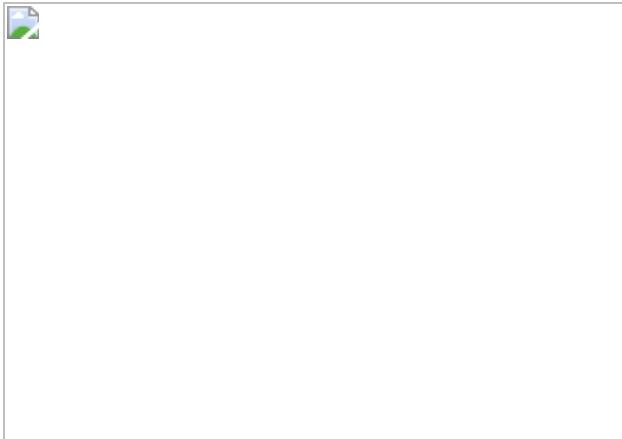
Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Simplified process chain of the solar fuel system.



Fig. 2: Representative day run of the solar redox unit for co-splitting H₂O and CO₂.



Data availability

The main data supporting the findings of this study are available within the paper and its extended data figures. Source data are available with this paper.

References

1. 1.

Gewe, V. et al. Evaluating the climate impact of aviation emission scenarios towards the Paris agreement including COVID-19 effects. *Nat. Commun.* **12**, 3841 (2021).

2. 2.

Chen, J. et al. The relationship between the development of global maritime fleets and GHG emission from shipping. *J. Environ. Manage.* **242**, 31–39 (2019).

3. 3.

World Demand by Product Groups, 2018–2019 (International Energy Agency, 2021); <https://www.iea.org/data-and-statistics/charts/world-demand-by-product-groups-2018-2019>

4. 4.

Schäfer, A. et al. Technological, economic and environmental prospects of all-electric aircraft. *Nat. Energy*. **4**, 160–166 (2019).

5. 5.

Lewis, N. S., & Nocera, D. G. Powering the planet: chemical challenges in solar energy utilization. *Proc. Natl Acad. Sci. USA*. **103**, 15729–15735 (2006).

6. 6.

Ozin, G. A. Throwing new light on the reduction of CO₂. *Adv. Mater.* **27**, 1957–1963 (2015).

7. 7.

Detz, R. J. et al. The future of solar fuels: when could they become competitive? *Energy Environ. Sci.* **11**, 1653–1669 (2018).

8. 8.

Romero, M. & Steinfield, A. Concentrating solar thermal power and thermochemical fuels. *Energy Environ. Sci.* **5**, 9234–9245 (2012).

9. 9.

Zeman, F. S., & Keith, D. W. Carbon neutral hydrocarbons. *Phil. Trans. R. Soc. A*. **366**, 3901–3918 (2008).

10. 10.

Wurzbacher, J. et al. Concurrent separation of CO₂ and H₂O from air by a temperature-vacuum swing adsorption/desorption cycle. *Environ. Sci. Technol.* **46**, 9191–9198 (2012).

11. 11.

Brady, C. et al. Integration of thermochemical water splitting with CO₂ direct air capture. *Proc. Natl Acad. Sci. USA.* **116**, 25001–25007 (2019).

12. 12.

Jia, J. et al. Solar water splitting by photovoltaic-electrolysis with a solar-to-hydrogen efficiency over 30%. *Nat. Commun.* **7**, 13237 (2016).

13. 13.

Chalmin, A. Direct air capture: recent developments and future plans. *Geoengineering Monitor*
<https://www.geoengineeringmonitor.org/2019/07/direct-air-capture-recent-developments-and-future-plans/> (2019).

14. 14.

Roeb, M. et al. Test operation of a 100 kW pilot plant for solar hydrogen production from water on a solar tower. *Sol. Energy.* **85**, 634–644 (2011).

15. 15.

Vázquez, F. V. et al. Power-to-X technology using renewable electricity and carbon dioxide from ambient air: SOLETAIR proof-of-concept and improved process concept. *J. CO₂ Util.* **28**, 235–246 (2018).

16. 16.

Chueh, W. C. & Haile, S. M. A thermochemical study of ceria: exploiting an old material for new modes of energy conversion and CO₂ mitigation. *Phil. Trans. R. Soc. A.* **368**, 3269–3294 (2010).

17. 17.

Abanades, S. & Flamant, G. Thermochemical hydrogen production from a two-step solar-driven water-splitting cycle based on cerium oxides. *Sol. Energy*. **80**, 1611–1623 (2006).

18. 18.

Madhusudhan, J. et al. Reversible phase transformations in novel Ce-substituted perovskite oxide composites for solar thermochemical redox splitting of CO₂. *Adv. Energy Mater.* **11**, 2003532 (2021).

19. 19.

McDaniel, A. H. et al. Sr- and Mn-doped LaAlO_{3-δ} for solar thermochemical H₂ and CO production. *Energy Environ. Sci.* **6**, 2424–2428 (2013).

20. 20.

Muhich, C. L. et al. Efficient generation of H₂ by splitting water with an isothermal redox cycle. *Science*. **341**, 540–542 (2013).

21. 21.

Bayon, A. et al. Operational limits of redox metal oxides performing thermochemical water splitting. *Energy Technol.*
<https://doi.org/10.1002/ente.202100222> (2021).

22. 22.

Ermanoski, I. et al. Efficiency maximization in solar-thermochemical fuel production: challenging the concept of isothermal water splitting. *Phys. Chem. Chem. Phys.* **16**, 8418–8427 (2014).

23. 23.

Abanades, S. et al., Synthesis and thermochemical redox cycling of porous ceria microspheres for renewable fuels production from solar-

aided water-splitting and CO₂ utilization. *Appl. Phys. Lett.* **119**, 023902 (2021).

24. 24.

Chueh, W. C. et al. High-flux solar-driven thermochemical dissociation of CO₂ and H₂O using nonstoichiometric ceria. *Science*. **330**, 1797–1801 (2010).

25. 25.

Marxer, D. et al. Solar thermochemical splitting of CO₂ into separate streams of CO and O₂ with high selectivity, stability, conversion, and efficiency. *Energy Environ. Sci.* **10**, 1142–1149 (2017).

26. 26.

Furler, P. et al. Thermochemical CO₂ splitting via redox cycling of ceria reticulated foam structures with dual-scale porosities. *Phys. Chem. Chem. Phys.* **16**, 10503–10511 (2014).

27. 27.

Dähler, F. et al. Optical design and experimental characterization of a solar concentrating dish system for fuel production via thermochemical redox cycles. *Sol. Energy*. **170**, 568–575 (2018).

28. 28.

Furler, P. et al. Syngas production by simultaneous splitting of H₂O and CO₂ via ceria redox reactions in a high-temperature solar reactor. *Energy Environ. Sci.* **5**, 6098–6103 (2012).

29. 29.

Schäppi, R. et al. Solar thermochemical splitting of CO₂ in a modular solar dish-reactor system. In *Proc. ISES-SWC2019 Solar World*

Congress (2019); <https://doi.org/10.18086/swc.2019.24.08>.

30. 30.

Marxer, D. et al. Demonstration of the entire production chain to renewable kerosene via solar thermochemical splitting of H₂O and CO₂. *Energy Fuels.* **29**, 3241–3250 (2015).

31. 31.

Zoller, S. *A 50 kW Solar Thermochemical Reactor for Syngas Production Utilizing Porous Ceria Structures*. Dissertation ETH No. 26451, ETH Zurich (2020).

32. 32.

Romero, R. et al. Solar-driven thermochemical production of sustainable liquid fuels from H₂O and CO₂ in a heliostat field. *Proc. ISES-SWC2019 Solar World Congress* (2019); <https://doi.org/10.18086/swc.2019.23.02>.

33. 33.

Corporan, E. et al. Emissions characteristics of a turbine engine and research combustor burning a Fischer-Tropsch jet fuel. *Energy Fuels.* **21**, 2615–2626 (2007).

34. 34.

SolarPILOT software v.2015.10.5 (National Renewable Energy Laboratory, 2015); <https://www.nrel.gov/csp/solarpilot.html>.

35. 35.

Geissbühler, L. *Thermocline Thermal Energy Storage: Advances and Applications to CSP, Compressed Air Energy Storage, and Solar Fuels*. Dissertation ETH No. 24555, ETH Zurich (2017).

36. 36.

Scheffe, J. R. & Steinfeld, A. Thermodynamic analysis of cerium-based oxides for solar thermochemical fuel production. *Energy Fuels*. **26**, 1928–1936 (2012).

37. 37.

Lapp, J. et al. Efficiency of two-step solar thermochemical non-stoichiometric redox cycles with heat recovery. *Energy*. **37**, 591–600 (2012).

38. 38.

Hoes, M. et al. Additive-manufactured ordered porous structures made of ceria for concentrating solar applications. *Energy Technol.* **7**, 1900484 (2019).

39. 39.

Wim, H. & Geerlings, H. Efficient production of solar fuel using existing large scale production technologies. *Environ. Sci. Technol.* **45**, 8609–8610 (2011).

40. 40.

Kim, J. et al. Fuel production from CO₂ using solar-thermal energy: system level analysis. *Energy Environ. Sci.* **5**, 8417–8429 (2012).

41. 41.

Falter, C. et al. Geographical potential of solar thermochemical jet fuel production. *Energies*. **13**, 802 (2020).

42. 42.

Falter, C. et al. An integrated techno-economic, environmental and social assessment of the solar thermochemical fuel pathway. *Sustain. Energy Fuels*. **4**, 3992–4002 (2020).

43. 43.

Sutherland, B. R. Pricing CO₂ direct air capture. *Joule*. **3**, 1571–1573 (2019).

44. 44.

Perner, J. & Bothe, D. *International Aspects of a Power-to-X Roadmap* (Frontier Economics, 2018).

45. 45.

Patt, A. & Lilliestam, J. The case against carbon prices. *Joule*. **2**, 2494–2498 (2018).

46. 46.

Rosenbloom, D. et al. Why carbon pricing is not sufficient to mitigate climate change – and how “sustainability transition policy” can help. *Proc. Natl Acad. Sci. USA*. **117**, 8664–8668 (2020).

47. 47.

Lilliestam, J. et al. The near- to mid-term outlook for concentrating solar power: mostly cloudy, chance of sun. *Energ. Source. B*. <https://doi.org/10.1080/15567249.2020.1773580> (2020).

48. 48.

Lilliestam, J. et al. Empirically observed learning rates for concentrating solar power and their responses to regime change. *Nat. Energy*. **2**, 17094 (2017).

49. 49.

Villasmil, W., Meier, A. & Steinfeld, A. Dynamic modeling of a solar reactor for zinc oxide thermal dissociation and experimental validation using IR thermography. *J. Sol. Energy Eng.* **136**, 010901 (2014).

50. 50.

Haring, H. W. *Industrial Gases Processing* (Wiley-VCH, 2008).

51. 51.

Haeussler, A. et al. Solar thermochemical fuel production from H₂O and CO₂ splitting via two-step redox cycling of reticulated porous ceria structures integrated in a monolithic cavity-type reactor. *Energy*. **201**, 117649 (2020).

52. 52.

Ash-Kurlander, U. et al. Impact of daily startup-shutdown conditions on the production of solar methanol over a commercial Cu-ZnO-Al₂O₃ catalyst. *Energy Technol.* **4**, 565–572 (2016).

53. 53.

Mehos, M. et al. *Concentrating Solar Power Best Practices Study* NREL/TP-5500-75763 (National Renewable Energy Laboratory, 2020); <https://www.nrel.gov/docs/fy20osti/75763.pdf>.

54. 54.

Malhotra, A. & Schmidt, T. Accelerating low-carbon innovation. *Joule*. **4**, 2259–2267 (2020).

55. 55.

Trancik, J. Renewable energy: back the renewables boom. *Nature*. **507**, 300–302 (2014).

56. 56.

Ferioli, F. et al. Use and limitations of learning curves for energy technology policy: a component-learning hypothesis. *Energy Policy*. **37**, 2525–2535 (2009).

57. 57.

Nordhaus, W. *The Climate Casino* (Yale Univ. Press, 2013).

58. 58.

Patt, A. et al. Will policies to promote energy efficiency help or hinder achieving a 1.5 °C climate target? *Energ. Effic.* **12**, 551–565 (2019).

59. 59.

Nemet, G. Beyond the learning curve: factors influencing cost reductions in photovoltaics. *Energy Policy*. **34**, 3218–3232 (2006).

60. 60.

Haas, R. et al. A historical review of promotion strategies for electricity from renewable energy sources in EU countries. *Renew. Sustain. Energy Rev.* **15**, 1003–1034 (2011).

61. 61.

Lilliestam, J. et al. The effect of carbon pricing on technological change for full decarbonization: a review of ex-post evidence. *Wiley Interdiscip. Rev. Clim. Change*. **12**, e681 (2021).

62. 62.

Eskeland, G. et al. in *Making Climate Change Work For Us: European Perspectives on Adaptation and Mitigation Strategies* (eds Hulme, M. & Neufeldt, H.) 165–199 (Cambridge Univ. Press, 2010).

63. 63.

Kunreuther, H. et al. in *Climate Change 2014: Mitigation of Climate Change* (eds Edenhofer, O. et al.) 151–205 (IPCC, Cambridge Univ. Press, 2014).

64. 64.

Battke, B. et al. Internal or external spillovers – which kind of knowledge is more likely to flow within or across technologies. *Res. Policy.* **45**, 27–41 (2016).

65. 65.

Patt, A. *Transforming Energy: Solving Climate Change With Technology Policy* (Cambridge Univ. Press, 2015).

Acknowledgements

This work was funded in part by the Swiss Federal Office of Energy (grant no. SI/501213-01), the Swiss National Science Foundation (grant no. 200021-162435) and the European Research Council under the European Union's ERC Advanced Grant (SUNFUELS, grant no. 320541) and ERC Starting Grant (TRIPOD, grant no. 715132). We thank P. Basler, T. Cooper, Y. Gracia, P. Good, G. Ambrosetti, A. Pedretti, D. Rast, M. Schmitz, N. Tzouganatos and M. Wild for their contributions to the technology development.

Author information

Affiliations

1. Department of Mechanical and Process Engineering, ETH Zurich, Zurich, Switzerland

Remo Schäppi, David Rutz, Fabian Dähler, Alexander Muroyama, Philipp Haueter, Philipp Furler & Aldo Steinfeld

2. Institute for Advanced Sustainability Studies (IASS), Potsdam, Germany

Johan Lilliestam

3. Faculty of Economics and Social Sciences, University of Potsdam, Potsdam, Germany

Johan Lilliestam

4. Department of Environmental Systems Science, ETH Zurich, Zurich,
Switzerland

Anthony Patt

5. Synhelion SA, Lugano, Switzerland

Philipp Furler

Contributions

R.S., D.R., F.D., P.H., A.M., P.F. and A.S. designed the system's components. R.S., A.M. and D.R. executed the experiments. J.L. and A.P. performed the economic/policy analyses. P.F. and A.S. managed and co-supervised the project. A.S. conceived the project idea and wrote the manuscript with input from all authors.

Corresponding authors

Correspondence to [Philipp Furler](#) or [Aldo Steinfeld](#).

Ethics declarations

Competing interests

ETH Zurich has license agreements with its spinoff companies Climeworks and Synhelion, and owns the following patents: EP 09007467.5, WO2010/091831: Gebald, C., Wurzbacher, J. & Steinfeld, A., Amine containing fibrous structure for CO₂ capture; PCT/EP2014/001082: Steinfeld, A., Scheffe, J., Furler, P., Vogt, U. & Gorbar, M., Open-cell materials for use in thermochemical fuel production processes; EP16194074, WO2018/073049: Steinfeld, A., Furler, P., Haselbacher, A. & Geissbühler, L., A thermochemical reactor system for a temperature swing cyclic process with integrated heat recovery; EP18195213.68: Ackermann, S., Dieringer, P., Furler, P., Steinfeld, A. & Bulfin, B., Process for the

production of syngas. P.F. is the CTO of Synhelion; P.F. and A.S. are shareholders of Synhelion.

Additional information

Peer review information *Nature* thanks Albert Harvey, Christian Sattler and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

[Extended Data Fig. 1 Photographs of the solar fuel system at ETH Zurich.](#)

a) The solar redox unit, comprising the primary sun-tracking solar paraboloidal concentrator coupled to a secondary planar rotating reflector, and the two solar reactors at the foci. b) The two solar reactors, water-cooled calorimeter, and Lambertian target for solar radiative power measurements (seen via the secondary reflector).

[Extended Data Fig. 2 Representative solar redox cycle producing syngas with composition suitable for methanol synthesis.](#)

a) Temporal variation of the nominal cavity temperature, total pressure, and outlet gas flow rates during a single redox cycle. b) Temporal variation of the cumulative species concentration and yield of solar syngas collected during the oxidation step. Operation conditions – During the reduction step: $Q_{\text{solar}} = 5.1 \text{ kW}$, inlet flow 0.5 L/min Ar , $T_{\text{reduction-end}} = 1450 \text{ }^{\circ}\text{C}$, total pressure $\leq 25 \text{ mbar}$. During the oxidation step: $Q_{\text{solar}} = 0 \text{ kW}$, inlet flows $0.4 \text{ L/min CO}_2 + 9.8 \text{ g/min H}_2\text{O}$, $T_{\text{oxidation-start}} = 900 \text{ }^{\circ}\text{C}$, total pressure = 1 bar.

[Source data](#)

Extended Data Fig. 3 Representative solar redox cycle producing syngas with composition suitable for FT synthesis.

a) Temporal variation of the nominal cavity temperature, total pressure, and outlet gas flow rates during a single redox cycle. b) Temporal variation of the cumulative species concentration and yield of solar syngas collected during the oxidation step. Operation conditions – During the reduction step: $Q_{\text{solar}} = 4.1 \text{ kW}$, inlet flow 0.5 L/min Ar, $T_{\text{reduction-end}} = 1450 \text{ }^{\circ}\text{C}$, total pressure $\leq 50 \text{ mbar}$. During the oxidation step: $Q_{\text{solar}} = 0 \text{ kW}$, inlet flows 0.2 L/min CO₂ + 9.8 g/min H₂O, $T_{\text{oxidation-start}} = 800 \text{ }^{\circ}\text{C}$, total pressure = 1 bar.

[Source data](#)

Extended Data Fig. 4 Syngas yield (H₂ in orange, CO in green, CO₂ in black) for each of the 152 consecutive solar redox cycles.

L denotes standard liters.

[Source data](#)

Extended Data Fig. 5

Cyclic variation (blue data points) and cumulative (black curve) molar ratio H₂:CO_x for the 152 consecutive redox cycles of Extended Data Fig. 4.

[Source data](#)

Extended Data Fig. 6 Simplified layout of a commercial-scale solar fuel plant with ten solar towers, each for 100 MW_{thermal}.

DAC: Direct Air Capture; GTL: Gas-to-Liquid.

Extended Data Table 1 Summary of syngas quality for the experimental runs of Extended Data Fig. 2 for methanol synthesis and Extended Data Fig. 3 for FT synthesis
Extended Data Table 2 Support policy instruments

Source data

[Source Data Fig. 2](#)

[Source Data Extended Data Fig. 2](#)

[Source Data Extended Data Fig. 3](#)

[Source Data Extended Data Fig. 4](#)

[Source Data Extended Data Fig. 5](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Schäppi, R., Rutz, D., Dähler, F. *et al.* Drop-in fuels from sunlight and air. *Nature* **601**, 63–68 (2022). <https://doi.org/10.1038/s41586-021-04174-y>

- Received: 17 May 2019
- Accepted: 21 October 2021
- Published: 03 November 2021
- Issue Date: 06 January 2022

- DOI: <https://doi.org/10.1038/s41586-021-04174-y>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04174-y>.

| [Section menu](#) | [Main menu](#) |

- Article
- Open Access
- [Published: 05 January 2022](#)

Depressed 660-km discontinuity caused by akimotoite–bridgmanite transition

- [Artem Chanyshев](#) [ORCID: orcid.org/0000-0002-4068-5433^{1,2}](#),
- [Takayuki Ishii](#) [ORCID: orcid.org/0000-0002-1494-2141^{2,3}](#),
- [Dmitry Bondar²](#),
- [Shrikant Bhat](#) [ORCID: orcid.org/0000-0002-1229-9842¹](#),
- [Eun Jeong Kim](#) [ORCID: orcid.org/0000-0003-0256-4785²](#),
- [Robert Farla¹](#),
- [Keisuke Nishida](#) [ORCID: orcid.org/0000-0002-8264-586X²](#),
- [Zhaodong Liu^{2,4}](#),
- [Lin Wang^{2,5}](#),
- [Ayano Nakajima⁶](#),
- [Bingmin Yan³](#),
- [Hu Tang](#) [ORCID: orcid.org/0000-0003-1571-8843³](#),
- [Zhen Chen³](#),
- [Yuji Higo⁷](#),
- [Yoshinori Tange](#) [ORCID: orcid.org/0000-0001-9579-0786¹](#) &
- [Tomoo Katsura^{2,3}](#)

[Nature](#) volume 601, pages 69–73 (2022)

- 204 Accesses
- 24 Altmetric
- [Metrics details](#)

Subjects

- [Geodynamics](#)
- [Mineralogy](#)

Abstract

The 660-kilometre seismic discontinuity is the boundary between the Earth's lower mantle and transition zone and is commonly interpreted as being due to the dissociation of ringwoodite to bridgmanite plus ferropericlase (post-spinel transition)^{1,2,3}. A distinct feature of the 660-kilometre discontinuity is its depression to 750 kilometres beneath subduction zones^{4,5,6,7,8,9,10}. However, in situ X-ray diffraction studies using multi-anvil techniques have demonstrated negative but gentle Clapeyron slopes (that is, the ratio between pressure and temperature changes) of the post-spinel transition that do not allow a significant depression^{11,12,13}. On the other hand, conventional high-pressure experiments face difficulties in accurate phase identification due to inevitable pressure changes during heating and the persistent presence of metastable phases^{1,3}. Here we determine the post-spinel and akimotoite–bridgmanite transition boundaries by multi-anvil experiments using in situ X-ray diffraction, with the boundaries strictly based on the definition of phase equilibrium. The post-spinel boundary has almost no temperature dependence, whereas the akimotoite–bridgmanite transition has a very steep negative boundary slope at temperatures lower than ambient mantle geotherms. The large depressions of the 660-kilometre discontinuity in cold subduction zones are thus interpreted as the akimotoite–bridgmanite transition. The steep negative boundary of the akimotoite–bridgmanite transition will cause slab stagnation (a stalling of the slab's descent) due to significant upward buoyancy^{14,15}.

[Download PDF](#)

Main

The 660-km discontinuity (D660) is one of the most important structural boundaries in the Earth's mantle, which plays an essential role in mantle dynamics because it determines whether slabs stagnate or penetrate into the lower mantle¹⁴. Seismological observations reveal significant depressions of the D660 down to 750 km, as well as a D660 multiple-discontinuity structure in subduction zones^{4,5,6,7,8,9,10}. Such observations should be related to the phase transitions of mantle minerals. Because slabs are generally colder than the ambient mantle, the temperature dependence of phase transitions (the Clapeyron slope in the pressure–temperature (P – T) phase diagram) is essential to interpret such depressions and understand subduction dynamics.

The D660 is usually attributed to the dissociation of $(\text{Mg},\text{Fe})_2\text{SiO}_4$ ringwoodite (Rw) into $(\text{Mg},\text{Fe})\text{SiO}_3$ bridgmanite (Brg) and $(\text{Mg},\text{Fe})\text{O}$ ferropericlase (fPc) (the RBP transition, hereafter)^{1,2,3}. Recent in situ X-ray diffraction studies using a multi-anvil

apparatus showed that this reaction has negative but gentle Clapeyron slopes (-1.3 MPa K^{-1} to -0.5 MPa K^{-1}), which can only vary the D660 depth between 630 and 670 km depth^{11,12,13,16}. An alternative explanation of the D660 depressions in subduction zones is therefore needed. Currently, the following three hypotheses are being discussed. One is steepening of the RBP phase boundary by water. One study¹² suggested that the addition of 2 wt% of water produces a phase boundary with a slope of -3.1 MPa K^{-1} to -3.2 MPa K^{-1} . However, this suggestion could be invalid because of improper experimental procedure; the formation of ringwoodite from bridgmanite plus ferropericlase was not demonstrated under wet conditions in this study. The second explanation is the persistence of metastable Rw in the Brg + fPc stability field owing to sluggish kinetics at low temperatures^{17,18}. One seismological study demonstrated that the D660 is extremely sharp and less than 2 km thick in cold subduction zones², and such sharpness can be interpreted as being due to the overpressure-induced RBP transition only at extremely low temperatures ($<1,000 \text{ K}$)¹⁸. Although this explanation is possible, there has been no determination of the transition kinetics at such low temperatures to provide a definite answer. The third is that the depressed D660 is caused by another phase transition with a steeper Clapeyron slope. The most likely candidate is the akimotoite (Ak)–Brg transition (AB transition, hereafter), as previously proposed¹⁹. High-pressure and high-temperature in situ X-ray studies have shown that this transition occurs at a similar pressure to the RBP transition, but has a steeper slope (-3.2 MPa K^{-1})^{20,21,22}. Hence, the AB transition could be responsible for the D660 depressions beneath subduction zones.

Combining the multi-anvil technique with in situ X-ray diffraction produces the most reliable high P – T phase relations data because of the well-controlled P – T conditions²³. Nevertheless, most previous experiments using this combination investigated phase stability based only on the formation of new phases from a starting material that is stable under ambient conditions. Such data indicate that the newly formed phases are more stable than the starting material but do not provide any information about phase equilibrium between high-pressure phases. The definition of phase equilibrium is the balance between the forward reaction (lower-pressure phase to higher-pressure phase) and reverse reaction (higher- to lower-pressure phase). The stability of the lower- and higher-pressure phases must therefore be determined by bracketing based on the results of a pair of forward and reverse reactions. Furthermore, the kinetics of phase transitions of mantle minerals becomes extremely sluggish once a high-pressure phase is formed^{13,22}. Our recent studies^{1,3} indicated that conventional high P – T experiments could have misinterpreted phase relations owing to sudden and large changes in sample pressure upon heating and a sluggish transformation of a newly formed phase. It is thus necessary to re-investigate the previously determined phase relations using a strategy that considers the above issues.

In this study, we determined the boundaries of the RBP and AB phase transitions in the MgO–SiO₂ systems over a temperature range of 1,250–2,085 K by using advanced multi-anvil techniques with *in situ* X-ray diffraction. The detailed experimental procedure is described in the Methods. The crucial point of our strategy is that the phase stability was determined from the relative change of the diffraction peak intensities of the coexisting lower-pressure and higher-pressure phases at a fixed press load and temperature to observe the growth of a stable phase. This method enables us to correctly determine the reaction direction by removing the kinetic effects and to avoid sudden pressure changes upon heating or cooling³ (Fig. 1). Although the determination of phase stability by comparing *in situ* X-ray diffraction intensities at a fixed temperatures was previously done¹⁶, that study lacked data on the paired forward and reverse reactions at the same temperatures and did not consider the effect of kinetics.

Fig. 1: Accurate phase identification by means of *in situ* X-ray diffraction in a multi-anvil press.

 **figure1**

a, An example of the change of intensity ratio between ringwoodite and bridgmanite plus periclase at 2,038 K and 23.31(2) GPa. **b**, An example of the change of intensity ratio between akimotoite and bridgmanite at 1,448 K and 22.14(7) GPa. The first diffraction patterns are shown in **(a)** blue and **(b)** pink, and the second patterns are

displayed in (a) orange and (b) dark grey. Magnified areas of the diffraction patterns are shown on the right. The numbers above the peaks indicate the Miller indexes of bridgmanite, ringwoodite, akimotoite and stishovite. The upward and downward arrows indicate peaks with increased and decreased intensities in the second diffraction patterns, respectively. Because the intensities of most peaks of the low-pressure phases increase in the second diffraction patterns, the stable phases were defined as (a) ringwoodite at 23.29(3) GPa and 2,038 K, and (b) akimotoite at 22.11(6) GPa and 1,448 K. In b, the fluorescence lines of Pb K α and K β are shown by the Siegbahn notation. Av., average.

Source data

The RBP boundary has a slightly concave curve, whereas the AB boundary has a steep convex curve (Fig. 2). The RBP boundary is located at pressures of 23.2–23.7 GPa in the temperature range of 1,250–2,040 K. Its slope varies from -0.1 MPa K^{-1} at temperatures less than 1,700 K to -0.9 MPa K^{-1} at 2,000 K with an averaged value of -0.5 MPa K^{-1} (Fig. 2). The slope of the AB boundary gradually changes from -8.1 MPa K^{-1} at low temperatures up to 1,300 K to -3.2 MPa K^{-1} above 1,600 K. The AB boundary is located at higher pressure than the RBP boundary at $T < 1,260 \text{ K}$ (based on linear extrapolation), where $P = 23.8 \text{ GPa}$ (Fig. 3). In this temperature range, RW should dissociate not into Brg but Ak plus periclase (Pc). The AB boundary will be located at $P = 26.8 \text{ GPa}$ and $T = 900 \text{ K}$ by linear extrapolation of our data below 1,350 K.

Fig. 2: Phase relations of bridgmanite-forming transitions in MgO-SiO₂ systems.

 figure2

a, RBP transition in the Mg_2SiO_4 system. Green diamonds and purple triangles indicate the P - T conditions at which the ratio of bridgmanite plus periclase to ringwoodite increases or decreases, respectively. The grey shaded area indicates an allowed region for the RBP boundary. The dot-dashed curve represents the most

probable transition curve within this area. **b**, AB transition in the MgSiO_3 system. Red squares and blue circles indicate the P - T conditions at which the ratio of bridgmanite to akimotoite increases or decreases, respectively. The violet shaded area indicates an allowed region for the AB boundary for the vast majority of experimental data points. The solid violet curve represents the most probable transition curve within this area. The probable reason for variation of the AB phase transition boundary is given in the text. Pressures were determined from the MgO unit-cell volumes using the Birch–Murnaghan and Vinet equations of state from ref. ⁴⁷. Error bars originate from the pressure uncertainty of the MgO equation of state suggested from ref. ⁴⁷. The temperatures and pressures were corrected based on the pressure effects on the thermoelectromotive force of a W_{97}Re_3 – $\text{W}_{75}\text{Re}_{25}$ thermocouple⁴⁸.

[Source data](#)

Fig. 3: Comparison of the RBP and AB phase transition boundaries determined in the present study.



Cold and warm subduction geotherms are from ref. [36,37](#) and the average mantle geotherm is after ref. [49](#). RBP, dot-dashed black curve and grey shaded area; AB, solid violet curve and violet shaded area. The solid orange line is 660 km depth^{[50](#)}. Ak, MgSiO₃ akimotoite; Brg, MgSiO₃ bridgmanite; Pc, MgO periclase; Rw, Mg₂SiO₄ ringwoodite. Dotted lines are the result of linear extrapolation of the RBP boundary below 1,500 K and the AB boundary below 1,290 K.

The high precision of pressure and temperature determination, high density of experimental data points and application of the advanced techniques to avoid kinetic problems allowed us to determine phase boundaries far more accurately than in previous studies. The majority of previous studies after 2003 on the RBP boundary yielded similar slopes but a lower-pressure location than the present study (Extended Data Fig. [1a](#)). The lower-pressure boundaries^{[12,13,16](#)} were obtained using gold pressure scales, whereas the previous study that used a MgO scale^{[11](#)}, which is also adopted in this study, was located at similar pressure. The quench (ex situ) experiments^{[2,21](#)} exhibited steep slopes and higher pressures. The previous experimentally determined AB boundaries^{[2,20,22,24,25](#)} yielded similar slopes of the high-temperature side of the present boundary (Extended Data Fig. [1b](#)), whereas the steep slope in the low-temperature side has not been investigated previously. Earlier experimental studies were not able to overcome the sluggish kinetics of the AB transition at low temperature and obtain the correct phase boundary, as explained above. Theoretical studies predicted steeper slopes of the AB boundary, $-3.5 \pm 0.8 \text{ MPa K}^{-1}$ (ref. [26](#)) and $-6.0 \pm 1.0 \text{ MPa K}^{-1}$ (ref. [27](#)), than those determined experimentally. Thus, our results are in better agreement with theoretical data than are previous experimental studies.

Curvature of the AB transition boundary

We find that the Clapeyron slope of the AB phase transition boundary gradually varies with temperature. The slope of a phase boundary is equal to the ratio of the entropy change (ΔS_{tr}) to the volume change (ΔV_{tr}) associated with the phase transition according to the Clausius–Clapeyron relation: $dP/dT = \Delta S_{\text{tr}}/\Delta V_{\text{tr}}$. The volume changes of the AB transition vary only slightly over the entire investigated temperature range because both phases have similar thermal expansion coefficients and bulk moduli (Extended Data Table [1](#)). The entropy is a function of isobaric heat capacity (C_P) and temperature $\Delta S_{\text{tr}} = \int_0^T \frac{C_P}{T} dT$. Calorimetric studies have reported Ak and Brg heat capacities only up to 700 K (ref. [28](#)) and 295 K (ref. [29](#)), respectively, owing to the limited thermal stability of these minerals at ambient pressure. A few theoretical studies calculated their heat capacities^{[26,30](#)}, but according to these calculations the AB boundary should have a relatively constant Clapeyron slope of $-3.5 \pm 0.8 \text{ MPa K}^{-1}$ (ref. [26](#)), which

contradicts our results. Thus, further explanation is required. The curved AB boundary can be explained by different curvatures of Ak and Brg heat capacities with temperature (Extended Data Fig. 2), as suggested by one computational study³⁰. We estimate that the Ak heat capacity should be lower than that of Brg at $T < 860$ K and higher at higher temperatures (Extended Data Fig. 2). This difference in curvature could result from temperature-dependent cation disorder in Ak. Although possible cation disorder in Ak has been previously suggested^{31,32,33}, no X-ray diffraction measurement has experimentally shown the Mg–Si disorder population³⁴. Nevertheless, the measured distances of Mg–O (2.077 Å; ref. 34) and Si–O (1.799 Å; ref. 34) in Ak and Shannon’s ionic radii of Mg (0.72 Å; ref. 35), Si (0.4 Å; ref. 35) and O (1.4 Å; ref. 35) imply that Si may occupy up to 15% of Mg sites, leading to cation disorder. The effect of cation disorder on the heat capacity in ilmenite-structured minerals has not been investigated, but disorder causes an increase in entropy, leading to a curved phase boundary. Detailed studies regarding Ak cation disorder at high P – T are required to understand the variation of the Clapeyron slope of the AB boundary.

Topography of the D660

The present results predict the topography of the D660, especially its depression, beneath subduction zones. At $T < 1,260$ K, which is appropriate for cold subduction zones, the AB boundary is located at higher pressures than the extension of the RBP boundary (Fig. 3). Under this circumstance, Rw should first dissociate not to Brg plus fPc but to Ak plus fPc (the RAP transition), and Ak should then transform to Brg at higher pressures. In other words, the RBP transition should not be present in cold subduction zones, and the AB transition should cause the depressed D660, as proposed in previous theoretical^{26,27} and seismological¹⁹ studies. If the temperature is as low as 900 K, which some geodynamic studies^{36,37} have predicted for cold subducted slabs, the D660 will be depressed to a depth of 740 km. Thus, the combination of the RAP and AB transitions will produce a double D660 at 660–740 km depths. The seismic wave velocity contrasts associated with these transitions are 40% and 60% from those associated with the RBP transition, respectively, as shown in Extended Data Fig. 3, which should be sufficiently large to detect each of them separately by seismological observations. At $T > 1,260$ K, which is appropriate for warm subduction zones, the RBP boundary should cause the D660 at approximately 660 km depth (Fig. 3).

The seismological observations of cold and warm slabs support the above predictions. One example is the very old (110 Myr³⁷) Tonga slab, for which the calculated temperature at depths of 660–740 km is 870–1,270 K in the central part of the slab³⁷. Seismological observations revealed a double D660 at depths of 738 km and 669 km for this slab¹⁰, which can be interpreted as the AB and RAP transitions, respectively. A contrasting example is the Peru slab, which is younger (41 Myr; ref. 37) and therefore

warmer. The calculated temperature of the Peru slab at approximately 660 km depth is 1,270–1,570 K in the central part of the slab³⁷, and seismological observations revealed the D660 at 656 km (ref. ¹⁵), which can be interpreted as the RBP transition. Thus, the D660 depths calculated based on the AB, RAP and RBP boundaries and expected temperatures agree with the seismological observations beneath subduction zones. The D660 is globally observed near 660 km depth except for subduction zones, which should correspond to the RBP transition located at 23.4–23.8 GPa. We note that a depression of the D660 down to 690 km depth is observed in a hot plume³⁸, which neither the RBP or AB transition can explain. This observation can be interpreted as being caused by another phase transition, probably the garnet (Grt)–Brg phase transition²⁰. Further study of this phase transition following our strategy is required to understand the origin of the D660 depression in the hot plume.

Influence of Al and Fe

Although we did not investigate the effects of Al and Fe on the phase boundaries, we can predict their effects on the geophysical applications of our results. Because Ak and Brg form by the dissociation of Rw, they will contain no Al. When no Al is present, Fe strongly partitions into fPc, but neither Ak³⁹ nor Brg⁴⁰. Our previous study also demonstrated that the pressures of Fe-free and Fe-bearing RBP transitions at 1,700 K are almost identical¹. Although the Ak formed from Rw in cold slabs could react with Al-rich Grt to incorporate Al into Ak during further slab sinking (at 24–27 GPa and 1,100–1,250 K), elemental diffusivity should be too low at such temperatures to allow compositional equilibrium in Grt and Ak. Thus, the AB transition in cold slabs should essentially occur in the MgO–SiO₂ system, as is the case in our study.

Geodynamic implications

Several geodynamics studies have shown that steep phase boundaries produce large buoyancy forces^{41,42}. When the upward buoyancy produced by a negative phase boundary exceeds the downward force due to thermal expansion, mantle flows are hampered and slabs stagnate above the D660⁴³. The boundary values of the Clapeyron slope, below which the slab will stagnate above the D660, were calculated as -6 MPa K^{-1} (ref. ⁴²) and -3 MPa K^{-1} (ref. ⁴⁴) in different geodynamics studies. Seismic tomography has revealed that warm young slabs (for example beneath Peru, the Marianas and Central America) penetrate the lower mantle^{15,45}, whereas cold old slabs (for example beneath the Izu–Bonin region, South Kurile and Japan) stagnate above D660⁴⁶. The extremely steep negative Clapeyron slope of the AB phase transition at low temperatures (-8.1 MPa K^{-1}) will cause strong upward buoyancy and stagnation of cold subducted slabs, although the nearly neutral Clapeyron slope of the RBP phase

transition at $T > 1,260$ K (-0.1 MPa K $^{-1}$) will allow the downward movement of warm slabs by thermal expansion.

Methods

Preparation of starting materials

MgSiO₃ bridgmanite and a mixture of MgSiO₃ bridgmanite and MgO periclase were used as starting materials for the determination of the AB and RBP phase transition boundaries, respectively. The starting material of bridgmanite was synthesized from enstatite powder at a pressure of 27 GPa and temperature of 1,300 K for 1 h. The mixture of bridgmanite + periclase was prepared using forsterite powder at a pressure of 27 GPa and temperature of 1,200 K for 5 min. The starting powders of enstatite and forsterite prepared by ref. ³ were used. High-pressure experiments to prepare the starting materials were conducted using a 15-MN Kawai-type multi-anvil press with the Osugi module at the Bayerisches Geoinstitut (BGI), University of Bayreuth (IRIS-15)^{1,51}. Tungsten carbide (WC) anvils of grade TF05 (Fujilloy Co. Ltd) with a truncated edge length of 3.0 mm were used to compress the high-pressure cell. Pyrophyllite gaskets were used to seal the compressed volume and generate quasi-hydrostatic pressures inside the cell. A Cr-doped MgO 7-mm octahedron was used as a pressure medium. A Mo foil with a 25-μm thickness was used as a sample container and heating element. The furnace was surrounded by a ZrO₂ thermal insulator. Mo rods were placed at both ends of the heater to connect the heater/capsule with the WC anvils. The temperature was monitored on the surface of the heater with a W₉₇Re₃–W₇₅Re₂₅ thermocouple.

Reagent-grade MgO was used as a pressure marker for the in situ X-ray diffraction experiments and was sintered at 2 GPa and 770 K for 1 h using the 10-MN Kawai type multi-anvil press at BGI (Hymag). Tungsten carbide anvils with 15-mm truncated edge lengths were used to generate high pressure together with a Cr-doped MgO octahedron with a 25-mm edge length as the pressure medium. The inside of the pressure medium consisted of a stepped cylindrical graphite heater and ZrO₂ thermal insulator. A Mo foil (25-μm) was used as the sample capsule. The temperature was measured on the surface of the capsule using a W₉₇Re₃–W₇₅Re₂₅ thermocouple.

The recovered samples were analysed using a micro-focused X-ray diffractometer (Bruker AXS Discover 8) with a two-dimensional solid-state detector (VANTEC500) and micro-focus source (I μ S) with Co-K α radiation operated at 40 kV and 500 μA. We found a small amount of stishovite together with bridgmanite (Extended Data Fig. 4), which was also observed later in the diffraction spectra at high pressures and temperatures (Fig. 1b). A small amount of stishovite can be explained as excess SiO₂.

cristobalite in the starting material (enstatite powder). The synthesized samples and sintered MgO were cut into half-disks of 0.8 mm in diameter and 0.5 mm in thickness.

In situ X-ray diffraction experiments

The AB and RBP phase transition boundaries were determined at the DESY synchrotron radiation facility (Hamburg, Germany) and SPring-8 synchrotron radiation facility (Hyogo Prefecture, Japan), respectively. The experiments were performed at beamline P61b at DESY using the 3×5 -MN six-axis multi-anvil press, and at beamline BL04B1 at SPring-8 using the 15-MN Kawai-type multi-anvil press, SPEED-*Mk.II*. The press loads of the six-axis press were the press loads of one axis multiplied by three, which corresponds to the press load of uniaxial presses such as SPEED-*Mk.II*. X-ray diffraction patterns were collected for 150–300 s for the pressure marker and 600–7,200 s for the sample using a Ge solid-state detector (SSD) with a 4,096-channel analyser. The SSD analyser was calibrated using the X-ray fluorescence lines of ^{55}Fe , ^{57}Co and ^{133}Ba before the measurements. The diffraction angle (2θ) was calibrated before each experiment with a precision of 0.0003° using MgO as a standard.

We used the same type of high-pressure cell described in ref. [1](#). The high-pressure cell contained a Cr_2O_3 –MgO pressure medium, cylindrical LaCrO_3 heater, ZrO_2 thermal insulator sleeve and Ta electrodes. The sample and pressure marker with a half-disk shape were located in a 25-μm-thick Re cylindrical foil in the center of the experimental cell. Diamond/epoxy and boron/epoxy rods were put on both sides of the sample to minimize X-ray absorption. The sample temperature was measured on the surface of the Re capsule using a W_{97}Re_3 – $\text{W}_{75}\text{Re}_{25}$ thermocouple. The thermocouple was isolated from the heater by Al_2O_3 insulator tubes. The pressure effect of the thermoelectromotive force of the thermocouple was corrected using the equations determined by ref. [48](#) after the experiments (Supplementary Tables [1](#) and [2](#)).

The incident X-ray beam collimated to dimensions of 30–50 μm horizontally and 200–300 μm vertically was directed at the sample through the gaps between the second-stage anvils. All of the experiments were carried out with a press oscillation around the vertical press axis between 0° and 6° during the X-ray diffraction measurement to suppress intensity heterogeneities of the diffracted peaks due to possible grain growth at high temperature [52](#). The pressure was obtained from the MgO unit cell volumes using the equations of state proposed by ref. [47](#) based on the third-order Birch–Murnaghan (3BM) and Vinet equations of state. To calculate the MgO unit-cell volumes, we usually used eight diffraction peaks (111, 200, 220, 311, 222, 400, 420, 422), which produce relatively high precision in pressure [1,3](#).

Although the experiments were carried out at different synchrotron facilities, the applied techniques were identical, with minor differences, which are described in detail in the next section (the AB boundary and the RBP boundary) and did not affect the overall result of the study.

Determination of phase boundaries

As argued in refs. [1.3](#), the procedure to determine the phase boundaries in this study is based on the following ideas.

1. 1.

The starting materials are very reactive owing to the high-density defects produced during cold compression. The high-density defects can be inferred from the peak broadening during compression.

2. 2.

A phase that forms *in situ* is much more inert than the starting material due to low-defect density, which is inferred from the sharper diffraction peaks.

3. 3.

The phase that forms *in situ* is still reactive unless it is annealed at high temperatures for a long duration. This is likely because of the small grain size, which is inferred from the sharpening of diffraction peaks with increasing temperature.

4. 4.

It is difficult to form another new phase from a phase that forms *in situ*, probably because the surface energy hampers the formation of new grains. On the other hand, it is relatively easy to allow one phase to increase by consuming the other phase if the two phases are already present.

5. 5.

Reaction rates decrease when approaching the phase boundaries. It is therefore possible to maintain the coexistence of higher- and lower-pressure phases if the sample P - T conditions are held close to the phase boundaries.

6. 6.

The reaction rate from one phase to the other decreases by annealing at high temperatures. This is likely because, although the transition from one phase to the other occurs through grain-boundary diffusion, the grain-boundary density should decrease due to grain growth by annealing.

For these reasons, we achieved coexistence of the higher- and lower-pressure phases at as low a temperature as possible at the beginning of the experiment. The sample temperature was then increased stepwise, and the phase boundaries were bracketed by small proceedings of the forward and reverse reactions in which the sample pressure was held near the phase boundary.

The AB boundary

The starting materials of bridgmanite were initially compressed to press loads of 10–12 MN at room temperature, at which the sample pressures were 30–32 GPa. They were then gradually heated to temperatures of 1,000–1,200 K, during which the sample pressure spontaneously decreased to 24–26 GPa. At temperatures of 1,100–1,150 K, the starting materials of pure bridgmanite partially transformed into akimotoite. The temperature and press load were then maintained and diffraction patterns of the samples and pressure markers were successively collected. In many cases, the sample pressure spontaneously decreased. A stable phase was determined based on the relative increase or decrease of the diffraction peak intensities of the higher-pressure (Brg) to lower-pressure (Ak) phases. We typically used 6 Ak peaks (012, 104, 110, 113, 024, 116) and 15 Brg peaks (002, 110, 111, 020, 112, 200, 120, 210, 022, 202, 113, 122, 212, 220, 023) to estimate the diffraction peak intensities change. The change in the intensities of such a large number of diffraction peaks is associated with a quantitative fraction volume change of the coexisting phases. On the other hand, using only one or few peaks could be easily associated with grain growth or preferred orientation development, and therefore the reliability in such cases would be questionable. By crossing the phase transition boundary, the Brg/Ak ratio initially increased but then decreased. After the temperature and pressure stabilized, we considered the samples to be in the Brg stability field if the Brg/Ak ratio increased, but in the Ak stability field if the ratio decreased. If the spontaneous pressure decrease seemingly terminated, the press load was reduced slightly (by 0.05–0.15 MN) while maintaining the temperature to enhance the pressure drop until the Brg/Ak intensity ratio started to decrease. The lowest pressure at which the Brg/Ak ratio increased and highest pressure at which the ratio decreased were used to obtain the transition pressures. After the transition pressures were determined at this temperature, the sample temperature was increased by 50–100 K, and the above procedure was repeated to bracket the transition pressures at the higher temperature. If the Brg/Ak ratio decreased at the beginning of a new temperature stage, the sample temperature was reduced by 50–100 K, the press load was increased by 0.5–1.5 MN, and the

sample was heated to the original temperature to enter the Brg stability field. After the samples were exposed at 2,085 K (2,000 K before correction), they were quenched and decompressed to ambient pressure.

The RBP boundary

The RBP boundary was determined in a similar way to the AB boundary. The procedural differences between these two boundaries are as follows. The mixtures of bridgmanite and periclase (starting material) were partially converted to ringwoodite at 1,200 K. The stable phase at high P - T conditions was estimated mainly by the relative intensities of bridgmanite and ringwoodite diffraction peaks due to the relatively small proportion of periclase (Fig. 1a). We usually used 7 Rw peaks (220, 311, 222, 400, 422, 511, 333) and 18 Brg peaks (002, 110, 111, 020, 112, 120, 210, 121, 202, 113, 122, 212, 220, 023, 221, 130, 131, 311) to estimate the diffraction peak intensity changes. Because the total peak number is even larger than that used for the AB transition determination, the intensity changes of the diffraction peaks are also associated with a quantitative fraction volume change of the coexisting phases. The maximum experimental temperature was 2,040 K (1,950 K before correction).

Data availability

The X-ray diffraction data that support the findings of this study are available on Zenodo (<https://doi.org/10.5281/zenodo.5532573>). [Source data](#) are provided with this paper.

References

1. 1.
Ishii, T. et al. Sharp 660-km discontinuity controlled by extremely narrow binary post-spinel transition. *Nat. Geosci.* **12**, 869–872 (2019).
2. 2.
Ito, E. & Takahashi, E. Postspinel transformations in the system Mg_2SiO_4 – Fe_2SiO_4 and some geophysical implications. *J. Geophys. Res. Solid Earth* **94**, 10637–10646 (1989).
3. 3.

Ishii, T. et al. Complete agreement of the post-spinel transition with the 660-km seismic discontinuity. *Sci. Rep.* **8**, 6358 (2018).

4. 4.

Ai, Y., Zheng, T., Xu, W., He, Y. & Dong, D. A complex 660 km discontinuity beneath northeast China. *Earth Planet. Sci. Lett.* **212**, 63–71 (2003).

5. 5.

Chen, L. & Ai, Y. Discontinuity structure of the mantle transition zone beneath the North China Craton from receiver function migration. *J. Geophys. Res. Solid Earth* <https://doi.org/10.1029/2008JB006221> (2009).

6. 6.

Deuss, A., Redfern, S. A., Chambers, K. & Woodhouse, J. H. The nature of the 660-kilometer discontinuity in Earth's mantle from global seismic observations of PP precursors. *Science* **311**, 198–201 (2006).

7. 7.

Gao, Y. et al. Seismic discontinuities in the mantle transition zone and at the top of the lower mantle beneath eastern China and Korea: influence of the stagnant Pacific slab. *Phys. Earth Planet. Inter.* **183**, 288–295 (2010).

8. 8.

Schultz, R. & Gu, Y. J. Multiresolution imaging of mantle reflectivity structure using SS and P'P' precursors. *Geophys. J. Int.* **195**, 668–683 (2013).

9. 9.

Tibi, R., Wiens, D. A., Shiobara, H., Sugioka, H. & Yuan, X. Double seismic discontinuities at the base of the mantle transition zone near the Mariana slab. *Geophys. Res. Lett.* <https://doi.org/10.1029/2007GL030527> (2007).

10. 10.

Zang, S. X., Zhou, Y. Z., Ning, J. Y. & Wei, R. Q. Multiple discontinuities near 660 km beneath Tonga area. *Geophys. Res. Lett.* **33**, L20312 (2006).

11. 11.

Fei, Y. et al. Experimentally determined postspinel transformation boundary in Mg_2SiO_4 using MgO as an internal pressure standard and its geophysical implications. *J. Geophys. Res. Solid Earth* <https://doi.org/10.1029/2003JB002562> (2004).

12. 12.

Ghosh, S. et al. Effect of water in depleted mantle on post-spinel transition and implication for 660 km seismic discontinuity. *Earth Planet. Sci. Lett.* **371**, 103–111 (2013).

13. 13.

Katsura, T. et al. Post-spinel transition in Mg_2SiO_4 determined by high $P-T$ *in situ* X-ray diffractometry. *Phys. Earth Planet. Inter.* **136**, 11–24 (2003).

14. 14.

Fukao, Y., Obayashi, M., Nakakuki, T. & Group, D. S. P. Stagnant slab: a review. *Annu. Rev. Earth Planet. Sci.* **37**, 19–46 (2009).

15. 15.

Fukao, Y. & Obayashi, M. Subducted slabs stagnant above, penetrating through, and trapped below the 660 km discontinuity. *J. Geophys. Res. Solid Earth* **118**, 5920–5938 (2013).

16. 16.

Litasov, K., Ohtani, E., Sano, A., Suzuki, A. & Funakoshi, K. *In situ* X-ray diffraction study of post-spinel transformation in a peridotite mantle: implication for the 660-km discontinuity. *Earth Planet. Sci. Lett.* **238**, 311–328 (2005).

17. 17.

Tetzlaff, M. & Schmeling, H. The influence of olivine metastability on deep subduction of oceanic lithosphere. *Phys. Earth Planet. Inter.* **120**, 29–38 (2000).

18. 18.

Kubo, T. et al. Mechanisms and kinetics of the post-spinel transformation in Mg_2SiO_4 . *Phys. Earth Planet. Inter.* **129**, 153–171 (2002).

19. 19.

Cottaar, S. & Deuss, A. Large-scale mantle discontinuity topography beneath Europe: signature of akimotoite in subducting slabs. *J. Geophys. Res. Solid Earth* **121**, 279–292 (2016).

20. 20.

Hirose, K., Komabayashi, T., Murakami, M. & Funakoshi, K. I. In situ measurements of the majorite–akimotoite–perovskite phase transition boundaries in MgSiO_3 . *Geophys. Res. Lett.* **28**, 4351–4354 (2001).

21. 21.

Ishii, T., Kojitani, H. & Akaogi, M. Post-spinel transitions in pyrolite and Mg_2SiO_4 and akimotoite-perovskite transition in MgSiO_3 : precise comparison by high-pressure high-temperature experiments with multi-sample cell technique. *Earth Planet. Sci. Lett.* **309**, 185–197 (2011).

22. 22.

Ono, S. et al. In situ observation of ilmenite–perovskite phase transition in MgSiO_3 using synchrotron radiation. *Geophys. Res. Lett.* **28**, 835–838 (2001).

23. 23.

Katsura, T. Phase-relation studies of mantle minerals by in situ X-ray diffraction using multianvil apparatus. *Spec. Pap. Geol. Soc. Am.* **421**, 189–205 (2007).

24. 24.

Kato, T. et al. In situ X ray observation of high-pressure phase transitions of MgSiO_3 and thermal expansion of MgSiO_3 perovskite at 25 GPa by double-stage multianvil system. *J. Geophys. Res. Solid Earth* **100**, 20475–20481 (1995).

25. 25.

Kulka, B. L., Dolinschi, J. D., Leinenweber, K. D., Prakapenka, V. B. & Shim, S.-H. The bridgmanite–akimotoite–majorite triple point determined in large volume press and laser-heated diamond anvil cell. *Minerals* **10**, 67 (2020).

26. 26.

Hernández, E. R., Brodholt, J. & Alfè, D. Structural, vibrational and thermodynamic properties of Mg_2SiO_4 and MgSiO_3 minerals from first-

principles simulations. *Phys. Earth Planet. Inter.* **240**, 1–24 (2015).

27. 27.

Yu, Y. G., Wentzcovitch, R. M., Vinograd, V. L. & Angel, R. J. Thermodynamic properties of MgSiO_3 majorite and phase transitions near 660 km depth in MgSiO_3 and Mg_2SiO_4 : a first principles study. *J. Geophys. Res. Solid Earth* <https://doi.org/10.1029/2010JB007912> (2011).

28. 28.

Ashida, T., Kume, S., Ito, E. & Navrotsky, A. MgSiO_3 ilmenite: heat capacity, thermal expansivity, and enthalpy of transformation. *Phys. Chem. Miner.* **16**, 239–245 (1988).

29. 29.

Akaogi, M. & Ito, E. Refinement of enthalpy measurement of MgSiO_3 perovskite and negative pressure–temperature slopes for perovskite-forming reactions. *Geophys. Res. Lett.* **20**, 1839–1842 (1993).

30. 30.

Sokolova, T. S., Dorogokupets, P. I., Litasov, K. D., Danilov, B. S. & Dymshits, A. M. Spreadsheets to calculate P – V – T relations, thermodynamic and thermoelastic properties of silicates in the MgSiO_3 – MgO system. *High Pressure Res.* **38**, 193–211 (2018).

31. 31.

Stixrude, L., Wentzcovitch, R. & Kiefer, B. First principles investigation of the MgSiO_3 akimotoite to perovskite phase transition. In *APS March Meeting Abstracts* W34.009 (APS, 2003).

32. 32.

Kiefer, B. & Stixrude, L. Cation disorder and elasticity in MgSiO_3 akimotoite. In *AGU Fall Meeting Abstracts* MR11A-12 (AGU, 2002).

33. 33.

Wentzcovitch, R. M., Stixrude, L., Karki, B. & Kiefer, B. Akimotoite to perovskite phase transition in MgSiO_3 . *Geophys. Res. Lett.* **31**, L10611 (2004).

34. 34.

Horiuchi, H., Hirano, M., Ito, E. & Matsui, Y. MgSiO_3 (ilmenite-type): single crystal X-ray diffraction study. *Am. Mineral.* **67**, 788–793 (1982).

35. 35.

Ahrens, L. H. The use of ionization potentials. Part 1. Ionic radii of the elements. *Geochim. Cosmochim. Acta* **2**, 155–169 (1952).

36. 36.

Thompson, A. B. Water in the Earth's upper mantle. *Nature* **358**, 295–302 (1992).

37. 37.

Ganguly, J., Freed, A. M. & Saxena, S. K. Density profiles of oceanic slabs and surrounding mantle: integrated thermodynamic and thermal modeling, and implications for the fate of slabs at the 660 km discontinuity. *Phys. Earth Planet. Inter.* **172**, 257–267 (2009).

38. 38.

Jenkins, J., Cottaar, S., White, R. & Deuss, A. Depressed mantle discontinuities beneath Iceland: evidence of a garnet controlled 660 km discontinuity? *Earth Planet. Sci. Lett.* **433**, 159–168 (2016).

39. 39.

Ito, E. In *High Pressure Research in Geophysics* (eds Akimoto, S. & Manghnani, M. H.) 405–419 (Tokyo Center of Academic Publications, 1982).

40. 40.

Nakajima, Y., Frost, D. J. & Rubie, D. C. Ferrous iron partitioning between magnesium silicate perovskite and ferropericlase and the composition of perovskite in the Earth's lower mantle. *J. Geophys. Res. Solid Earth* <https://doi.org/10.1029/2012JB009151> (2012).

41. 41.

Schubert, G., Yuen, D. A. & Turcotte, D. L. Role of phase transitions in a dynamic mantle. *Geophys. J. Int.* **42**, 705–735 (1975).

42. 42.

Christensen, U. R. & Yuen, D. A. The interaction of a subducting lithospheric slab with a chemical or phase boundary. *J. Geophys. Res. Solid Earth* **89**, 4389–4402 (1984).

43. 43.

Torii, Y. & Yoshioka, S. Physical conditions producing slab stagnation: constraints of the Clapeyron slope, mantle viscosity, trench retreat, and dip angles. *Tectonophysics* **445**, 200–209 (2007).

44. 44.

Agrusta, R., Goes, S. & van Hunen, J. Subducting-slab transition-zone interaction: stagnation, penetration and mode switches. *Earth Planet. Sci. Lett.* **464**, 10–23 (2017).

45. 45.

Van der Hilst, R. D., Widjiantoro, S. & Engdahl, E. Evidence for deep mantle circulation from global tomography. *Nature* **386**, 578–584 (1997).

46. 46.

Van der Hilst, R. & Seno, T. Effects of relative plate motion on the deep structure and penetration depth of slabs below the Izu–Bonin and Mariana island arcs. *Earth Planet. Sci. Lett.* **120**, 395–407 (1993).

47. 47.

Tange, Y., Nishihara, Y. & Tsuchiya, T. Unified analyses for P – V – T equation of state of MgO: a solution for pressure-scale problems in high P – T experiments. *J. Geophys. Res. Solid Earth* <https://doi.org/10.1029/2008JB005813> (2009).

48. 48.

Nishihara, Y., Doi, S., Kakizawa, S., Higo, Y. & Tange, Y. Effect of pressure on temperature measurements using WRe thermocouple and its geophysical impact. *Phys. Earth Planet. Inter.* **298**, 106348 (2020).

49. 49.

Katsura, T., Yoneda, A., Yamazaki, D., Yoshino, T. & Ito, E. Adiabatic temperature profile in the mantle. *Phys. Earth Planet. Inter.* **183**, 212–218 (2010).

50. 50.

Dziewonski, A. M. & Anderson, D. L. Preliminary reference Earth model. *Phys. Earth Planet. Inter.* **25**, 297–356 (1981).

51. 51.

Ishii, T. et al. Generation of pressures over 40 GPa using Kawai-type multi-anvil press with tungsten carbide anvils. *Rev. Sci. Instrum.* **87**, 024501 (2016).

52. 52.

Katsura, T. et al. A large-volume high-pressure and high-temperature apparatus for in situ X-ray observation, ‘SPEED-Mk. II’. *Phys. Earth Planet. Inter.* **143**, 497–506 (2004).

53. 53.

Irifune, T. et al. The postspinel phase boundary in Mg_2SiO_4 determined by in situ X-ray diffraction. *Science* **279**, 1698–1700 (1998).

54. 54.

Hao, S., Wang, W., Qian, W. & Wu, Z. Elasticity of akimotoite under the mantle conditions: implications for multiple discontinuities and seismic anisotropies at the depth of ~ 600 –750 km in subduction zones. *Earth Planet. Sci. Lett.* **528**, 115830 (2019).

55. 55.

Valdez, M. N., Wu, Z., Yonggang, G. Y., Revenaugh, J. & Wentzcovitch, R. M. Thermoelastic properties of ringwoodite $(\text{Fe}_x\text{Mg}_{1-x})_2\text{SiO}_4$: its relationship to the 520 km seismic discontinuity. *Earth Planet. Sci. Lett.* **351**, 115–122 (2012).

56. 56.

Shukla, G. et al. Thermoelasticity of Fe^{2+} -bearing bridgmanite. *Geophys. Res. Lett.* **42**, 1741–1749 (2015).

57. 57.

Wu, Z. & Wentzcovitch, R. M. Quasiharmonic thermal elasticity of crystals: an analytical approach. *Phys. Rev. B* **83**, 184115 (2011).

58. 58.

Zhou, C., Greaux, S., Nishiyama, N., Irifune, T. & Higo, Y. Sound velocities measurement on MgSiO_3 akimotoite at high pressures and high temperatures with simultaneous in situ X-ray diffraction and ultrasonic study. *Phys. Earth Planet. Inter.* **228**, 97–105 (2014).

59. 59.

Wang, Y., Uchida, T., Zhang, J., Rivers, M. L. & Sutton, S. R. Thermal equation of state of akimotoite MgSiO_3 and effects of the akimotoite–garnet transformation on seismic structure near the 660 km discontinuity. *Phys. Earth Planet. Inter.* **143**, 57–80 (2004).

60. 60.

Funamori, N. et al. Thermoelastic properties of MgSiO_3 perovskite determined by in situ X ray observations up to 30 GPa and 2000 K. *J. Geophys. Res. Solid Earth* **101**, 8257–8269 (1996).

Acknowledgements

We thank H. Fischer, S. Übelhack, R. Njul, U. Trenz and S. Linhardt at the Bayerisches Geoinstitut and S. Sonntag at DESY for their technical assistance, K. Glazyrin for insightful discussions, and E. Posner for English correction of the manuscript. This work was funded by a research project approved by the Federal Ministry of Education and Research (BMBF) (05K16WC2), the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (proposal no. 787 527), the German Research Foundation (DFG) (KA3434/9-1) to T.K., and DFG (IS350/1-1) to T.I. The synchrotron X-ray diffraction experiments were performed at the PETRA III beamline P61B at DESY (Hamburg, Germany), a member of the Helmholtz Association HGF, and at the beamline BL04B1 at SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (proposal nos. 2019B1133, 2019A1353, 2018B1209, 2018B1218 and 2017B1078).

Funding

Open access funding provided by Deutsches Elektronen-Synchrotron (DESY).

Author information

Affiliations

1. Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany

Artem Chanyshhev, Shrikant Bhat & Robert Farla

2. Bayerisches Geoinstitut, University of Bayreuth, Bayreuth, Germany

Artem Chanyshhev, Takayuki Ishii, Dmitry Bondar, Eun Jeong Kim, Keisuke Nishida, Zhaodong Liu, Lin Wang & Tomoo Katsura

3. Center for High Pressure Science and Technology Advanced Research, Beijing, China

Takayuki Ishii, Bingmin Yan, Hu Tang, Zhen Chen & Tomoo Katsura

4. State Key Laboratory of Superhard Materials, Jilin University, Changchun, China

Zhaodong Liu

5. Earth and Planets Laboratory, Carnegie Institution, Washington, DC, USA

Lin Wang

6. Department of Earth Sciences, Graduate School of Science, Tohoku University, Sendai, Japan

Ayano Nakajima

7. Japan Synchrotron Radiation Research Institute (JASRI), Sayo, Japan

Yuji Higo & Yoshinori Tange

Contributions

A.C. and T.I. conducted experiments for the determinations of AB and RBP phase-transition boundaries, respectively. The current experimental strategy was developed by T.K. The experimental method was developed by T.I. T.I. instructed A.C. about the experimental method. D.B., Z.L., L.W., A.N., B.Y., H.T., Z.C., Y.H. and Y.T. operated the synchrotron radiation experiments at beamline BL04B1 at SPring-8. D.B., S.B., E.J.K. and R.F. operated synchrotron radiation experiments at beamline P61B at

DESY. A.C. and T.I. analysed the data. A.C. wrote the first draft. All authors commented on the manuscript. The whole project was directed by T.K.

Corresponding authors

Correspondence to [Artem Chanyshев](#) or [Takayuki Ishii](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review information

Nature thanks Ian Jackson and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

[Extended Data Fig. 1 Comparison of the present results with previous data.](#)

a, RBP transition boundary. **b**, AB transition boundary. The present data are shown with solid black lines and grey translucent areas. Phase transition boundaries determined in *in situ* X-ray multi-anvil experiments^{[11,12,13,16,20,22,24,53](#)} are shown by dashed and dotted lines, those determined in multi-anvil laboratory experiments^{[2,21,25](#)} are shown by dashed lines with dots, and those determined computationally^{[26,27](#)} are shown by thin solid lines.

[Extended Data Fig. 2 Heat capacities of MgSiO₃ Ak and Brg.](#)

Comparison of the heat capacities of MgSiO₃ Ak and Brg explaining the present AB boundary (dotted blue and red lines) with those experimentally measured (thick solid

blue and red lines)^{28,29} and those calculated (dashed blue and dark red lines)³⁰. 3nR denotes the limiting value of the Debye heat capacity.

Extended Data Fig. 3 Seismic velocity contrasts caused by phase transitions.

ΔV_p and ΔV_s caused by the AB (red and blue solid lines and red arrows), RAP (purple dashed and orange dot-dashed lines and orange arrows), RBP (purple dashed and green dotted lines and green arrows), and $Ak + Pc \rightarrow Brg + Pc$ (orange dot-dashed and green dotted lines and green arrows) transitions along the cold geotherm. Data sources: Ak, ref. ⁵⁴; Rw, ref. ⁵⁵; Brg, ref. ⁵⁶; Pc, ref. ⁵⁷. Figure redrawn from ref. ⁵⁴.

Extended Data Fig. 4 X-ray diffraction patterns of starting materials.

X-ray diffraction patterns of bridgmanite and bridgmanite plus periclase aggregates used as starting material to determine the AB and RBP phase transition boundaries, respectively. The samples were synthesized at 27 GPa and 1,000 K for 1 h in a multi-anvil press from enstatite powder to obtain bridgmanite, and from forsterite powder to obtain the mixture of bridgmanite and periclase. The diffraction peaks of stishovite and periclase are marked by a blue asterisk and pink plus, respectively.

Source data

Extended Data Table 1 Compressibility and thermal expansion coefficients of akimotoite and bridgmanite

Supplementary information

Peer Review File

Supplementary Tables

Supplementary Tables 1 and 2.

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Extended Data Fig. 4

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Reprints and Permissions

About this article

Cite this article

Chanyshев, A., Ishii, T., Bondar, D. *et al.* Depressed 660-km discontinuity caused by akimotoite–bridgmanite transition. *Nature* **601**, 69–73 (2022).
<https://doi.org/10.1038/s41586-021-04157-z>

- Received: 26 November 2020
- Accepted: 18 October 2021
- Published: 05 January 2022
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04157-z>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04157-z>

| [Section menu](#) | [Main menu](#) |

- Article
- Open Access
- [Published: 15 December 2021](#)

Experimental evidence for recovery of mercury-contaminated fish populations

- [Paul J. Blanchfield](#) [ORCID: orcid.org/0000-0003-0886-5642^{1,2,3}](#),
- [John W. M. Rudd](#) [ORCID: orcid.org/0000-0002-4805-4337^{1 nAff17}](#),
- [Lee E. Hrenchuk](#) [ORCID: orcid.org/0000-0002-6194-094X^{1,3}](#),
- [Marc Amyot](#) [ORCID: orcid.org/0000-0002-0340-3249⁴](#),
- [Christopher L. Babiarz⁵](#),
- [Ken G. Beaty¹](#),
- [R. A. Drew Bodaly¹](#),
- [Brian A. Branfireun⁶](#),
- [Cynthia C. Gilmour](#) [ORCID: orcid.org/0000-0002-1720-9498⁷](#),
- [Jennifer A. Graydon](#) [ORCID: orcid.org/0000-0001-5243-6891⁸](#),
- [Britt D. Hall⁹](#),
- [Reed C. Harris¹⁰](#),
- [Andrew Heyes](#) [ORCID: orcid.org/0000-0001-7855-8512¹¹](#),
- [Holger Hintelmann](#) [ORCID: orcid.org/0000-0002-5287-483X¹²](#),
- [James P. Hurley¹³](#),
- [Carol A. Kelly](#) [ORCID: orcid.org/0000-0003-2473-7560^{1 nAff17}](#),
- [David P. Krabbenhoft¹⁴](#),
- [Steve E. Lindberg¹⁵](#),
- [Robert P. Mason](#) [ORCID: orcid.org/0000-0002-7443-4931¹⁶](#),
- [Michael J. Paterson^{1,3}](#),
- [Cheryl L. Podemski](#) [ORCID: orcid.org/0000-0001-6168-7326¹](#),
- [Ken A. Sandilands^{1,3}](#),
- [George R. Southworth¹⁵](#),
- [Vincent L. St Louis](#) [ORCID: orcid.org/0000-0001-5405-1522⁸](#),
- [Lori S. Tate^{1 nAff18}](#) &
- [Michael T. Tate](#) [ORCID: orcid.org/0000-0003-1525-1219¹⁴](#)

Nature volume 601, pages 74–78 (2022)

- 6192 Accesses
- 148 Altmetric
- [Metrics details](#)

Subjects

- [Boreal ecology](#)
- [Ecosystem ecology](#)
- [Environmental impact](#)
- [Freshwater ecology](#)

Abstract

Anthropogenic releases of mercury (Hg)^{1,2,3} are a human health issue⁴ because the potent toxicant methylmercury (MeHg), formed primarily by microbial methylation of inorganic Hg in aquatic ecosystems, bioaccumulates to high concentrations in fish consumed by humans^{5,6}. Predicting the efficacy of Hg pollution controls on fish MeHg concentrations is complex because many factors influence the production and bioaccumulation of MeHg^{7,8,9}. Here we conducted a 15-year whole-ecosystem, single-factor experiment to determine the magnitude and timing of reductions in fish MeHg concentrations following reductions in Hg additions to a boreal lake and its watershed. During the seven-year addition phase, we applied enriched Hg isotopes to increase local Hg wet deposition rates fivefold. The Hg isotopes became increasingly incorporated into the food web as MeHg, predominantly from additions to the lake because most of those in the watershed remained there. Thereafter, isotopic additions were stopped, resulting in an approximately 100% reduction in Hg loading to the lake. The concentration of labelled MeHg quickly decreased by up to 91% in lower trophic level organisms, initiating rapid decreases of 38–76% of MeHg concentration in large-bodied fish populations in eight years. Although Hg loading from watersheds may not decline in step with lowering deposition rates, this experiment clearly demonstrates that any reduction in Hg loadings to lakes, whether from direct deposition or runoff, will have immediate benefits to fish consumers.

[Download PDF](#)

Main

The Minamata Convention on Mercury is an international treaty that aims to protect human health and the environment from adverse effects of MeHg by controlling Hg

emissions, which should then decrease deposition and loading of anthropogenic Hg to aquatic environments¹⁰. Yet there is little direct evidence for how quickly fish MeHg concentrations will decline following reductions in current rates of Hg loading owing, in part, to a range of ecological factors that can influence both the microbial production and the bioaccumulation of MeHg in aquatic food webs^{9,11}. Further complicating this relationship are human activities such as commercial fishing, introduction of exotic species and enhanced nutrient additions that trigger large-scale trophic disruptions¹², which can in turn substantially alter fish tissue MeHg concentrations^{13,14,15}, because fish acquire most of their MeHg through their diet¹⁶. Changing climatic conditions can also influence MeHg production⁸, as well as restructure food webs, alter the dominant pathways of energy flow and cause size-dependent changes in fish growth rates that shift population size structure^{14,17,18}. In addition, until now there has been no way to evaluate the relative contribution of newly deposited Hg to contemporary MeHg production. Consequently, it is exceedingly difficult to unambiguously assess the recovery of contaminated fish populations due specifically to Hg control measures⁹.

Over a 15-year period (2001–2015) we conducted a whole-ecosystem Hg loading and recovery experiment (Mercury Experiment To Assess Atmospheric Loading In Canada and the United States (METAALICUS)) in a pristine boreal watershed¹⁹. METAALICUS addresses the relationship between changes in inorganic Hg loadings to a lake and MeHg concentrations in fish using highly enriched inorganic Hg isotopes (termed ‘spikes’) that enabled us to specifically follow a change in loading against a background of previously deposited Hg and present-day, relatively constant, Hg inputs from direct deposition to the lake surface and from the watershed. In our experiment, these are defined as ‘ambient Hg’. By adding a different spike Hg to the lake (²⁰²Hg), wetland (¹⁹⁸Hg) and upland (²⁰⁰Hg) compartments of the Lake 658 watershed (52 ha in total) during a 6- to 7-year addition phase (Fig. 1a) we could follow the uptake of MeHg in fish derived solely from newly deposited Hg¹⁹. We then ceased all experimental additions to determine the magnitude and timing of reductions in fish MeHg concentrations to reductions in Hg loading to the lake, which we tested by tracking the decline in spike MeHg in fish, their prey and other compartments of the lake ecosystem over an eight-year recovery phase. The diverse fish community of the METAALICUS lake enabled assessment of contaminant bioaccumulation and recovery across different trophic guilds and different exposure pathways (that is, sediment versus water) for three species important to freshwater fisheries across the boreal ecoregion²⁰ (planktivore: yellow perch (*Perca flavescens*); benthivore: lake whitefish (*Coregonus clupeaformis*); and piscivore: northern pike (*Esox lucius*)).

Fig. 1: Temporal dynamics of mercury addition and recovery in the Lake 658 ecosystem.

 **figure1**

a, Location (inset) of the Experimental Lakes Area (ELA), Canada, where Hg enriched with different isotopes was applied to the wetland, upland and lake surface of Lake 658 to simulate enhanced wet deposition of Hg (dark blue shaded area). **b**, Inorganic Hg added to the lake was methylated and measured as MeHg concentration in water (in ng l^{-1} ; $n = 516$), sediments (in ng g^{-1} dry weight; $n = 1,627$) and invertebrates (in ng g^{-1} wet weight; $n = 211$), and as total Hg in fishes (in ng g^{-1} wet weight; $n = 1,052$). Mean annual concentrations for the open-water season are shown for all lake components except for fish populations, which were collected each autumn. Concentration data for large-bodied fish are derived from body-length standardization (pike, 475 mm; whitefish, 535 mm). **c**, Hg loading to the lake increased MeHg concentrations (per cent increase = [lake spike MeHg]/[ambient MeHg] $\times 100$) during the addition phase (2001–2007), then decreased during the recovery phase (2008–2015), when experimental Hg additions to the ecosystem ceased (light blue shaded area in **a**). Dotted lines indicate missing data.

[Source data](#).

The METAALICUS watershed is in an undisturbed remote region of Canada, such that our experimental addition rate increased wet Hg deposition approximately fivefold

(from approximately 3.6 to 19 $\mu\text{g m}^{-2} \text{ yr}^{-1}$), to levels similar to more polluted regions of the world²¹. Most of the Hg added to the wetland and upland areas of the watershed either remained bound to vegetation and soils or evaded back to the atmosphere^{22,23}. The wetland spike was below the detection level in all fish species. The Hg applied to the upland catchment accounted for only a small fraction (less than 1%) of all Hg in runoff to the lake¹⁹ and consequently contributed little (less than 2%) to the changes in MeHg concentrations of fish populations throughout the study (Extended Data Fig. 1). Hence, after six years of increased additions to the watershed (Fig. 1a), the observed increases in fish MeHg were due almost entirely to Hg added directly to the lake surface. This did not appear to be caused by preferential methylation of lake spike Hg. The evidence for this is that the seasonal production of lake spike MeHg, all of which had to be formed within the lake because it was added as inorganic Hg, varied synchronously with ambient MeHg in the lake and in biota¹⁹. This finding also indicates that in this headwater lake ambient MeHg is mainly derived from in-lake methylation of inorganic ambient Hg, most of which came from the upland catchment.

Delivery of lake spike Hg to the sediments and anoxic bottom waters, which are the dominant sites of methylation in the study lake^{19,24}, resulted in formation of spike MeHg in these compartments, from where spike MeHg also migrated to surface waters. Lake spike MeHg rapidly accumulated in all lake biota (Fig. 1b), with concentrations in fish muscle increasing with continued loading for all species (Extended Data Table 1; linear regression, $P < 0.05$), apart from young-of-year (YOY) yellow perch, which showed high inter-annual variability in spike MeHg concentrations after an initial increase (Fig. 1b). By contrast, ambient MeHg concentrations in all fish species did not show any consistent trends during the addition phase (Fig. 2a–c), nor did they in a nearby reference lake (Extended Data Table 1; $P > 0.05$). Steady ambient MeHg concentrations in fish through time are indicative of relatively stable watershed inputs of Hg, which is the main source of ambient inorganic Hg for methylation in both the experimental and reference lakes^{19,25}.

Fig. 2: Accumulation and trophic transfer of lake spike and ambient mercury.

 **figure2**

a–c, Annual fish muscle MeHg concentrations (total MeHg = lake spike MeHg + ambient MeHg; grey circles) increased above background concentrations (ambient MeHg; white circles) during the addition phase (dark blue shaded area) from uptake of isotope enriched Hg added to Lake 658 (lake spike; black circles) for planktivorous (age 1+ yellow perch; $n = 140$) (**a**), piscivorous (northern pike; $n = 442$) (**b**) and benthivorous (lake whitefish; $n = 189$) (**c**) populations, then declined during the recovery phase (light blue shaded area). **d–f**, Biomagnification factors (BMF = $[MeHg_{\text{predator}}]/[MeHg_{\text{prey}}]$) of lake spike MeHg and ambient MeHg from dominant prey items for each of these fish species were as follows: zooplankton ($n = 127$) to yellow perch (**d**); forage fish ($n = 421$) to northern pike (**e**); and *Chaoborus* ($n = 62$) to lake whitefish (**f**). Fish data are means from autumn sampling (sample sizes in Extended Data Tables [2](#), [3](#)). Concentration data for pike and whitefish are derived from body-length standardization; dotted lines indicate missing data.

[Source data](#).

A critical question for the addition period was how much higher were MeHg concentrations than they would have been in the absence of the experimentally increased loading of Hg. The addition of lake spike Hg was roughly equivalent to all ambient Hg inputs (runoff plus direct deposition) to the lake, resulting in a doubling, or about 100% increase, in Hg loading to the lake^{[19](#)}. In response to seven years of experimental additions to the lake, per cent increases in lake spike MeHg concentrations were highest in water (60%) and least in the upper 2 cm of sediments (30%) where large stores of ambient Hg existed (Fig. [1c](#)). The response of food web organisms was intermediate to that of water and sediments, such that spike Hg additions to the lake raised MeHg concentrations by 45–57% in invertebrates and forage fishes and by more than 40% for large-bodied fish species (Fig. [1c](#)).

Temporal patterns of biomagnification for spike MeHg relative to ambient MeHg inform how quickly the different fish species came into equilibrium with their respective prey. For small-bodied yellow perch (1 year of age) feeding on zooplankton, it took three years before the biomagnification of lake spike resembled that of ambient MeHg (Fig. 2d), and a further two years for both the apex predator, northern pike, feeding on forage fishes (Fig. 2e), and lake whitefish feeding on *Chaoborus* (Fig. 2f). Relative to planktivorous yellow perch, final addition phase concentrations of spike MeHg were slightly higher for benthivorous lake whitefish (1.2 \times) and further increased for piscivorous northern pike (3.9 \times), similar to ambient MeHg (whitefish (1.5 \times) and pike (3.9 \times); relative to perch; Fig. 2a–c) and consistent with expectations of contaminant biomagnification among trophic guilds^{5,20}. These findings imply that the key in-lake processes leading to the formation and trophic transfer of MeHg to the different fish populations became comparable for spike Hg and ambient Hg during the addition phase.

To then directly test the hypothesis that MeHg concentrations in fishes would decline following reductions in Hg loading to the lake, we ceased all experimental additions of enriched Hg isotopes (Fig. 1a). This resulted in a 100% reduction in loading of lake spike. Average concentrations of lake spike MeHg in fish populations rapidly declined (within less than ten years) in concert with the decline in the availability of spike MeHg through dietary and waterborne pathways (Fig. 1b). Within the first 3 years, the relative amount of lake spike MeHg declined by 81% in water, 35% in sediments, 66% in zooplankton and 67% in *Chaoborus* (Fig. 1c), leading to marked reductions (85–91%) in the concentration of spike MeHg in forage fish species by the end of the recovery phase (Fig. 1b). Eight years after addition, lake spike Hg contributed just a small fraction (approximately 6%) to MeHg concentrations in forage fishes and invertebrate prey (Fig. 1c). The more rapid decline in per cent spike MeHg in water compared with sediments, even in this relatively long water residence time lake (about 6 years), emphasizes that the magnitude and timing of responses by fish to Hg loading reductions could be influenced by their relative reliance upon pelagic versus benthic dietary pathways²⁶.

The notably fast response of the lower food web to the cessation of lake spike Hg loadings initiated rapid recovery of large-bodied fish species. Within 8 years, lake spike MeHg concentrations declined by 76% in the northern pike population and by 38% in the lake whitefish population (Fig. 1b). During the recovery phase, spike MeHg concentrations for these large-bodied species initially increased for both populations before showing steady declines (Fig. 2b,c). The rate of decline in spike MeHg for northern pike, however, was roughly twice that of lake whitefish (Fig. 1c).

Differences in the lifespan of the fish populations had a key role in the rates of recovery following the reduction of Hg loadings to the lake. Lake whitefish were

much older (median age = 17 years versus 3 years for pike) and larger (Extended Data Tables 2, 3) than northern pike, and more individuals in that population would have lived through some or all of the addition and recovery phases of the experiment. Lake whitefish had the coldest thermal preferences and greatest association with benthic habitats of any fish population, which probably also contributed to their delayed recovery.

Boreal fishes are known to eliminate MeHg very slowly once accumulated²⁷. To further explain the recovery of the apex predator population, we tracked changes in the body burdens of spike MeHg in individual northern pike over time while also monitoring the population as a whole. As expected, individual responses were variable, but lake spike MeHg burdens in northern pike mostly increased during the early recovery phase with overall little to no loss of the spike MeHg 6–8 years after cessation of spike additions (Fig. 3). These findings parallel those observed for lake spike MeHg in individual northern pike moved from the study lake to a nearby reference lake²⁸ and underscore how the prolonged retention of MeHg in fish muscle tissue can delay recovery of some fisheries^{7,29}. Thus, it was the annual recruitment of new fish with low MeHg concentrations into the population, along with the loss of older fish (as evidenced by a stable population size structure; Extended Data Fig. 2), that enabled the swift recovery of the population from Hg contamination as a whole. Consequently, average burdens of spike MeHg in the northern pike population were reduced by 50% in less than five years, in spite of the efficient retention of spike MeHg by some older fish (Fig. 3, Extended Data Fig. 2).

Fig. 3: Recovery of the apex predator from mercury loading.



Comparison of changes in body burdens of lake spike MeHg during the recovery phase for the northern pike population (annual mean, black circles) to that of individual northern pike (grey lines and triangles). Individual northern pike were sampled at the end of the addition phase (in 2007; $n = 16$) and subsequently recaptured during the recovery phase (each line represents an individual fish). Population data are based on all fish captured each autumn ($n = 280$). All northern pike were sampled using a non-lethal biopsy (represented in images) in the autumn of each year and returned to the lake. Fish body burdens of lake spike MeHg (body burden = lake spike MeHg (ng g^{-1}) \times fish mass (g)) were normalized to concentrations in the autumn of 2007 (t_0 ; the final time isotope-enriched Hg was added to Lake 658 and the beginning of the recovery period). Exponential decay regression starting in the second year of recovery estimated a 50% reduction in lake spike MeHg burden in the population in 4.2 years (data are mean (black circle) \pm 95% confidence interval (shaded band); line fit: $y = 1.7439 \times e^{-0.2928x}$, $R^2 = 0.95$, $F_{1,6} = 95.5$, $P = 0.0002$).

Source data.

Differentiating the relative importance of present day inputs versus previously deposited Hg to overall fish MeHg concentrations is a key uncertainty when predicting the efficacy of Hg pollution reduction^{29,30}. Here we demonstrate that within a few years of abatement, the experimental Hg added previously to the lake was no longer an important source of MeHg to the lower food web or to forage fishes (Fig. 1c). Long-lived fish species of subsistence, commercial and recreational importance lagged behind their prey, but the contribution of recently deposited Hg to fish MeHg steadily diminished for these populations as well. There was a similar, rapid response for the upland spike when loading ceased, even though only a small amount appeared in the fish during the experiment (Extended Data Fig. 1). The small contribution of the terrestrial spike to fish MeHg supports our former conclusion¹⁹ that lakes with large watersheds will respond more slowly to changes in atmospheric deposition.

The most important outcome of this whole-ecosystem experiment is the demonstration that a decrease in a single factor (Hg loading to the lake) has a clear and timely effect on average MeHg concentrations in fish populations, even for long-lived species that eliminate MeHg slowly. The spike MeHg data show that fish populations will respond quickly to any change in loading rates—whether from direct deposition to the lake (Fig. 1) or runoff (Extended Data Fig. 1). Decreases in loading to the lake from these two sources will follow different time courses in response to lower atmospheric deposition¹⁹. However, as these two loads decrease, the fish populations in the receiving lake will soon afterwards have lower MeHg than they would have if nothing were done, thereby reducing human exposure.

Methods

Mercury additions to the study catchment

METAALICUS was conducted on the Lake 658 catchment at the Experimental Lakes Area (ELA; now IISD-ELA), a remote area in the Precambrian Shield of northwestern Ontario, Canada ($49^{\circ} 43' 95''$ N, $93^{\circ} 44' 20''$ W) set aside for whole-ecosystem research³¹. The Lake 658 catchment includes upland (41.2 ha), wetland (1.7 ha) and lake surface (8.4 ha) areas. Lake 658 is a double basin (13 m depth), circumneutral, headwater lake, with a fish community consisting of forage (yellow perch (*P. flavescens*) and blacknose shiner (*Notropis heterolepis*)), benthivorous (lake whitefish (*C. clupeaformis*) and white sucker (*Catostomus commersonii*)), and piscivorous (northern pike (*E. lucius*)) fishes. The lake is closed to fishing.

Hg addition methods used in METAALICUS have been described in detail elsewhere^{19,32,33}. In brief, three Hg spikes, each enriched with a different stable Hg isotope, were applied separately to the lake surface, upland and wetland areas. Upland and wetland spikes were applied once per year (when possible; Fig. 1a) by fixed-wing aircraft (Cessna 188 AGtruck). Mercury spikes (as HgNO_3) were diluted in acidified water (pH 4) in a 500 l fiberglass tank and sprayed with a stainless-steel boom on upland (approximately 79.9% ^{200}Hg) and wetland (approximately 90.1% ^{198}Hg) areas. Spraying was completed during or immediately before a rain event, with wind speeds less than 15 km h^{-1} to minimize drift of spike Hg outside of target areas. Aerial spraying of upland and wetland areas left a 20-m buffer to the shoreline, which was sprayed by hand with a gas-powered pump and fire hose to within about 5 m of the lake³². Average net application rates of isotopically labelled Hg to the upland and wetland areas were $18.5 \mu\text{g m}^{-2} \text{ yr}^{-1}$ and $17.8 \mu\text{g m}^{-2} \text{ yr}^{-1}$, respectively.

The average net application rate for lake spike Hg was $22.0 \mu\text{g m}^{-2} \text{ yr}^{-1}$. For each lake addition, inorganic Hg enriched with approximately 89.7% ^{202}Hg was added as HgNO_3 from four 20-l carboys filled with acidified lake water (pH 4). Nine lake additions were conducted bi-weekly at dusk over an 18-week (wk) period during the open-water season of each year (2001–2007) by injecting at 70-cm depth into the propeller wash of trolling electric motors of two boats crisscrossing each basin of the lake^{32,33}. It was previously demonstrated with ^{14}C additions to an ELA lake that this approach evenly distributed spike added in the evening by the next morning³⁴.

We did not attempt to simulate Hg in rainfall for isotopic lake additions because it is impossible to simulate natural rainfall concentrations (about 10 ng l^{-1}) in the 20-l carboys used for additions. Instead, our starting point for the experiment was to ensure that the spike was behaving as closely as possible to ambient surface water Hg very

soon after it entered the lake. Several factors support this assertion. By the next morning each spike addition had increased epilimnetic Hg concentrations by only 1 ng l⁻¹ ²⁰²Hg. Average ambient concentrations were 2 ng l⁻¹. Thus, while the Hg concentrations in the carboys were high (2.6 mg l⁻¹), the receiving waters were soon at trace levels. Furthermore, we investigated if the additions altered the degree of bioavailability or photoreactivity of Hg(ii) in the receiving surface water. We examined the bioavailability of spike Hg(ii) as compared to ambient Hg in the lake itself using a genetically engineered bioreporter bacterium³⁵. On seven occasions, epilimnetic samples were collected on the day before and within 12 h of spike additions. The spike was added to the lake as Hg(NO₃)₂, which is bioavailable to the bioreporter bacterium (detection limit = 0.1 ng Hg(ii) l⁻¹), but we never saw bioavailable ambient or spike Hg(ii) in the lake, presumably because it was quickly bound to dissolved organic carbon (DOC). This indicates that, in terms of bioavailability, the spike Hg was behaving like ambient Hg soon after additions. Photoreactivity in the surface water was examined on seven occasions, by measuring the % of total Hg(ii) that was dissolved gaseous Hg for spike and ambient Hg, either 24 h or 48 h after the lake was spiked³⁶. There was no significant difference (paired *t*-test, *P* > 0.05), demonstrating that by then the lake spike was behaving in the same way as ambient Hg during gaseous Hg production.

Lake, food web and fish sampling

Water samples were collected from May to October every four weeks at the deepest point of Lake 658. Water was pumped from six depths through acid-cleaned Teflon tubing into acid-cleaned Teflon or glass bottles. Water samples were filtered in-line using pre-ashed quartz fibre filters (Whatman GFQ, 0.7 µm). Subsequently, Hg species were measured in the filtered water samples (dissolved Hg and MeHg) and in particles collected on the quartz fibre filter (particulate Hg and MeHg).

From 2001 to 2012, Lake 658 sediments were sampled at 4 fixed sites up to 5 times per year. Sampling frequency was highest in 2001, with monthly sampling from May to September, and declined over the course of the study. Fixed sites were located at depths of 0.5, 2, 3 and 7 m. A sediment survey of up to 12 additional sites was also conducted once or twice each year. Survey sites were selected to represent the full range of water depths in both basins. Cores were collected by hand by divers, or by subsampling sediments collected using a small box corer. Cores were capped and returned to the field station for processing within a few hours. For each site, three separate cores were sectioned and composited in zipper lock bags for a 0- to 2-cm depth sampling horizon, and then frozen at -20 °C.

Bulk zooplankton and *Chaoborus* samples were collected from Lake 658 for MeHg analysis. Zooplankton were collected during the day from May to October (bi-weekly:

2001–2007; monthly: 2008–2015). A plankton net (150 µm, 0.5 m diameter) was towed vertically through the water column from 1 m above the lake bottom at the deepest point to the surface of the lake. Samples were frozen in plastic Whirl-Pak bags after removal of any *Chaoborus* using acid-washed tweezers. Dominant zooplankton taxa in Lake 658 included calanoid copepods (*Diaptomus oregonensis*) and Cladocera (*Holopedium glacialis*, *Daphnia pulicaria* and *Daphnia mendotae*). *Chaoborus* samples were collected monthly in the same manner at least 1 h after sunset. After collection, *Chaoborus* were picked from the sample using forceps and frozen in Whirl-Pak bags. *Chaoborus* were not separated by species for MeHg analyses, but both *C. flavicans* and *C. punctipennis* occur in the lake. Profundal chironomids were sampled at the deepest part of the lake using a standard Ekman grab sampler. Grab material was washed using water from a nearby lake and individual chironomids were picked by hand.

All work with vertebrate animals was approved by Animal Care Committees (ACC) through the Canadian Council on Animal Care (Freshwater Institute ACC for Fisheries and Oceans Canada, 2001–2013; University of Manitoba ACC for IISD-ELA, 2014–2015). Licenses to Collect Fish for Scientific Purposes were granted annually by the Ontario Ministry of Natural Resources and Forestry. Prior to any Hg additions, a small-mesh fence was installed at the outlet of Lake 658 to the downstream lake to prevent movement of fish between lakes. Sampling for determination of MeHg concentrations (measured as total mercury (THg), see below) occurred each autumn (August–October; that is, the end of the growing season in north temperate lakes) for all fish species in Lake 658, and for northern pike and yellow perch in nearby reference Lake 240 (Extended Data Tables [2](#), [3](#)). Fish collections occurred randomly throughout the lakes. Forage fish (YOY and 1+ yellow perch, and blacknose shiner) were captured using small mesh gillnets (6–10 mm) set for <20 min, seine nets, and hoop nets. A small number of fish (up to $n = 20$) of each species and age class (determined by visual inspection) were euthanized immediately following capture in an overdose bath of 0.25 g l⁻¹ tricaine methanesulfonate (TMS; Syndel Laboratories). After transport to the field station, fish were measured for fork length (FL; in mm) and mass (to 0.1 g), then immediately frozen (at -20 °C) in individual WhirlPak bags. A year class failure of yellow perch resulted in a single YOY collected in 2008 (data not presented) and no age 1+ fish in 2009 (Extended Data Table [3](#)).

Large-bodied fish were captured by angling and multi-mesh gill nets (2.5–11.4 cm mesh) set for 20–30 min. Upon capture, each fish was anaesthetized with 0.06 g l⁻¹ TMS, measured for FL (mm), weighed (to 1 g), tagged (Passive Integrated Transponder; Biomark), and a small biopsy of dorsal muscle (0.091 ± 0.002 g wet weight (mean \pm s.e.m)) was collected using a dermal punch^{[37](#)}. Only fish large enough for the biopsy procedure were sampled, such that our analyses include very few juveniles (pike: 317–850 mm FL; whitefish: 344–874 mm FL). Muscle samples were

inserted into 0.6-ml polypropylene vials (Rose Scientific), immediately put on ice, and frozen within 4 h (-20°C). This non-lethal method permitted repeated sampling of individual fish over time^{28,37}. The first ray of either the pectoral or pelvic fin was collected for aging purposes upon first capture. Fish recovered from anaesthesia in a tub of fresh lake water (~ 15 min) before being released back into the lakes. From 2001–2015, we collected 690 biopsy muscle samples from 390 fish (238 northern pike, 114 lake whitefish and 38 white sucker) in Lake 658; 149 fish (90 northern pike, 38 lake whitefish and 21 white sucker) were biopsied more than once (2 to 6 per individual). Because of consistently low annual catches of white sucker (< 10 individuals) across sampling years, we have excluded them from our analyses, but note here that their patterns of lake spike MeHg accumulation and recovery were similar to those of lake whitefish. We were unable to sex most fish because they were either immature or captured outside of their spawning season.

Sample processing and analytical methods

Detailed methods on sample preparation and MeHg or THg analysis, as well as interlaboratory calibrations, have been reported elsewhere for the METAALICUS project^{19,38,39}. In brief, MeHg was distilled from water samples and from sediment using atmospheric pressure water vapour distillation and measured after aqueous phase ethylation using sodium tetraethylborate (NaBEt₄). Volatile Hg species were purged and trapped onto Tenax and MeHg was measured after thermodesorption and GC separation using inductively coupled plasma mass spectrometry (ICP-MS) detection (Micromass Platform or Perkin-Elmer Elan DRC II, respectively)³⁹ and quantification by species specific isotope dilution mass spectrometry. The MeHg isotope dilution standards were synthesized and calibrated in-house. Isotope-dilution spikes were added prior to distillation, and MeHg external standards were routinely calibrated against degradation by measuring the standard against inorganic Hg before and after BrCl digestion. The QC strategy include the regular analysis of blanks, laboratory duplicates and certified reference materials (CRMs) IAEA 405 (International Atomic Energy Agency, Vienna, Austria) and NIST 1566b (National Institute of Standards and Technology, Gaithersburg, Maryland) for MeHg. No CRMs are commercially available for MeHg in water.

All biota samples were handled using clean techniques with Teflon or stainless steel tools cleaned with 95% ethanol^{19,38}. Zooplankton and *Chaoborus* were freeze dried, ground with an acid-washed mortar and pestle, subsampled, and weighed to the nearest 0.00001 g. For determination of MeHg concentrations (ambient, lake spike, upland spike and wetland spike) in invertebrate samples, MeHg was solubilized by treatment with a solution of KOH in ethanol (20 % w/v), ethylated by additions of NaBEt₄, and the resulting volatile Hg species were purged and trapped on carbotrap³⁹.

Samples were thermally desorbed and separated by gas chromatography before quantification by ICP-MS as above³⁹. Samples of CRMs (TORT2 (2001–2013), IAEA452 (2014–2015); National Research Council of Canada, Ottawa, Ontario) were subjected to the same procedures; measured MeHg concentrations in the reference materials were not statistically different from certified values ($P > 0.05$).

Prey fish were kept frozen to maintain consistent wet weights. Approximately 0.2 g of skinless dorsal muscle was removed from each fish, weighed (0.0001 g), and placed in an acid-washed glass vial with a Teflon-lined cap (National Scientific Company). Muscle biopsy samples were weighed to the nearest 0.00001 g (Sartorius BP211D, Data Weighing Systems) before and after freeze-drying (Lyph-lock 12-l freeze dry system Model 77545, Labconco) to obtain wet and dry sample masses, and dry weight proportion²⁸. Fish samples were analysed for THg, which is the sum of organic and inorganic Hg. Because we had previously determined that >90% of the Hg in muscle tissue from yellow perch in Lake 658 is MeHg^{40,41}, here we report fish mercury data as MeHg.

THg concentrations (ambient, lake spike, upland spike and wetland spike) in fish muscle samples were quantified by ICP-MS³⁹. Samples were digested with $\text{HNO}_3/\text{H}_2\text{SO}_4$ (7:3 v/v) and heated at 80 °C until brown NOx gases no longer formed. The THg in sample digests was reduced by SnCl_2 to Hg^0 which was then quantified by ICP-MS (Thermo-Finnigan Element2) using a continuous flow cold vapour generation technique⁴¹. To correct for procedural recoveries, all samples were spiked with $^{201}\text{HgCl}_2$ prior to sample analysis. Samples of CRMs (DORM2 (2001–2011), DORM3 (2012–2013), DORM4 (2014–2015); National Research Council of Canada) were submitted to the same procedures; measured THg concentrations in the reference materials were not statistically different from certified values ($P > 0.05$). Detection limit for each of the spikes was 0.5% of ambient Hg.

Calculations and statistical methods

Analyses were completed with Statistica (6.1, Statsoft) and SigmaPlot (11.0, Systat Software). We present wet weight (w.w.) MeHg concentrations for all samples, except sediments which are dry weight (d.w.) concentrations. For zooplankton, *Chaoborus*, and profundal chironomids, d.w. MeHg concentrations were multiplied by a standard proportion (0.15) to yield w.w. concentrations for each sample⁴². The resulting w.w. concentrations were averaged over each open water season to determine annual means. For fish muscle biopsies, d.w. MeHg concentrations were multiplied by individual d.w. proportions to yield w.w. MeHg concentrations for each sample. To avoid any size-related biases, we calculated standardized annual MeHg concentrations (ambient and lake spike) for northern pike and lake whitefish by determining best-fit relationships

between FL and MeHg concentrations for each year (quadratic polynomial, except for a linear fit for lake whitefish in 2004), and using the resulting regression equations to estimate MeHg concentrations at a standard FL⁴³ (the mean FL of all fish sampled for each species: northern pike, 475 mm; lake whitefish, 530 mm). Square root transformation of raw northern pike data was required to satisfy assumptions of normality and homoscedasticity prior to standardization. The resulting data represent standardized concentrations of lake spike and ambient MeHg for each species each year.

We used the ratio of lake spike and ambient Hg in each sample as a measure of the amount by which Hg concentrations were changed with the addition of isotopically enriched Hg:

$$\frac{[\text{lake spike Hg}]_i}{[\text{ambient Hg}]_i} = \frac{[\text{lake spike Hg}]_i}{[\text{ambient Hg}]_i} \times 100\% \quad (1)$$

where $[\text{lake spike Hg}]_i$ is the concentration of lake spike MeHg in sample i , and $[\text{ambient Hg}]_i$ is the concentration of ambient MeHg in sample i . For northern pike and lake whitefish, we calculated the mean annual relative increase from all individuals (not the size-standardized concentration data).

Biomagnification factors (BMF) were calculated to describe differences in Hg concentrations between predator and prey⁵:

$$\text{BMF} = \log_{10} \left(\frac{[\text{MeHg}]_{\text{predator}}}{[\text{MeHg}]_{\text{prey}}} \right) \quad (2)$$

where $[\text{MeHg}]_{\text{predator}}$ is the mean (forage fish) or standardized (large-bodied fish) concentration of MeHg in the predator (ng g^{-1} w.w.) and $[\text{MeHg}]_{\text{prey}}$ is the mean concentration of MeHg in the prey (ng g^{-1} w.w.). MeHg concentration of prey items were averaged from samples collected throughout the open-water season immediately prior to autumn sampling of fish species to represent an integrated exposure for calculation of BMF. We used a dominant prey item to represent the diet of each fish species. For age 1+ yellow perch, northern pike, and lake whitefish, dominant prey items were zooplankton, forage fishes (YOY and 1+ yellow perch, and blacknose shiner) and *Chaoborus*, respectively.

To assess loss of lake spike MeHg by northern pike during the recovery period (2008–2015), we calculated²⁸ whole body burdens (in µg) of lake spike MeHg for the standardized population and for individuals that had been sampled in autumn 2007 (t_0 is the final time spike Hg was added to the lake) and again in at least one subsequent year during annual autumn sampling ($n = 16$ fish, of which 1–9 individuals were recaptured annually from 2008–2015). This calculation of MeHg burden is a relative measure of whole fish Hg content because MeHg is higher in muscle tissue than in other tissue types^{28,40}. For the standardized population data, we used best-fit relationships between FL (in mm) and body weight (in g; quadratic polynomial) to determine body weight at the standard FL. We multiplied this body weight by standard ambient and spike MeHg concentrations (in ng g⁻¹ w.w.) in muscle tissue for each year to determine body burdens over time (in ng). For individual fish, we multiplied spike MeHg concentration (in ng g⁻¹ w.w.) by body weight (in g) to yield individual body burdens (in ng). To account for differences among individuals and between individuals and the population, we normalized the data to examine the mean proportion of original (t_0) lake spike MeHg burden present in northern pike each year of the recovery period (2008–2015).

$$\text{\$}\$\{\text{rm}\{\text{change}\}\},\{\text{rm}\{\text{in}\}\},\{\text{rm}\{\text{burden}\}\},\{\text{rm}\{\text{from}\}\},\{t\}_0=\\ \{\text{rm}\{\text{burden}\}\}_0/\{\text{rm}\{\text{burden}\}\}_0\\ (3)$$

We used a best fit regression (exponential decay, beginning in the second year of recovery) to estimate the half-life (50% of original burden) of lake spike MeHg for the population.

Northern pike and lake whitefish ages were determined by cleithra and otoliths, respectively, if mortality had occurred, but most ages were quantified using fin rays collected from live fish⁴⁴ (K. H. Mills, DFO or North/South Consultants). Northern pike of the sizes selected for biopsy sampling had a median age of 3 years (range: 2–12 years; $n = 305$); the median age of lake whitefish was 17 years (range: 3–38 years; $n = 86$).

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

Datasets generated in this study are available at <https://doi.org/10.5061/dryad.nzs7h44sf>. [Source data](#) are provided with this paper.

References

1. 1. Streets, D. G. et al. Five hundred years of anthropogenic mercury: spatial and temporal release profiles. *Environ. Res. Lett.* **14**, 084044 (2019).
2. 2. Fitzgerald, W. F., Engstrom, D. R., Mason, R. P. & Nater, E. A. The case for atmospheric mercury contamination in remote areas. *Environ. Sci. Technol.* **32**, 1–7 (1998).
3. 3. Driscoll, C. T., Mason, R. P., Chan, H. M., Jacob, D. J. & Pirrone, N. Mercury as a global pollutant: sources, pathways, and effect. *Environ. Sci. Technol.* **47**, 4967–4983 (2013).
4. 4. Mergler, D. et al. Methylmercury exposure and health effects in humans: a worldwide concern. *Ambio* **36**, 3–11 (2007).
5. 5. Lavoie, R. A., Jardine, T. D., Chumchal, M. M., Kidd, K. A. & Campbell, L. M. Biomagnification of mercury in aquatic food webs: a worldwide meta-analysis. *Environ. Sci. Technol.* **47**, 13385–13394 (2013).
6. 6. Sunderland, E. M., Li, M. & Bullard, K. Decadal changes in the edible supply of seafood and methylmercury exposures in the United States. *Environ. Health Perspect.* **126**, 017006 (2018).
7. 7. Munthe, J. et al. Recovery of mercury-contaminated fisheries. *Ambio* **36**, 33–44 (2007).

8. 8.

Krabbenhoft, D. P. & Sunderland, E. M. Global change and mercury. *Science* **341**, 1457–1458 (2013).

9. 9.

Eagles-Smith, C. A. et al. Modulators of mercury risk to wildlife and humans in the context of rapid global change. *Ambio* **47**, 170–197 (2018).

10. 10.

Global Mercury Assessment 2013: Sources, Emissions, Releases and Environmental Transport,
http://www.unep.org/hazardoussubstances/Mercury/Informationmaterials/Reports_andPublications/tabcid/3593/Default.aspx (UNEP Chemicals Branch, 2013).

11. 11.

Clayden, M. G. et al. Mercury biomagnification through food webs is affected by physical and chemical characteristics of lakes. *Environ. Sci. Technol.* **47**, 12047–12053 (2013).

12. 12.

Vander Zanden, M. J., Casselman, J. M. & Rasmussen, J. B. Stable isotope evidence for the food web consequences of species invasions in lakes. *Nature* **401**, 464–467 (1999).

13. 13.

Kelly, E. N., Schindler, D. W., St. Louis, V. L., Donald, D. B. & Vladicka, K. E. Forest fire increases mercury accumulation by fishes via food web restructuring and increased mercury inputs. *Proc. Natl Acad. Sci. USA* **103**, 19380–19385 (2006).

14. 14.

Schartup, A. T. et al. Climate change and overfishing increase neurotoxicant in marine predators. *Nature* **572**, 648–650 (2019).

15. 15.

Lepak, R. F. et al. Mercury source changes and food web shifts alter contamination signatures of predatory fish from Lake Michigan. *Proc. Natl Acad. Sci. USA* **116**, 23600–23608 (2019).

16. 16.

Hall, B. D., Bodaly, R. A., Fudge, R. J. P., Rudd, J. W. M. & Rosenberg, D. M. Food as the dominant pathway of methylmercury uptake by fish. *Water Air Soil Pollut.* **100**, 13–24 (1997).

17. 17.

Guzzo, M. M., Blanchfield, P. J. & Rennie, M. D. Behavioral responses to annual temperature variation alter the dominant energy pathway, growth, and condition of a cold-water predator. *Proc. Natl Acad. Sci. USA*, **114**, 9912–9917 (2017).

18. 18.

Gårdmark, A. & Huss, M. Individual variation and interactions explain food web responses to global warming. *Phil. Trans. R. Soc. B* **375**, 20190449 (2020).

19. 19.

Harris, R. C. et al. Whole-ecosystem study shows rapid fish-mercury response to changes in mercury deposition. *Proc. Natl Acad. Sci. USA* **104**, 16586–16591 (2007).

20. 20.

Depew, D. C. et al. An overview of mercury concentrations in freshwater fish species: a national fish mercury dataset for Canada. *Can. J. Fish. Aquat. Sci.* **70**, 436–451 (2013).

21. 21.

St. Louis, V. L. et al. Atmospheric concentrations and wet/dry loadings of mercury at the remote Experimental Lakes Area, Northwestern Ontario, Canada. *Environ. Sci. Technol.* **53**, 8017–8026 (2019).

22. 22.

Graydon, J. A. et al. The role of terrestrial vegetation in atmospheric Hg deposition: Pools and fluxes of spike and ambient Hg from the METAALICUS experiment. *Global Biogeochem. Cycles* **26**, GB1022 (2012).

23. 23.
- Oswald, C. J., Heyes, A. & Branfireun, B. A. Fate and transport of ambient mercury and applied mercury isotope in terrestrial upland soils: insights from the METAALICUS watershed. *Environ. Sci. Technol.* **48**, 1023–1031 (2014).
24. 24.
- Eckley, C. S. & Hintelmann, H. Determination of mercury methylation potentials in the water column of lakes across Canada. *Sci. Total Environ.* **368**, 111–125 (2006).
25. 25.
- Sellers, P., Kelly, C. A. & Rudd, J. W. M. Fluxes of methylmercury to the water column of a drainage lake: The relative importance of internal and external sources. *Limnol. Oceanogr.* **46**, 623–631 (2001).
26. 26.
- Vander Zanden, M. J. & Vadeboncoeur, Y. Fishes as integrators of benthic and pelagic food webs in lakes. *Ecology* **83**, 2152–2161 (2002).
27. 27.
- Madenjian, C. P., Chipps, S. R. & Blanchfield, P. J. Time to refine mercury mass balance models for fish. *FACETS* **6**, 272–286 (2021).
28. 28.
- Van Walleghem, J. L. A., Blanchfield, P. J., Hrenchuk, L. E. & Hintelmann, H. Mercury elimination by a top predator, *Esox lucius*. *Environ. Sci. Technol.* **47**, 4147–4154 (2013).
29. 29.
- Knightes, C. D., Sunderland, E. M., Barber, M. C., Johnston, J. M. & Ambrose, R. B. Application of ecosystem-scale fate and bioaccumulation models to predict fish mercury response times to changes in atmospheric deposition. *Environ. Toxicol. Chem.* **28**, 881–893 (2009).
30. 30.

Chen, C. Y. et al. A critical time for mercury science to inform global policy. *Environ. Sci. Technol.* **52**, 9556–9561 (2018).

31. 31.

Blanchfield, P. J., Paterson, M. J., Shearer, J. A. & Schindler, D. W. Johnson and Vallentyne's legacy: 40 years of aquatic research at the Experimental Lakes Area. *Can. J. Fish. Aquat. Sci.* **66**, 1831–1836 (2009).

32. 32.

Sandilands, K. A. et al. *Application of Enriched Stable Mercury Isotopes to the Lake 658 Watershed for the METAALICUS Project, at the Experimental Lakes Area, Northwestern Ontario, Canada*. Can. Tech. Rep. Fish. Aquat. Sci. vol. 2597 (Fisheries and Oceans Canada, 2005).

33. 33.

Sandilands, K. A. et al. *Application of Enriched Stable Mercury Isotopes to the Lake 658 Watershed for the METAALICUS Project, at the Experimental Lakes Area, Northwestern Ontario, Canada, 2001–2007*. Can. Tech. Rep. Fish. Aquat. Sci. vol. 2813 (Fisheries and Oceans Canada, 2008).

34. 34.

Bower, P. M. et al. Simultaneous measurement of primary production by whole-lake and bottle radiocarbon additions. *Limnol. Oceanogr.* **32**, 299–312 (1987).

35. 35.

Golding, G. R. et al. Evidence for facilitated uptake of Hg(ii) by *Vibrio anguillarum* and *Escherichia coli* under anaerobic and aerobic conditions. *Limnol. Oceanogr.* **47**, 967–975 (2002).

36. 36.

Southworth, G. et al. Evasion of added isotopic mercury from a northern temperate lake. *Environ. Technol. Chem.* **26**, 53–60 (2007).

37. 37.

Baker, R. F., Blanchfield, P. J., Paterson, M. J., Flett, R. J. & Wesson, L. Evaluation of non-destructive biopsy methods for the analysis of mercury in fish tissue. *Trans. Amer. Fish. Soc.* **133**, 568–576 (2004).

38. 38.

Paterson, M. J. et al. Bioaccumulation of newly-deposited mercury by fish and invertebrates: an enclosure study using stable mercury isotopes. *Can. J. Fish. Aquat. Sci.* **63**, 2213–2224 (2006).

39. 39.

Hintelmann, H. & Ogrinc, N. in *Biogeochemistry of Environmentally Important Trace Elements* (eds Cai, Y. & Braids, O. C.) 321–338 (ACS, 2003).

40. 40.

Van Walleghem, J. L. A., Blanchfield, P. J. & Hintelmann, H. Elimination of mercury by yellow perch in the wild. *Environ. Sci. Technol.* **41**, 5895–5901 (2007).

41. 41.

Hrenchuk, L. E., Blanchfield, P. J., Paterson, M. J. & Hintelmann, H. H. Dietary and waterborne mercury accumulation by yellow perch: a field experiment. *Environ. Sci. Technol.* **46**, 509–516 (2012).

42. 42.

Howmiller, R. P. Effects of preservatives on weight of some common macrobenthic invertebrates. *Trans. Amer. Fish. Soc.* **101**, 743–746 (1972).

43. 43.

Tremblay, G., Legendre, P., Doyon, J.-F., Verdon, R. & Schetagne, R. The use of polynomial regression analysis with indicator variables for interpretation of mercury in fish data. *Biogeochemistry* **40**, 189–201 (1998).

44. 44.

Mills, K. H. & Beamish, R. J. Comparison of fin-ray and scale age determinations for lake whitefish (*Coregonus clupeaformis*) and their implications for estimates of growth and annual survival. *Can. J. Fish. Aquat. Sci.* **37**, 534–544 (1980).

Acknowledgements

We thank A. Robinson for aircraft support, M. Stainton and R. Hesslein for the preparation of mercury isotopes, M. Lyng, J. T. Bell, R. Burrell, B. Dimock, M. Dobrin, K. Fetterly, C. Miller, M. Rearick, J. Reid, G. Riedel, J. Shead, J. Zhou and staff and students at the Experimental Lakes Area for assistance with field collections and analysis of samples. K. Kidd, D. Orihel and J. Smol provided reviews of this manuscript. We thank the following agencies for their financial support for this study: Electric Power Research Institute, Environment and Climate Change Canada, Fisheries and Oceans Canada, Natural Sciences and Engineering Research Council of Canada, National Science Foundation Grant DEB 0451345 (C.C.G.) and DEB 0351050 (A.H. and C.C.G.), Southern Company, University of Alberta, U.S. Department of Energy, U.S. Environmental Protection Agency, and Wisconsin Focus on Energy Program Project Grant no. 4900-02-03 (J.P.H.). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Author information

Author notes

1. John W. M. Rudd & Carol A. Kelly

Present address: R&K Research, Salt Spring Island, British Columbia, Canada

2. Lori S. Tate

Present address: Wisconsin Department of Natural Resources, Madison, WI, USA

Affiliations

1. Fisheries and Oceans Canada, Freshwater Institute, Winnipeg, Manitoba, Canada

Paul J. Blanchfield, John W. M. Rudd, Lee E. Hrenchuk, Ken G. Beaty, R. A. Drew Bodaly, Carol A. Kelly, Michael J. Paterson, Cheryl L. Podemski, Ken A. Sandilands & Lori S. Tate

2. Department of Biology, Queen's University, Kingston, Ontario, Canada

Paul J. Blanchfield

3. IISD Experimental Lakes Area, Winnipeg, Manitoba, Canada

Paul J. Blanchfield, Lee E. Hrenchuk, Michael J. Paterson & Ken A. Sandilands

4. Département de Sciences Biologiques, Université de Montréal, Montreal, Quebec, Canada
Marc Amyot
5. Environmental Chemistry and Technology Program, University of Wisconsin-Madison, Madison, WI, USA
Christopher L. Babiarz
6. Department of Biology, Biological and Geological Sciences Building, University of Western Ontario, London, Ontario, Canada
Brian A. Branfireun
7. Smithsonian Environmental Research Center, Edgewater, MD, USA
Cynthia C. Gilmour
8. Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada
Jennifer A. Graydon & Vincent L. St Louis
9. Department of Biology, University of Regina, Regina, Saskatchewan, Canada
Britt D. Hall
10. Reed Harris Environmental, Oakville, Ontario, Canada
Reed C. Harris
11. University of Maryland Center for Environmental Science, Chesapeake Biological Laboratory, Solomons, MD, USA
Andrew Heyes
12. Water Quality Center, Trent University, Peterborough, Ontario, Canada
Holger Hintemann
13. University of Wisconsin-Madison, Department of Civil and Environmental Engineering, Environmental Chemistry and Technology Program, Madison, WI, USA

James P. Hurley

14. US Geological Survey, Middleton, WI, USA

David P. Krabbenhoft & Michael T. Tate

15. Oak Ridge National Laboratory, Oak Ridge, TN, USA

Steve E. Lindberg & George R. Southworth

16. Department of Marine Sciences, University of Connecticut, Groton, CT, USA

Robert P. Mason

Contributions

P.J.B., J.W.M.R., M.A., K.G.B., R.A.B., B.A.B., C.C.G., R.C.H., A.H., H.H., J.P.H., C.A.K., D.P.K., S.E.L., R.P.M., M.J.P., C.L.P. and V.L.S.L. contributed to the design of the whole-ecosystem experiment. The METAALICUS project was overseen by J.W.M.R. and R.C.H., who along with C.L.B., R.A.B., J.A.G., H.H., J.P.H., C.A.K., D.P.K., K.A.S., V.L.S.L. and M.T.T., applied mercury to the lake and watershed. P.J.B., L.E.H. and L.S.T. conducted the fish sampling. Water, sediment and lower trophic level data were collected and prepared by C.C.G., B.D.H., H.H., C.A.K., M.J.P., C.L.P. and K.A.S. Field samples in this study were analysed for mercury by C.C.G., H.H., D.P.K. and M.T.T. Data analyses were performed by P.J.B. and L.E.H. All authors collected and discussed project-level data that contributed to the interpretation of the data presented in this study. L.E.H. produced the figures and wrote the methods. P.J.B., J.W.M.R., C.A.K., and V.L.S.L. wrote the manuscript with input from all authors.

Corresponding author

Correspondence to [Paul J. Blanchfield](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review information

Nature thanks Bridget Bergquist and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

[Extended Data Fig. 1 Temporal dynamics of upland mercury in fish.](#)

Isotopic ^{198}Hg added to the upland area of Lake 658 was above the detection limit (0.5% of ambient MeHg; dashed line) in all fish species, but contributed little (<2%) to overall MeHg concentrations (percent increase = [upland spike MeHg]/[ambient MeHg] \times 100). Mean annual MeHg concentration data for each species or age-class is presented and based on fish collected during fall population sampling ($n = 1,052$; sample size details in Extended Data Tables [2](#) and [3](#)); dotted lines indicate missing data

[Extended Data Fig. 2 Comparison of individual and population body sizes of northern pike.](#)

Mean (\pm s.e.m.) body size of all northern pike sampled in the fall of each year (population, black circles; $n = 442$) for muscle MeHg concentration using a biopsy method was stable over time. Individual northern pike (grey triangles) captured in 2007 and again in at least one subsequent year ($n = 16$ fish with 1–9 individuals recaptured each year 2008–2015) were used to determine individual losses of lake spike MeHg during recovery (see Fig. [3](#)). These individual fish, which were also captured prior to 2007, showed an increase in body size over time (linear regression: $y = 17.43x - 34470.0$, $R^2 = 0.55$, $F_{1,51} = 61.1$, $P < 0.0001$)

Extended Data Table 1 Fish mercury concentrations in the experimental and reference lakes over time

Extended Data Table 2 Annual fish metrics and mercury concentrations during the addition phase

Extended Data Table 3 Annual fish metrics and mercury concentrations during the recovery phase

Supplementary information

Reporting Summary

Peer Review File

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Extended Data Fig. 1

Source Data Extended Data Fig. 2

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Reprints and Permissions

About this article

Cite this article

Blanchfield, P.J., Rudd, J.W.M., Hrenchuk, L.E. *et al.* Experimental evidence for recovery of mercury-contaminated fish populations. *Nature* **601**, 74–78 (2022). <https://doi.org/10.1038/s41586-021-04222-7>

- Received: 15 January 2021
- Accepted: 09 November 2021
- Published: 15 December 2021
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04222-7>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04222-7>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 01 December 2021](#)

Cyclic evolution of phytoplankton forced by changes in tropical seasonality

- [Luc Beaufort](#) [ORCID: orcid.org/0000-0001-6055-9373¹](#),
- [Clara T. Bolton](#) [ORCID: orcid.org/0000-0002-3078-1253¹](#),
- [Anta-Clarisse Sarr](#) [ORCID: orcid.org/0000-0002-9495-5355¹](#),
- [Baptiste Suchéras-Marx](#) [ORCID: orcid.org/0000-0002-8107-8242¹](#),
- [Yair Rosenthal](#) [ORCID: orcid.org/0000-0002-7546-6011²](#),
- [Yannick Donnadieu¹](#),
- [Nicolas Barbarin](#) [ORCID: orcid.org/0000-0001-8641-0649¹](#) nAff3,
- [Samantha Bova](#) [ORCID: orcid.org/0000-0002-5064-8775²](#) nAff4,
- [Pauline Cornuault](#) [ORCID: orcid.org/0000-0002-6697-4735¹](#) nAff5,
- [Yves Gally¹](#),
- [Emmeline Gray](#) [ORCID: orcid.org/0000-0002-1664-4930¹](#) nAff6,
- [Jean-Charles Mazur¹](#) &
- [Martin Tetard](#) [ORCID: orcid.org/0000-0003-0487-1949¹](#)

[Nature](#) volume **601**, pages 79–84 (2022)

- 4248 Accesses
- 1 Citations
- 402 Altmetric
- [Metrics details](#)

Subjects

- [Palaeoceanography](#)
- [Palaeontology](#)

Abstract

Although the role of Earth's orbital variations in driving global climate cycles has long been recognized, their effect on evolution is hitherto unknown. The fossil remains of coccolithophores, a key calcifying phytoplankton group, enable a detailed assessment of the effect of cyclic orbital-scale climate changes on evolution because of their abundance in marine sediments and the preservation of their morphological adaptation to the changing environment^{1,2}. Evolutionary genetic analyses have linked broad changes in Pleistocene fossil coccolith morphology to species radiation events³. Here, using high-resolution coccolith data, we show that during the last 2.8 million years the morphological evolution of coccolithophores was forced by Earth's orbital eccentricity with rhythms of around 100,000 years and 405,000 years—a distinct spectral signature to that of coeval global climate cycles⁴. Simulations with an Earth System Model⁵ coupled with an ocean biogeochemical model⁶ show a strong eccentricity modulation of the seasonal cycle, which we suggest directly affects the diversity of ecological niches that occur over the annual cycle in the tropical ocean. Reduced seasonality in surface ocean conditions favours species with mid-size coccoliths, increasing coccolith carbonate export and burial; whereas enhanced seasonality favours a larger range of coccolith sizes and reduced carbonate export. We posit that eccentricity pacing of phytoplankton evolution contributed to the strong 405,000-year cyclicity that is seen in global carbon cycle records.

This is a preview of subscription content

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Noelaerhabdaceae coccolith morphology and accumulation, eccentricity, and climate over the last 2.8 million years.

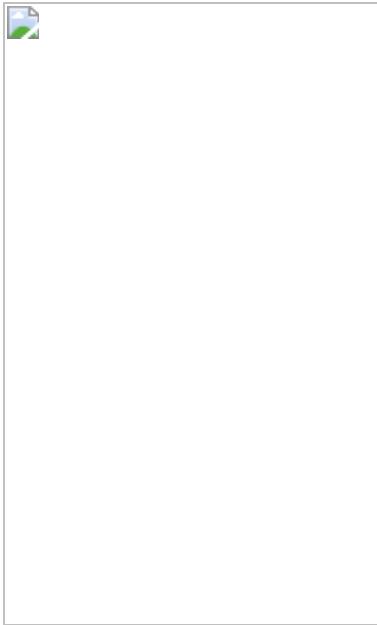


Fig. 2: MDI concept.

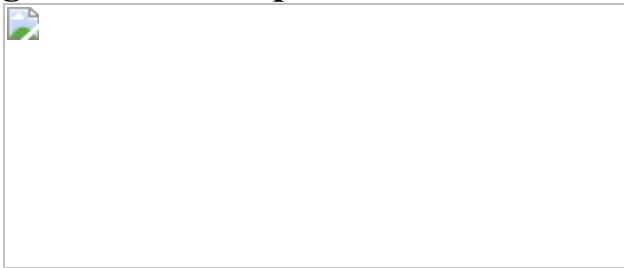


Fig. 3: Modelled NPP seasonal contrast under different eccentricity configurations and MDI.



Data availability

All coccolith morphological data, as well as all model outputs described in the paper (including NPP and main oceanic and atmospheric variables) are archived at the SEANOE open access data repository:

<https://doi.org/10.17882/84031>. LMDZ, XIOS, NEMO and ORCHIDEE are released under the terms of the CeCILL license. OASIS-MCT is released under the terms of the Lesser GNU General Public License (LGPL). IPSLCM5A2 source code is publicly available through svn, with the following commands line : svn co

http://forge.ipsl.jussieu.fr/igcmg/svn/modipsl/branches/publications/IPSLCM5A2.1_11192019 modipsl ; cd modipsl/util ; ./model IPSLCM5A2.1. The mod.def file provides information regarding the different revisions used, namely: NEMOGCM branch nemo_v3_6_STABLE revision 6665; XIOS2 branches/xios-2.5 revision 1763; IOIPSL/src svn tags/v2_2_2; LMDZ5 branches/IPSLCM5A2.1 rev 3591; branches/publications/ORCHIDEE_IPSLCM5A2.1.r5307 rev 6336; and OASIS3-MCT 2.0_branch (rev 4775 IPSL server). The login/password combination requested at first use to download the ORCHIDEE component is anonymous/anonymous. We recommend that you refer to the project website:

http://forge.ipsl.jussieu.fr/igcmg_doc/wiki/Doc/Config/IPSLCM5A2 for a proper installation and compilation of the environment.

References

1. 1.

Beaufort, L. et al. Sensitivity of coccolithophores to carbonate chemistry and ocean acidification. *Nature* **476**, 80–84 (2011).

2. 2.

Henderiks, J. & Bollmann, J. The *Gephyrocapsa* sea surface palaeothermometer put to the test: comparison with alkenone and foraminifera proxies off NW Africa. *Mar. Micropaleontol.* **50**, 161–184 (2004).

3. 3.

Bendif, E. M. et al. Repeated species radiations in the recent evolution of the key marine phytoplankton lineage *Gephyrocapsa*. *Nat. Commun.* **10**, 4234 (2019).

4. 4.

Ruddiman, W. F. & Raymo, M. E. Northern Hemisphere climate regimes during the past 3 Ma: possible tectonic connections. *Phil. Trans. R. Soc. B. Biol. Sci.* **318**, 411–430 (1988).

5. 5.

Sepulchre, P. et al. IPSL-CM5A2—an Earth System Model designed for multi-millennial climate simulations. *Geosci. Model Dev.* **13**, 3011–3053 (2020).

6. 6.

Aumont, O., Éthé, C., Tagliabue, A., Bopp, L. & Gehlen, M. PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies. *Geosci. Model Dev.* **8**, 2465–2513 (2015).

7. 7.

Broecker, W. & Clark, E. Ratio of coccolith CaCO_3 to foraminifera CaCO_3 in late Holocene deep sea sediments. *Paleoceanography* **24**, PA3205 (2009).

8. 8.

Suchéras-Marx, B. et al. The colonization of the oceans by calcifying pelagic algae. *Biogeosciences* **16**, 2501–2510 (2019).

9. 9.

Ridgwell, A. & Zeebe, R. E. The role of the global carbonate cycle in the regulation and evolution of the Earth system. *Earth Planet. Sci. Lett.* **234**, 299–315 (2005).

10. 10.

Young, J. et al. A guide to extant coccolithophore taxonomy. *Journal of Nannoplankton Research, Special Issue* **1**, 1–132 (2003).

11. 11.

Henderiks, J. Coccolithophore size rules—reconstructing ancient cell geometry and cellular calcite quota from fossil coccoliths. *Mar. Micropaleontol.* **67**, 143–154 (2008).

12. 12.

Filatov, D. A. Extreme Lewontin’s paradox in ubiquitous marine phytoplankton species. *Mol. Biol. Evol.* **36**, 4–14 (2019).

13. 13.

Bown, P. R. Calcareous nannoplankton evolution: a tale of two oceans. *Micropaleontology* **51**, 299–308 (2005).

14. 14.

Si, W. & Rosenthal, Y. Reduced continental weathering and marine calcification linked to late Neogene decline in atmospheric CO₂. *Nat. Geosci.* **12**, 833–838 (2019).

15. 15.

Bolton, C. T. et al. Decrease in coccolithophore calcification and CO₂ since the middle Miocene. *Nat. Commun.* **7**, 10284 (2016).

16. 16.

Gibbs, S. J., Shackleton, N. & Young, J. Orbitally forced climate signals in mid-Pliocene nannofossil assemblages. *Mar. Micropaleontol.* **51**, 39–56 (2004).

17. 17.

Beaufort, L. et al. Insolation cycles as a major control of equatorial Indian Ocean primary production. *Science* **278**, 1451–1454 (1997).

18. 18.

Chepstow-Lusty, A., Backman, J. & Shackleton, N. J. Comparison of upper Pliocene Discoaster abundance variations from North Atlantic Sites 552, 607, 658, 659 and 662: further evidence for marine plankton responding to orbital forcing. *Proc. ODP Sci. Results* **108**, 121–141 (1989).

19. 19.

Thierstein, H. R., Geitzenauer, K. R., Molfino, B. & Shackleton, N. J. Global synchronicity of late Quaternary coccolith datum levels : validation by oxygen isotopes. *Geology* **5**, 400–404 (1977).

20. 20.

Bollmann, J., Baumann, K. H. & Thierstein, H. R. Global dominance of *Gephyrocapsa* coccoliths in the Late Pleistocene: selective dissolution, evolution, or global environmental change? *Paleoceanography* **13**, 517–529 (1998).

21. 21.

Raffi, I. et al. A review of calcareous nannofossil astrobiochronology encompassing the past 25 million years. *Quat. Sci. Rev.* **25**, 3113–3137 (2006).

22. 22.

Lisiecki, L. E. & Raymo, M. E. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}\text{O}$ records. *Paleoceanography* **20**, PA1003 (2005).

23. 23.

Laskar, J. et al. A long-term numerical solution for the insolation quantities of the Earth. *Astron. Astrophys.* **428**, 261–285 (2004).

24. 24.

Okada, H. & McIntyre, A. Seasonal distribution of modern Coccolithophores in the western north Atlantic ocean. *Mar. Biol.* **54**, 319–328 (1979).

25. 25.

Longhurst, A. *Ecological Geography of the Sea* Vol. 1 (Academic Press, 1998).

26. 26.

Sexton, P. F. & Barker, S. Onset of ‘Pacific-style’ deep-sea sedimentary carbonate cycles at the mid-Pleistocene transition. *Earth Planet. Sci. Lett.* **321–322**, 81–94 (2012).

27. 27.

Rickaby, R. E. M. et al. Coccolith chemistry reveals secular variations in the global ocean carbon cycle? *Earth Planet. Sci. Lett.* **253**, 83–95 (2007).

28. 28.

Wang, P., Tian, J. & Lourens, L. J. Obscuring of long eccentricity cyclicity in Pleistocene oceanic carbon isotope records. *Earth Planet. Sci. Lett.* **290**, 319–330 (2010).

29. 29.

Suchéras-Marx, B. & Henderiks, J. Downsizing the pelagic carbonate factory: impacts of calcareous nannoplankton evolution on carbonate burial over the past 17 million years. *Glob. Planet. Change* **123**, 97–109 (2014).

30. 30.

Holligan, P. M. et al. A biogeochemical study of the coccolithophore, *Emiliania huxleyi*, in the North Atlantic. *Global Biogeochem. Cycles* **7**, 879–900 (1993).

31. 31.

Laepple, T. & Lohmann, G. Seasonal cycle as template for climate variability on astronomical timescales. *Paleoceanography* **24**, PA4201 (2009).

32. 32.

Righetti, D., Vogt, M., Gruber, N., Psomas, A. & Zimmermann, N. E. Global pattern of phytoplankton diversity driven by temperature and environmental variability. *Sci. Adv.* **5**, eaau6253 (2019).

33. 33.

Ruddiman, W. F., Raymo, M. & McIntyre, A. Matuyama 41,000-year cycles : North Atlantic Ocean and Northern hemisphere ice sheets. *Earth Plan. Sci. Lett.* **80**, 117–129 (1986).

34. 34.

Zeebe, R. E., Westerhold, T., Littler, K. & Zachos, J. C. Orbital forcing of the Paleocene and Eocene carbon cycle. *Paleoceanography* **32**, 440–465 (2017).

35. 35.

Pälike, H. et al. The heartbeat of the Oligocene climate system. *Science* **314**, 1894–1898 (2006).

36. 36.

Herbert, T. D. A long marine history of carbon cycle modulation by orbital-climatic changes. *Proc. Natl Acad. Sci. USA* **94**, 8362–8369 (1997).

37. 37.

Ma, W., Tian, J., Li, Q. & Wang, P. Simulation of long eccentricity (400-kyr) cycle in ocean carbon reservoir during Miocene Climate Optimum: weathering and nutrient response to orbital change. *Geophys. Res. Lett.* **38**, L10701 (2011).

38. 38.

Archer, D. & Maier-Reimer, E. Effect of deep-sea sedimentary calcite preservation on atmospheric CO₂ concentration. *Nature* **367**, 260–263 (1994).

39. 39.

Beaufort, L. Adaptation of the random settling method for quantitative studies of calcareous nannofossils. *Micropaleontology* **37**, 415–418 (1992).

40. 40.

Beaufort, L., Barbarin, N. & Gally, Y. Optical measurements to determine the thickness of calcite crystals and the mass of thin carbonate particles such as coccoliths. *Nat. Protoc.* **9**, 633–642 (2014).

41. 41.

Beaufort, L. & Dollfus, D. Automatic recognition of coccolith by dynamical neural network. *Mar. Micropaleont.* **51**, 57–73 (2004).

42. 42.

Barbarin, N. *La Reconnaissance Automatisée des Nannofossiles Calcaires du Cénozoïque*. PhD thesis, Aix-Marseille Univ. (2014).

43. 43.

Beaufort, L. Weight estimates of coccoliths using the optical properties (birefringence) of calcite. *Micropaleontology* **51**, 289–298 (2005).

44. 44.

Dollfus, D. *Reconnaissance des Formes Naturelles par des Réseaux de Neurones Artificiels: Application au Nannoplancton Calcaire*. PhD thesis, Aix-Marseille Univ. (1997).

45. 45.

Dollfus, D. & Beaufort, L. Fat neural network for recognition of position-normalised objects. *Neural Netw.* **12**, 553–560 (1999).

46. 46.

Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).

47. 47.

de Garidel-Thoron, T., Rosenthal, Y., Bassinot, F. C. & Beaufort, L. Stable sea surface temperatures in the western Pacific warm pool over the past 1.75 million years. *Nature* **433**, 294–298 (2005).

48. 48.

Shackleton, N. J. & Al, E. An alternative astronomical calibration of the lower Pleistocene timescale based on ODP Site 677. *Trans. R. Soc. Edinb. Earth Sci.* **81**, 251–261 (1990).

49. 49.

Carcaillet, J. T., Thouveny, N. & Bourlès, D. L. Geomagnetic moment instability between 0.6 and 1.3 Ma from cosmonuclide evidence.

Geophys. Res. Lett. **30**, 1792 (2003).

50. 50.

Beaufort, L. & Party, S. S. *MD148/PECTEN* (Institut Polaire Francais, 2005).

51. 51.

Tachikawa, K. et al. The precession phase of hydrological variability in the Western Pacific Warm Pool during the past 400 ka. *Quat. Sci. Rev.* **30**, 3716–3727 (2011).

52. 52.

Regoli, F. et al. Progressive shoaling of the equatorial Pacific thermocline over the last eight glacial periods. *Paleoceanography* **30**, 439–455 (2015).

53. 53.

Clemens, S. C. et al. *Indian Monsoon Rainfall*. International Ocean Discovery Program Preliminary Report 353 (IODP Publications, 2015).

54. 54.

Clemens, S. C. et al. Site U1443. In *Proc. International Ocean Discovery Program* Vol. 353, <https://doi.org/10.14379/iodp.proc.353.103.2016> (2016).

55. 55.

Gebregiorgis, D. et al. Southern Hemisphere forcing of South Asian monsoon precipitation over the past ~1 million years. *Nat. Commun.* **9**, 4702 (2018).

56. 56.

Gebregiorgis, D. et al. What can we learn from X-ray fluorescence core scanning data? A paleomonsoon case study. *Geochem. Geophys. Geosyst.* **21**, e2019GC008414 (2020).

57. 57.

McGrath, S. M., Clemens, S. C., Huang, Y. & Yamamoto, M. Greenhouse gas and ice volume drive Pleistocene Indian summer monsoon precipitation isotope variability. *Geophys. Res. Lett.* **48**, e2020GL092249 (2021).

58. 58.

Rosenthal, Y. et al. Site U1485. In *Proc. International Ocean Discovery Program Vol. 363*, <https://doi.org/10.14379/iodp.proc.363.106.2018> (2018).

59. 59.

Beaufort, L., Bassinot, F. C. & Vincent, E. in *Reconstructing Ocean History : a Window into the Future* (eds Abrantes, F. & Mix, A. C.) 245–272 (Kluwer Academic/Plenum Publisher, 1999).

60. 60.

Young, J. R., P.R., Bown & J.A., Lees *Nannotax3* <http://www.mikrotax.org/Nannotax3> (2021).

61. 61.

Perch-Nielsen, K. in *Plankton Stratigraphy* Vol. 1 (eds Bolli, H. M., Saunders, J. B. & Perch-Nielsen, K.) (CUP Archive, 1989).

62. 62.

Beaufort, L., Probert, I. & Buchet, N. Effects of acidification and primary production on coccolith weight: Implications for carbonate transfer from the surface to the deep ocean. *Geochem. Geophys. Geosystems* **8**, Q08011 (2007).

63. 63.

Lyle, M. Neogene carbonate burial in the Pacific Ocean.
Paleoceanography **18**, 1059 (2003).

64. 64.

Paillard, D., Labeyrie, L. & Yiou, P. Macintosh program performs time-series analysis. *Eos* **77**, 379 (1996).

65. 65.

Li, M., Hinnov, L. & Kump, L. Acycle: time-series analysis software for paleoclimate research and education. *Comput. Geosci.* **127**, 12–22 (2019).

66. 66.

Blackman, R. B. & Tukey, J. W. The measurement of power spectra from the point of view of communications engineering—Part I. *Bell Syst. Tech. J.* **37**, 185–282 (1958).

67. 67.

Thomson, D. J. Spectrum estimation and harmonic analysis. In *Proc. IEEE* Vol. 70, 1055–1096 (1982).

68. 68.

Bollmann, J. Morphology and biogeography of the *Gephyrocapsa* coccoliths in Holocene sediments. *Mar. Micropaleontol.* **29**, 319–350 (1997).

69. 69.

Hourdin, F. et al. Impact of the LMDZ atmospheric grid configuration on the climate and sensitivity of the IPSL-CM5A coupled model. *Clim. Dyn.* **40**, 2167–2192 (2013).

70. 70.

Krinner, G., Ciais, P., Viovy, N. & Friedlingstein, P. A simple parameterization of nitrogen limitation on primary productivity for global vegetation models. *Biogeosciences Discuss.* **2**, 1243–1282 (2005).

71. 71.

Madec, G. & Nemo, T. *NEMO Ocean Engine*. Note du Pôle de modélisation de l’Institut Pierre-Simon Laplace no. 27 (IPSL, 2015).

72. 72.

Madec, G. *NEMO Reference Manual, Ocean Dynamics Component: NEMO-OPA. Preliminary Version*. Note du Pole de modélisation de l’Institut Pierre-Simon Laplace (IPSL, 2008).

73. 73.

Fichefet, T. & Maqueda, M. M. Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics. *J. Geophys. Res. Oceans* **102**, 12609–12646 (1997).

74. 74.

Valcke, S. et al. Coupling technologies for earth system modelling. *Geosci. Model Dev.* **5**, 1589–1596 (2012).

75. 75.

Dufresne, J.-L. et al. Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Clim. Dyn.* **40**, 2123–2165 (2013).

76. 76.

Mayorga, E. et al. Global nutrient export from WaterSheds 2 (NEWS 2): model development and implementation. *Environ. Modelling Softw.*

25, 837–853 (2010).

77. 77.

Bosmans, J. et al. Response of the Asian summer monsoons to idealized precession and obliquity forcing in a set of GCMs. *Quat. Sci. Rev.* **188**, 121–135 (2018).

78. 78.

Braconnot, P. & Marti, O. Impact of precession on monsoon characteristics from coupled ocean atmosphere experiments: changes in Indian monsoon and Indian ocean climatology. *Mar. Geol.* **201**, 23–34 (2003).

79. 79.

Prescott, C., Haywood, A., Dolan, A., Hunter, S. & Tindall, J. Indian monsoon variability in response to orbital forcing during the late Pliocene. *Glob. Planet. Change* **173**, 33–46 (2019).

80. 80.

Erb, M. P. et al. Response of the equatorial pacific seasonal cycle to orbital forcing. *J. Clim.* **28**, 9258–9276 (2015).

81. 81.

Chen, M.-T. & Beaufort, L. Exploring quaternary variability of the east Asia monsoon, Kuroshio Current, and western Pacific warm pool systems: high-resolution investigations of paleoceanography from the IMAGES III (MD106) IPHIS cruise. *Terr. Atmos. Ocean. Sci.* **9**, 129–142 (1998).

82. 82.

Gartner, S. Paleoceanography of the mid-Pleistocene. *Mar. Micropaleontol.* **13**, 23–46 (1988).

83. 83.

Pujos, A. & Giraudeau, J. Distribution of Noelaerhabdaceae (calcareous nannofossils) in the upper and middle Quaternary of the Atlantic and Pacific oceans. *Oceanolog. Acta* **16**, 349–362 (1993).

Acknowledgements

This paper is a contribution of the Climate research group at CEREGE. This research uses samples provided by the IODP. We thank the scientists, technical staff and crews of IODP expeditions 353 and 363 and IMAGES expeditions 3 and 13; A. Fruy and S. Sergi for sample preparation assistance; and the CEA–CCRT for providing access to the HPC resources of TGCC under the allocation 2019-A0070102212 made by GENCI. We acknowledge French ANR projects CALHIS (L.B.), iMonsoon (C.T.B.) and AMOR (Y.D.), and INSU project CALVE (C.T.B.) and FRB project COCCACE (L.B.) , which provided funding for this work. IODP France provided post-cruise funding to L.B. and C.T.B.

Author information

Author notes

1. Nicolas Barbarin

Present address: TOTAL S.A., Pau, France

2. Samantha Bova

Present address: Department of Geological Sciences, San Diego State University, San Diego, CA, USA

3. Pauline Cornuault

Present address: MARUM, Universität Bremen, Bremen, Germany

4. Emmeline Gray

Present address: The Open University, Milton Keynes, UK

Affiliations

1. Aix-Marseille University, CNRS, IRD, INRAE, CEREGE, Aix-en-Provence, France

Luc Beaufort, Clara T. Bolton, Anta-Clarissee Sarr, Baptiste Suchéras-Marx, Yannick Donnadieu, Nicolas Barbarin, Pauline Cornuault, Yves Gally, Emmeline Gray, Jean-Charles Mazur & Martin Tetard

2. Department of Marine and Coastal Sciences, and Earth and Planetary Sciences, Rutgers, State University of New Jersey, New Brunswick, NJ, USA

Yair Rosenthal & Samantha Bova

Contributions

L.B. designed the study. L.B., Y.G., N.B. and M.T. developed automated artificial intelligence methods. L.B., C.T.B., J.-C.M., P.C., E.G. and S.B. prepared samples and/or generated data. A.-C.S. designed and ran the model simulations, in collaboration with Y.D. L.B. and C.T.B. analysed the morphometric data. L.B., C.T.B., A.-C.S., B.S.-M., Y.D. and Y.R. discussed interpretations. L.B., C.T.B. and A.-C.S. wrote the manuscript with contributions from B.S.-M., Y.D. and Y.R.

Corresponding authors

Correspondence to [Luc Beaufort](#) or [Clara T. Bolton](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Claudia Agnini, Ying Guan, Rosalind Rickaby, Andy Ridgwell and Thomas Westerhold for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

[Extended Data Fig. 1 Size and MDI records for each core.](#)

Top panels **a** to **i**: Size frequency plots for each individual core used to create the composite record shown in Fig. [1a](#). Bottom panels **a** to **i**: Individual MDI records for each core (black lines and points) plotted with the stacked MDI record (red line). Pearson correlation coefficients between individual sites and the stack vary between 0.71 and 0.93 (*P* values are all <0.00001).

[Extended Data Fig. 2 Time-series analyses of individual records.](#)

a, b, Blackman Tukey cross-spectral analysis between eccentricity and stacked coccolith length (**a**) and eccentricity and stacked MDI (**b**). Top: coherency; Bottom: phase (radian). **c–k**, MTM and evolutive spectral analyses (see [Methods](#)) of detrended individual MDI series resampled at 2-kyr intervals (shown on left of each evolutive analysis). Primary orbital periods are shown by red lines.

[Extended Data Fig. 3 Decomposition of the Noelaerhabdaceae mass accumulation rate \(NoMAR\) record into its mass and flux components.](#)

a, Stacked NoMAR record, binned into 2-ky intervals (orange shading) and smoothed with a 30-kyr moving window (orange line), **b**, Noelaerhabdaceae coccolith flux (blue) and average Noelaerhabdaceae coccolith mass (red). Here, stacked mass and flux records are smoothed with a 30-kyr moving window as in **a**. **c**, NoMAR (orange) and MDI (purple) records, smoothed with a 30-kyr moving window. Grey shaded areas represent four described acmes of mid-size Noelaerhabdaceae species^{[19,20,21,82,83](#)}.

Extended Data Fig. 4 Ocean–atmosphere model outputs under different orbital configurations.

Top: Yearly maximum contrast in NPP ($\text{gC m}^{-2} \text{ day}^{-1}$) for **a**: $E_{\min}P_{\min}$, **b**: $E_{\max}P_{\min}$ and **c**: $E_{\max}P_{\max}$. Low eccentricity values minimize the amplitude of precession variability, thus we only show results for minimum precession value at minimum eccentricity ($E_{\min}P_{\min}$) but the reader can consider those results to be similar for the $E_{\min}P_{\max}$ simulation. **d** and **e** represent the anomaly of yearly maximum contrast in NPP. At $E_{\max}P_{\max}$, the eastern equatorial Indian Ocean exhibits moderate seasonality (**a**) due to inhibition of the summer productivity induced by lower nutrient concentrations in this area (Extended Data Fig. [6a](#)). In this case, high productivity areas during boreal summer are shifted to south-west of India. Bottom: Late summer (JASO) low-level winds for **f**: $E_{\min}P_{\min}$, **g**: $E_{\max}P_{\min}$, **h**: $E_{\max}P_{\max}$ simulations. **i** and **j** represent the anomaly in late summer low-level winds. At $E_{\max}P_{\max}$ the north-equatorial westerlies (**c**, **e**) are confined to south of 10° N owing to the extension above India of the low-pressure area.

Extended Data Fig. 5 Solar radiation and sea-level pressure in model simulations.

Seasonal latitudinal variations of solar radiation at the top of the atmosphere derived from the model (W.m^{-2}); **a**: $E_{\min}P_{\min}$, **b**: $E_{\max}P_{\min}$, **c**: $E_{\max}P_{\max}$. See Extended Data Table [3](#) for details of orbital configurations of each simulation. Late summer (JASO) low-level winds for **d**: $E_{\min}P_{\min}$, **e**:

$E_{\max}P_{\min}$, **f**: $E_{\max}P_{\max}$ simulations and anomaly in late summer low-level winds, **g**: $E_{\max}P_{\min}$ minus $E_{\min}P_{\min}$, **h**: $E_{\max}P_{\max}$ minus $E_{\min}P_{\max}$.

Extended Data Fig. 6 Nutrients, temperature and upwelling in model simulations.

a, NO₃ concentrations in the surface layer (0-100m). **b**, Upwelling velocity (averaged between 40 and 80m), **c**: Sea Surface Temperature (SST). All variables are averaged over JASO. Left: E_{\min} , Middle: $E_{\max}P_{\min}$ minus $E_{\min}P_{\min}$, Right: $E_{\max}P_{\max}$ minus $E_{\min}P_{\max}$.

Extended Data Fig. 7 Explanation of non-linearities in coccolithophore evolution.

a, Low-pass filter design for the delay between first appearance datum (FAD) and the beginning of the acme (BA) for *E. huxleyi* (blue line, lag of two eccentricity cycles) and another possible scenario for another species (red line, lag of one eccentricity cycle). The stepped green line represents *E. huxleyi*'s existence (0 = absence, 1 = presence). The blue and red curves in all panels are the output series of the 2 low-pass filters described in the [Methods](#). The black curve in **a** represents coeval eccentricity values. **b**, **c**, Bode plots of the 1-cycle lag filter (red) and the 2-cycle lag filter (blue) for magnitude (**b**) and phase (**c**) (see [Methods](#)). Earth's primary orbital periods are indicated by shading.

Extended Data Table 1 Characteristics of the nine marine records used in this study

Extended Data Table 2 Relative calcium carbonate mass contribution per calcareous nannofossil taxon or group for each sediment core

Extended Data Table 3 Summary of orbital parameters²³ used for each simulation and mean yearly contrast of radiation at equator (Wm⁻²) derived from IPSL-CM5A2

Supplementary information

Reporting Summary

Peer Review File

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Beaufort, L., Bolton, C.T., Sarr, AC. *et al.* Cyclic evolution of phytoplankton forced by changes in tropical seasonality. *Nature* **601**, 79–84 (2022). <https://doi.org/10.1038/s41586-021-04195-7>

- Received: 30 October 2020
- Accepted: 29 October 2021
- Published: 01 December 2021
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04195-7>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

Further reading

- [Earth's eccentric orbit paced the evolution of marine phytoplankton](#)
 - Rosalind E. M. Rickaby

Nature (2022)

[Earth's eccentric orbit paced the evolution of marine phytoplankton](#)

- Rosalind E. M. Rickaby

News & Views 01 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04195-7>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 15 December 2021](#)

Spatial genomics enables multi-modal study of clonal heterogeneity in tissues

- [Tongtong Zhao](#)^{1,2 na1},
- [Zachary D. Chiang](#)^{1,2,3 na1},
- [Julia W. Morriss](#)^{1,2},
- [Lindsay M. LaFave](#)^{2,4,5},
- [Evan M. Murray](#)^{1,2},
- [Isabella Del Priore](#)^{4,5},
- [Kevin Meli](#) [ORCID: orcid.org/0000-0002-9579-3829](#)^{4,5},
- [Caleb A. Lareau](#) [ORCID: orcid.org/0000-0003-4179-4807](#)^{1,2},
- [Naeem M. Nadaf](#)¹,
- [Jilong Li](#)¹,
- [Andrew S. Earl](#) [ORCID: orcid.org/0000-0002-5732-954X](#)^{1,2,3},
- [Evan Z. Macosko](#) [ORCID: orcid.org/0000-0002-2794-5165](#)^{1,6},
- [Tyler Jacks](#) [ORCID: orcid.org/0000-0001-5785-8911](#)^{1,4,5},
- [Jason D. Buenrostro](#) [ORCID: orcid.org/0000-0001-9958-3987](#)^{1,2,3 na2} &
- [Fei Chen](#) [ORCID: orcid.org/0000-0003-2308-3649](#)^{1,2,3 na2}

[Nature](#) volume 601, pages 85–91 (2022)

- 15k Accesses
- 251 Altmetric
- [Metrics details](#)

Subjects

- [Cancer genomics](#)
- [DNA sequencing](#)
- [Genomic analysis](#)
- [Genomics](#)

Abstract

The state and behaviour of a cell can be influenced by both genetic and environmental factors. In particular, tumour progression is determined by underlying genetic aberrations^{1,2,3,4} as well as the makeup of the tumour microenvironment^{5,6}. Quantifying the contributions of these factors requires new technologies that can accurately measure the spatial location of genomic sequence together with phenotypic readouts. Here we developed slide-DNA-seq, a method for capturing spatially resolved DNA sequences from intact tissue sections. We demonstrate that this method accurately preserves local tumour architecture and enables the de novo discovery of distinct tumour clones and their copy number alterations. We then apply slide-DNA-seq to a mouse model of metastasis and a primary human cancer, revealing that clonal populations are confined to distinct spatial regions. Moreover, through integration with spatial transcriptomics, we uncover distinct sets of genes that are associated with clone-specific genetic aberrations, the local tumour microenvironment, or both. Together, this multi-modal spatial genomics approach provides a versatile platform for quantifying how cell-intrinsic and cell-extrinsic factors contribute to gene expression, protein abundance and other cellular phenotypes.

This is a preview of subscription content

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Slide-DNA-seq enables spatially resolved DNA sequencing.

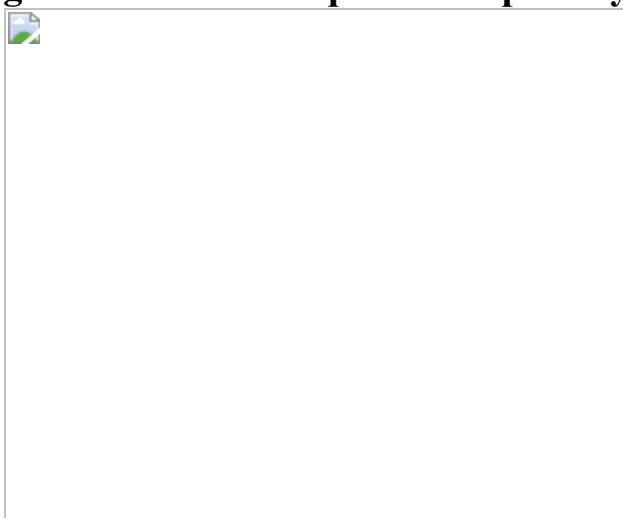


Fig. 2: Paired slide-DNA-seq and slide-RNA-seq characterize the genetics and transcriptomes of distinct metastatic clones.



Fig. 3: De novo identification of spatial tumour clones in primary human colorectal cancer.

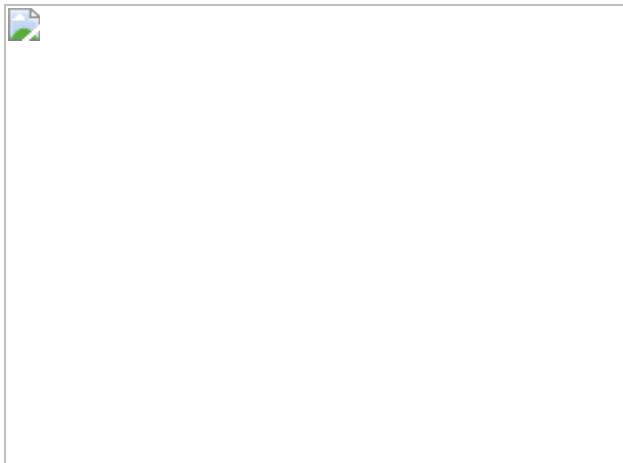
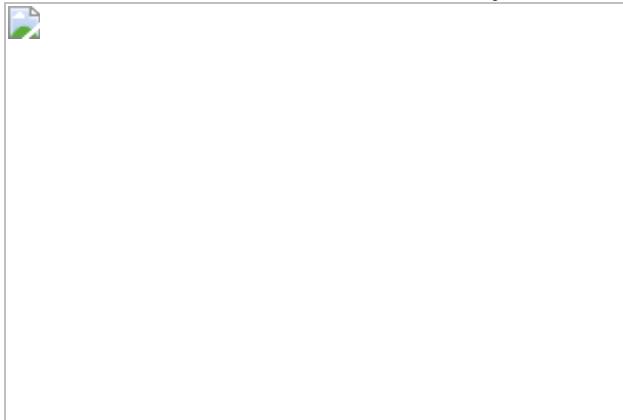


Fig. 4: Decomposition of transcriptional programs driven by genetic aberrations and tumour density.



Data availability

Raw sequencing data are available from the Sequence Read Archive under accession [PRJNA768453](#). Spatial barcode locations and counts matrices are available from the Broad Institute Single Cell Portal (https://singlecell.broadinstitute.org/single_cell/study/SCP1278). GC-content tracks for hg19 and mm10 were downloaded from the UC Santa Cruz Genome Browser. k36 mappability tracks for both genomes were downloaded from <https://bismap.hoffmanlab.org/>. Replication timing data were downloaded from Gene Expression Omnibus accession [GSM923451](#) for hg19 and [GSE137764](#) for mm10. Tn5 insertion bias tracks for both genomes were generated using the bias command from pyatac (<https://nucleoatac.readthedocs.io/en/latest/pyatac/>). Gene sets were downloaded from the Molecular Signatures Database Collections (<http://www.gsea-msigdb.org/gsea/msigdb/collections.jsp>).

Code availability

Code for the in situ bead indexing is available from <https://github.com/broadchenf/Slideseq>. Code for all analyses is available from https://github.com/buenrostrolab/slides_dna_seq_analysis and archived at <https://doi.org/10.5281/zenodo.5553305>.

References

1. 1.
McGranahan, N. & Swanton, C. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell* **168**, 613–628 (2017).
2. 2.
Turajlic, S., Sottoriva, A., Graham, T. & Swanton, C. Resolving genetic heterogeneity in cancer. *Nat. Rev. Genet.* **20**, 404–416 (2019).
3. 3.

Ramón y Cajal, S. et al. Clinical implications of intratumor heterogeneity: challenges and opportunities. *J. Mol. Med.* **98**, 161–177 (2020).

4. 4.

Pogrebniak, K. L. & Curtis, C. Harnessing tumor evolution to circumvent resistance. *Trends Genet.* **34**, 639–651 (2018).

5. 5.

Duan, Q., Zhang, H., Zheng, J. & Zhang, L. Turning cold into hot: firing up the tumor microenvironment. *Trends Cancer Res.* **6**, 605–618 (2020).

6. 6.

Jin, M.-Z. & Jin, W.-L. The updated landscape of tumor microenvironment and drug repurposing. *Signal Transduct. Target. Ther.* **5**, 166 (2020).

7. 7.

Tse, J. M. et al. Mechanical compression drives cancer cells toward invasive phenotype. *Proc. Natl Acad. Sci. USA* **109**, 911–916 (2012).

8. 8.

Zhao, Y. et al. Selection of metastasis competent subclones in the tumour interior. *Nat. Ecol. Evol.* **5**, 1033–1045 (2021).

9. 9.

Stylianopoulos, T., Munn, L. L. & Jain, R. K. Reengineering the physical microenvironment of tumors to improve drug delivery and efficacy: from mathematical modeling to bench to bedside. *Trends Cancer Res.* **4**, 292–319 (2018).

10. 10.

Gerstung, M. et al. The evolutionary history of 2,658 cancers. *Nature* **578**, 122–128 (2020).

11. 11.

Landau, D. A. et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* **152**, 714–726 (2013).

12. 12.

Wang, Y. et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* **512**, 155–160 (2014).

13. 13.

Xu, X. et al. Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell* **148**, 886–895 (2012).

14. 14.

Minussi, D. C. et al. Breast tumours maintain a reservoir of subclonal diversity during expansion. *Nature* **592**, 302–308 (2021).

15. 15.

Casasent, A. K. et al. Multiclonal invasion in breast tumors identified by topographic single cell sequencing. *Cell* **172**, 205–217.e12 (2018).

16. 16.

McPherson, A. et al. Divergent modes of clonal spread and intraperitoneal mixing in high-grade serous ovarian cancer. *Nat. Genet.* **48**, 758–767 (2016).

17. 17.

Jamal-Hanjani, M. et al. Tracking the evolution of non-small-cell lung cancer. *N. Engl. J. Med.* **376**, 2109–2121 (2017).

18. 18.

Payne, A. C. et al. In situ genome sequencing resolves DNA sequence and structure in intact biological samples. *Science* **371**, eaay3446 (2020).

19. 19.

Rodriques, S. G. et al. Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (2019).

20. 20.

Stickels, R. R. et al. Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat. Biotechnol.* **39**, 313–319 (2021).

21. 21.

Adey, A. et al. Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biol.* **11**, R119 (2010).

22. 22.

Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).

23. 23.

McFadden, D. G. et al. Genetic and clonal dissection of murine small cell lung carcinoma progression by genome sequencing. *Cell* **156**, 1298–1311 (2014).

24. 24.

Johnson, L. et al. Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. *Nature* **410**, 1111–1116 (2001).

25. 25.

LaFave, L. M. et al. Epigenomic state transitions characterize tumor progression in mouse lung adenocarcinoma. *Cancer Cell* **38**, 212–228.e13 (2020).

26. 26.

Morishita, A. et al. HMGA2 is a driver of tumor metastasis. *Cancer Res.* **73**, 4289–4299 (2013).

27. 27.

Thuault, S. et al. Transforming growth factor- β employs HMGA2 to elicit epithelial–mesenchymal transition. *J. Cell Biol.* **174**, 175–183 (2006).

28. 28.

Kumar, M. S. et al. HMGA2 functions as a competing endogenous RNA to promote lung cancer progression. *Nature* **505**, 212–217 (2014).

29. 29.

Cable, D. M., Murray, E., Zou, L. S., Goeva, A. & Macosko, E. Z. Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-021-00830-w> (2021).

30. 30.

Dienstmann, R. et al. Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. *Nat. Rev. Cancer* **17**, 268 (2017).

31. 31.

The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).

32. 32.

Kang, H. et al. Many private mutations originate from the first few divisions of a human colorectal adenoma. *J. Pathol.* **237**, 355–362 (2015).

33. 33.

Humphries, A. et al. Lineage tracing reveals multipotent stem cells maintain human adenomas and the pattern of clonal expansion in tumor evolution. *Proc. Natl Acad. Sci. USA* **110**, E2490–E2499 (2013).

34. 34.

Brayer, K. J., Frerich, C. A., Kang, H. & Ness, S. A. Recurrent fusions in MYB and MYBL1 define a common, transcription factor-driven oncogenic pathway in salivary gland adenoid cystic carcinoma. *Cancer Discov.* **6**, 176–187 (2016).

35. 35.

Wang, B., Matsuoka, S., Carpenter, P. B. & Elledge, S. J. 53BP1, a mediator of the DNA damage checkpoint. *Science* **298**, 1435–1438 (2002).

36. 36.

Long, D. T., Räsche, M., Joukov, V. & Walter, J. C. Mechanism of RAD51-dependent DNA interstrand cross-link repair. *Science* **333**, 84–87 (2011).

37. 37.

Lachaud, C. et al. Ubiquitinated Fancd2 recruits Fan1 to stalled replication forks to prevent genome instability. *Science* **351**, 846–849

(2016).

38. 38.

Sottoriva, A. et al. A Big Bang model of human colorectal tumor growth. *Nat. Genet.* **47**, 209–216 (2015).

39. 39.

Landrette, S. F. et al. *Plagl1* and *Plagl2* are oncogenes that induce acute myeloid leukemia in cooperation with *Cbfβ-MYH11*. *Blood* **105**, 2900–2907 (2005).

40. 40.

Ren, B. et al. *MCM7* amplification and overexpression are associated with prostate cancer progression. *Oncogene* **25**, 1090–1098 (2006).

41. 41.

Farhad, M., Rolig, A. S. & Redmond, W. L. The role of Galectin-3 in modulating tumor growth and immunosuppression within the tumor microenvironment. *Oncoimmunology* **7**, e1434467 (2018).

42. 42.

Karim, B. O., Rhee, K.-J., Liu, G., Yun, K. & Brant, S. R. Prom1 function in development, intestinal inflammation, and intestinal tumorigenesis. *Front. Oncol.* **4**, 323 (2014).

43. 43.

Winkler, J., Abisoye-Ogunniyan, A., Metcalf, K. J. & Werb, Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat. Commun.* **11**, 5120 (2020).

44. 44.

Wu, C.-Y. et al. Integrative single-cell analysis of allele-specific copy number alterations and chromatin accessibility in cancer. *Nat. Biotechnol.* **39**, 1259–1269 (2021).

45. 45.

Wu, S. et al. Circular ecDNA promotes accessible chromatin and high oncogene expression. *Nature* **575**, 699–703 (2019).

46. 46.

Integrative HMP (iHMP) Research Network Consortium. The Integrative Human Microbiome Project. *Nature* **569**, 641–648 (2019).

47. 47.

van Haasteren, J., Li, J., Scheideler, O. J., Murthy, N. & Schaffer, D. V. The delivery challenge: fulfilling the promise of therapeutic genome editing. *Nat. Biotechnol.* **38**, 845–855 (2020).

48. 48.

Goldman, N. et al. Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature* **494**, 77–80 (2013).

49. 49.

Ludwig, L. S. et al. Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. *Cell* **176**, 1325–1339.e22 (2019).

50. 50.

Liu, Y. et al. Bisulfite-free direct detection of 5-methylcytosine and 5-hydroxymethylcytosine at base resolution. *Nat. Biotechnol.* **37**, 424–429 (2019).

Acknowledgements

J.D.B. and F.C. acknowledge funding from the Allen Institute Distinguished Investigator award and funding from the NIH R21HG009749. F.C. also acknowledges funding from NIH DP50D024583, and R33CA246455. F.C. and E.Z.M acknowledge funding from NIH R01HG010647. F.C. also acknowledges the Searle Scholars Award and the Burroughs Wellcome CASI award. J.D.B. acknowledges support from the NIH New Innovator Award (DP2HL151353). Z.D.C. acknowledges funding from NHGRI training grant T32HG002295 and the Harvard Quantitative Biology Initiative. We thank S. Nagaraja for assistance in making figures. Components of our figures were created with BioRender.com. We thank J. Strecker for the gift of Tn5 enzyme, and the Buenrostro and Chen laboratories for helpful discussions. We thank the cancer patients and their families for their invaluable donations to science, making this work possible.

Author information

Author notes

1. These authors contributed equally: Tongtong Zhao, Zachary D. Chiang
2. These authors jointly supervised this work: Jason D. Buenrostro, Fei Chen

Affiliations

1. Broad Institute of MIT and Harvard, Cambridge, MA, USA

Tongtong Zhao, Zachary D. Chiang, Julia W. Morriss, Evan M. Murray, Caleb A. Lareau, Naeem M. Nadaf, Jilong Li, Andrew S. Earl, Evan Z. Macosko, Tyler Jacks, Jason D. Buenrostro & Fei Chen

2. Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA

Tongtong Zhao, Zachary D. Chiang, Julia W. Morriss, Lindsay M. LaFave, Evan M. Murray, Caleb A. Lareau, Andrew S. Earl, Jason D.

Buenrostro & Fei Chen

3. Gene Regulation Observatory, Broad Institute of MIT and Harvard,
Cambridge, MA, USA

Zachary D. Chiang, Andrew S. Earl, Jason D. Buenrostro & Fei Chen

4. Department of Biology, Massachusetts Institute of Technology,
Cambridge, MA, USA

Lindsay M. LaFave, Isabella Del Priore, Kevin Meli & Tyler Jacks

5. David H. Koch Institute, Massachusetts Institute of Technology,
Cambridge, MA, USA

Lindsay M. LaFave, Isabella Del Priore, Kevin Meli & Tyler Jacks

6. Department of Psychiatry, Massachusetts General Hospital, Boston,
MA, USA

Evan Z. Macosko

Contributions

T.Z. and J.W.M. developed the protocol and performed experiments. Z.D.C. developed the computational processing pipeline. T.Z., Z.D.C., J.D.B., and F.C. performed analyses. L.M.L., I.D.P. and K.M. assisted with the mouse experiments under the supervision of T.J. E.M.M. performed the *in situ* bead indexing and slide-RNA-seq experiments under the supervision of E.Z.M. and F.C. J.L. wrote the *in situ* bead-indexing pipeline. N.M.N. assisted with the 10x experiment. C.A.L. and A.S.E. assisted with the computational processing pipeline. T.Z., Z.D.C., J.D.B. and F.C. wrote the manuscript with input from all authors. E.Z.M., T.J., J.D.B. and F.C. supervised this work.

Corresponding authors

Correspondence to [Jason D. Buenrostro](#) or [Fei Chen](#).

Ethics declarations

Competing interests

E.Z.M. and F.C. are listed as inventors on a patent application related to slide-seq. T.J. is a member of the board of directors of Amgen and Thermo Fisher Scientific. He is also a co-founder of Dragonfly Therapeutics and T2 Biosystems. T.J. serves on the scientific advisory board of Dragonfly Therapeutics, SQZ Biotech and Skyhawk Therapeutics, and is the president of Break Through Cancer. J.D.B. holds patents related to ATAC-seq and is on the scientific advisory board for Camp4, Seqwell and Celsee. F.C. is a paid consultant for Celsius Therapeutics and Atlas Bio. E.Z.M is a paid consultant for Atlas Bio. T.Z., E.Z.M., J.D.B. and F.C. have filed a patent application based on this work.

Additional information

Peer review information *Nature* thanks Andrew Adey and Naveed Ishaque for their contribution to the peer review of this work.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

[Extended Data Fig. 1 Optimization of slide-DNA-seq protocol.](#)

a, Library preparation steps. **b–e**, Library size comparisons for live vs fixed tissue (**b**); histone extraction protocols (**c, d**); and varying lengths of a bridge oligo used to connect the barcoded bead oligo to genomic fragments **e**, either hybridized after tagmentation (left bar) or pre-loaded onto the Tn5 transposase prior to tagmentation (rest). All values are normalized to control condition (first column). **f**, Rate of ligation of genomic fragments to barcoded oligo either ordered in solution from IDT (left) or cleaved off from beads (center, right). **g–j**, Frequency of Tn5 insertions in the genome

relative to the nearest transcription start site (TSS) for slide-DNA-seq of mouse cerebellum (**g**), mouse liver metastases (**h**), human colon tumor (**i**), and for single-cell ATAC-seq of a mouse brain (**j**). Error bars, mean \pm s.d; n, number of replicate comparisons (generated from 4 biological samples); dots represent values of each replicate.

Extended Data Fig. 2 Comparison of fixation conditions during histone extraction.

Cerebellar sections are exposed to treatment as stated (with or without prior fixation) and stained with DAPI. Scale bars, 500 μm .

Extended Data Fig. 3 Quantification of DNA fragments per slide-DNA-seq array.

a, Nuclear (left) and mitochondrial (center) DNA fragments per bead obtained for tissues used in this study. Right, mitochondrial fraction of fragments. 4x, protocol variant with 4x fragmentation. Black lines on violin plots indicate the mean. **b**, slide-DNA-seq of the mouse cerebellum experiment in Fig. 1. Beads are colored by the number of nuclear fragments (left), mitochondrial DNA fragments (center), and fraction of mitochondrial DNA fragments (right). **c**, Visualization of representative convex hulls for different spatial bin values of k for k-nearest neighbor smoothing. Beads are colored by raw counts, insets show convex hulls for k = 1, 10, 25, and 50, centered on salmon colored beads. Hulls are generally circular except at the edge of the array. **d**, Distribution of mean fragments per 1 Mb genomic bin for different spatial bin values of k. The median diameter of the smoothed features is indicated in parentheses. **e**, Comparison of nuclear fragments (left) and effective diameter (right) per bead for different spatial bin values of k. Scale bars, 500 μm .

Extended Data Fig. 4 Estimation of slide-DNA-seq lateral diffusion.

a, Interpolated image showing the nuclear fraction of fragments of a 10 μm mouse cerebellar section processed for slide-DNA-seq. Cyan box indicates

magnified area (right). Smaller boxes indicate regions taken for linescans in **b** and **e**. **b**, Pseudo-intensity (representing nuclear fraction of fragments) of linescan as indicated by red box in **a**. Black dots, halfmax. Full Width at Half Maximum (FWHM) = 57.3 μm . **c**, 10 μm serial section of the same mouse cerebellum stained with DAPI. Blue box indicates magnified area (right). Smaller boxes indicate regions taken for linescans in **d** and **f**. **d**, Linescan of DAPI intensity as indicated in **c**. Black dots, halfmax. FWHM = 16.4 μm . **e**, Same as **b**, but for 3 different regions as indicated by the smaller non-red boxes in **a**. For the left (green) and middle (yellow) panel, FWHM is calculated as twice the distance between the peak and the halfmax (marked by black dots). **f**, Same as **d**, but for 3 different regions as indicated by the smaller non-red boxes in **b**. For the left (green) and middle (yellow) panel, FWHM is calculated as twice the distance between the peak and the halfmax (marked by black dots). **g**, Bar graph of average FWHM ($n=4$ regions). Error bars, mean \pm s.d. Upper bound for the diffusion measurement is half of largest FWHM (not taking into account the DAPI measurement). Scale bars, 500 μm (**a**, **c**), 200 μm (**b**, **d**).

Extended Data Fig. 5 Normalization of slide-DNA-seq sequencing biases.

a, Top, raw sequencing reads per 1 Mb bin for mouse cerebellum slide-DNA-seq are plotted for GC-content, mappability, replication timing score, and Tn5 bias. Pearson's r values are shown for each. Bottom, bias corrected coverage and correlation values after normalization. **b**, Same as **a** but for tagmentation-based bulk sequencing of mouse cerebellum (**Methods**). **c**, Same as **a** but for slide-DNA seq of mouse liver metastases. **d**, Same as **b** but for tagmentation-based bulk sequencing of mouse liver metastases. Blue points, bins from chrX (not included in the calculation of the fit).

Extended Data Fig. 6 Quantification of genomic coverage in a diploid sample.

Left (all panels), copy number profiles at 1 Mb genomic resolution of the mouse cerebellum for the sequencing modality and processing indicated. For this diploid sample, each copy number distribution is normalized to a

median of 2. Right (all panels), histogram of the number of bins per copy number. **a**, Raw coverage profile of slide-DNA-seq. **b**, Coverage profile of slide-DNA-seq normalized by GC-content and mappability. **c**, Coverage profile for bulk fragmentation-based sequencing. **d**, Coverage profile of bulk sequencing normalized by GC-content and mappability. **e**, Coverage profile of slide-DNA-seq normalized by the GC-content and mappability divided by bulk sequencing normalized by GC-content and mappability.

Extended Data Fig. 7 slide-DNA-seq clonal analysis workflow.

a, Principal components calculated from smoothed slide-DNA-seq beads, ordered by the percentage of variance explained for the mouse liver metastases array shown in Fig. 1. **b**, Weights per 1 Mb genomic bin for principal components 1 and 2. Red points indicate bins from chromosomes with an odd number, blue from chromosomes with an even number (and chrX). **c**, slide-DNA-seq array for the mouse liver metastases array shown in Fig. 1 with points colored by raw PC 1 scores (top left), smoothed PC 1 scores (top right), raw PC 2 scores (bottom left), smoothed PC 2 scores (bottom right). **d**, Calinski-Harabasz criterion values used to select the optimal value of k for k-means clustering. **e**, slide-DNA-seq array colored by cluster assignment using the value of k selected in **d**. **f–j**, Same as **a–e**, but for the mouse liver metastases array shown in Fig. 2.

Extended Data Fig. 8 Accuracy of clonal assignment via downsampling of bulk tumor cell lines.

a, Raw copy number profiles for four tumor cell lines profiled using fragmentation-based bulk sequencing. **b**, Representative 10,000 fragments samples of the cell lines shown in **a**. **c**, Clonal assignment accuracy for 10,000 fragment samples ($n=5,000$ samples of each cell line) using the analysis workflow shown in Extended Data Fig. 7. **d**, Same as **c** but for 1,000 fragment samples.

Extended Data Fig. 9 Reproducibility of slide-DNA-seq across serial sections.

a, Immunofluorescence (IF) against tumor marker HMGA2 (top) and two slide-DNA-seq replicates (center, bottom) were performed on two serial sections of a mouse liver metastasis. Beads colored by PC1 scores (left) and cluster assignment (right) show similar spatial architecture between replicates. Scale bars, 500 μ m. **b**, Aggregate copy number profiles of normal and tumor beads show high correlation (Pearson's $r = 0.986$ and 0.992) between the two replicates.

Extended Data Fig. 10 Quantification of genomic coverage by bin size and number of beads.

Each column represents normalized copy number profiles aggregated across the number of slide-DNA-seq beads indicated (10,000; 1,000; or 100), while each row indicates the genomic bin size (10 Mb, 5 Mb, 2.5 Mb, 1 Mb, and 500 kb) for the mouse cerebellum array.

Extended Data Fig. 11 Integrated slide-RNA-seq and single-nucleus RNA-seq analysis of clones.

a, H&E stain (left), IHC against tumor marker HMGA2 (center), and *Hmga2* expression from slide-RNA-seq (right) of three serial sections of mouse liver metastases. **b**, UMAP of unsupervised clustering of single nucleus RNA-seq performed on nuclei from mouse liver metastasis sample. **c**, Dot plot showing the expression of marker genes used to annotate clusters in **b**. **d**, Spatial projection of cell types from **b** onto the slide-RNA-seq array, colored in the same fashion. Black lines indicate spatial tumor clusters. **e**, Differential localization of cell types between clone A, clone B and normal regions. Heatmap shows signed (positive, enrichment; negative, depletion) $\log_{10}(p\text{-value})$ from permutation testing (two-sided, not adjusted for multiple comparisons). **f**, Spatial plot of monocyte localization on the array, which is significantly enriched for clone B. Black lines indicate spatial tumor clusters. VSMC, vascular smooth muscle cell; LSEC, liver sinusoidal endothelial cells.

Extended Data Fig. 12 Validation of ploidy and copy number of metastatic clones.

a, Assignment of beads to normal tissue, clone A, and clone B based on k-means clustering. **b–d**, Histogram of DNA content of single cells measured by propidium iodide (PI) fluorescence intensity through flow cytometry. (**b**) bone marrow cells (normal control); (**c**) clone A; (**d**) clone B. Diploid G1 (2N) and G2 (4N) gates are determined on bone marrow histogram and applied to clones A and B, revealing that the clone A genome is triploid; and the clone B genome is diploid with some amplifications (e.g. of chr. 15 and 19, see Fig. 2d). **e**, Aggregate copy number profiles of beads assigned to clone A. **f**, Aggregate copy number profiles of beads assigned to clone B.

Extended Data Fig. 13 Spatial projection of single-cell whole-genome sequencing (scWGS) clusters.

a, Genomic copy number profiles for 2,274 single cells obtained using scWGS, with cluster annotations colored. **b**, Top left: projection of scWGS clusters onto slide-DNA-seq. All other: three genomic regions of differential CNA profiles between the three projected clusters, shown are spatial heatmaps of signed p-value differences from the average profile (two-sided permutation test, not adjusted for multiple comparisons). **c**, Normalized copy number profiles for the three scWGS clusters, and the corresponding spatial clusters. Vertical lines denote variable regions from **b**. Single-cell cluster 2 (blue) shows complex CNA patterns that obscure cluster ploidy, nevertheless, copy number values are normalized to 2 for easy comparison to other clusters.

Extended Data Fig. 14 Tumor morphology of primary human colon cancer sample.

a, H&E stain of normal colon (left) and colon tumor (right) tissue from the same patient. Scale bars, 200 µm. **b**, Serial sections processed for H&E stain (left), slide-DNA-seq (center), and slide-RNA-seq (right). Scale bars, 500 µm. Yellow and black boxes indicate magnified areas in **c**, **d**, respectively. **c**, Magnified views of H&E stain, slide-DNA-seq and slide-RNA-seq reconstructions show concordant spatial tissue architecture across three modalities; scale bar, 200 µm. **d**, Magnified view of H&E stain of three regions that are assigned low, medium, and high tumor density by

slide-RNA-seq transcriptomic analysis (**b**, right). Arrows indicate regions of high tumor density identified through H&E stain. Scale bar, 100 μ m.

Extended Data Fig. 15 Biological pathways explained by subclone or tumor density.

a, Subclone-associated pathways identified through gene set enrichment analysis. **b**, Hallmark E2F target genes (n=200) plotted according to percent variance explained by clonal identity (x-axis) and tumor density (y-axis). Included genes colored by normalized density on the scatter plot, all other genes are shown in grey. **c**, Expression of highly subclone-associated E2F target genes (n=11, listed in **a**), plotted for spatial tumor regions of the slide-RNA-seq array from Fig. 4. **d**, MYC target genes (n=200) plotted according to percent variance explained by clonal identity (x-axis) and tumor density (y-axis). Included genes are colored by normalized density on the scatter plot, MYC is colored red, all other genes are shown in grey. **e**, Expression of highly subclone-associated MYC target genes (n=16, listed in **a**), plotted for spatial tumor regions (left). Box plot showing normalized MYC target gene expression by subclone assignment; each point represents a spatial tumor cluster (right). Red line, mean, red box, 95% confidence interval for mean, blue box, standard deviation. **f**, MYC expression plotted for spatial tumor regions (left). Scatter plot showing normalized MYC expression by tumor cell density; each point represents a spatial tumor cluster (right). **g**, Subclone-associated pathways identified through gene set enrichment analysis. **h**, Cell adhesion molecule binding genes (n=514) plotted according to percent variance explained by clonal identity (x-axis) and tumor density (y-axis). Included genes are colored by normalized density on the scatter plot, all other genes are shown in grey (reproduced from Fig. 4i). **i**, Expression of highly density-associated cell adhesion molecule binding genes (n=14, listed in **g**), plotted for spatial tumor regions. Scale bars, 500 μ m.

Supplementary information

Supplementary Information

This file contains Supplementary Materials and Methods, Supplementary text, legends for Extended Data Figures and Supplementary References

Reporting Summary

Supplementary Table 1

Oligos

Supplementary Table 2

Supplementary Table 3

Marker Genes

Supplementary Table 4

Mouse Clone DEGs

Supplementary Table 5

Human colon variance exp

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Zhao, T., Chiang, Z.D., Morriss, J.W. *et al.* Spatial genomics enables multi-modal study of clonal heterogeneity in tissues. *Nature* **601**, 85–91 (2022).
<https://doi.org/10.1038/s41586-021-04217-4>

- Received: 01 February 2021
- Accepted: 08 November 2021
- Published: 15 December 2021
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04217-4>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04217-4>

- Article
- [Published: 15 December 2021](#)

Building an allocentric travelling direction signal via vector computation

- [Cheng Lyu](#)¹,
- [L. F. Abbott](#)² &
- [Gaby Maimon](#) [ORCID: orcid.org/0000-0003-1219-5856](#)¹

[Nature](#) volume **601**, pages 92–97 (2022)

- 4824 Accesses
- 1 Citations
- 196 Altmetric
- [Metrics details](#)

Subjects

- [Navigation](#)
- [Network models](#)
- [Neural circuits](#)

Abstract

Many behavioural tasks require the manipulation of mathematical vectors, but, outside of computational models^{1,2,3,4,5,6,7}, it is not known how brains

perform vector operations. Here we show how the *Drosophila* central complex, a region implicated in goal-directed navigation^{7,8,9,10}, performs vector arithmetic. First, we describe a neural signal in the fan-shaped body that explicitly tracks the allocentric travelling angle of a fly, that is, the travelling angle in reference to external cues. Past work has identified neurons in *Drosophila*^{8,11,12,13} and mammals¹⁴ that track the heading angle of an animal referenced to external cues (for example, head direction cells), but this new signal illuminates how the sense of space is properly updated when travelling and heading angles differ (for example, when walking sideways). We then characterize a neuronal circuit that performs an egocentric-to-allocentric (that is, body-centred to world-centred) coordinate transformation and vector addition to compute the allocentric travelling direction. This circuit operates by mapping two-dimensional vectors onto sinusoidal patterns of activity across distinct neuronal populations, with the amplitude of the sinusoid representing the length of the vector and its phase representing the angle of the vector. The principles of this circuit may generalize to other brains and to domains beyond navigation where vector operations or reference-frame transformations are required.

This is a preview of subscription content

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Fig. 1: $h\Delta B$ neurons signal the allocentric travelling direction in *Drosophila*.

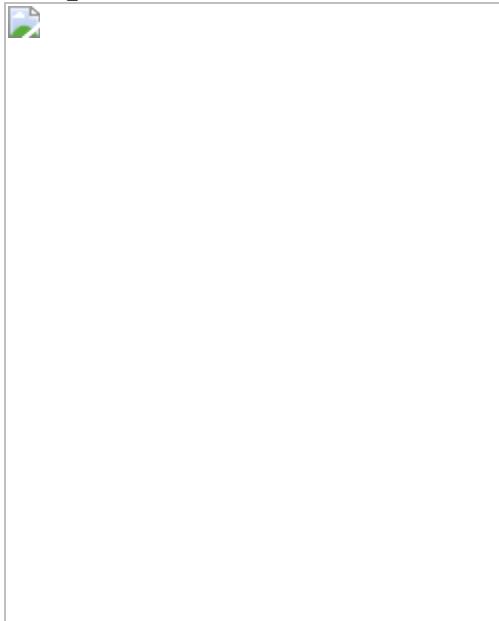


Fig. 2: The allocentric travelling direction can be computed by vector rotation and summation, which can be implemented by phasors.

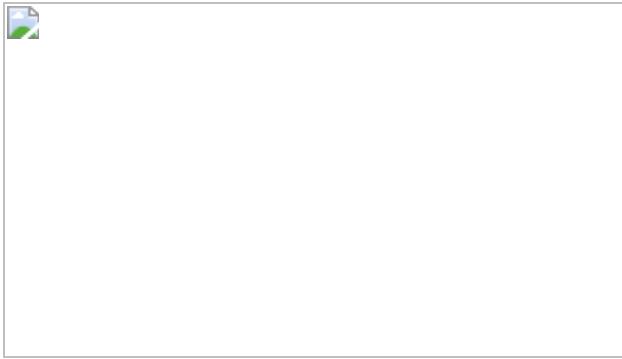


Fig. 3: PFN_d and PFN_v cells show physiological and anatomical patterns consistent with them functioning to build the travelling direction signal in hΔB cells.

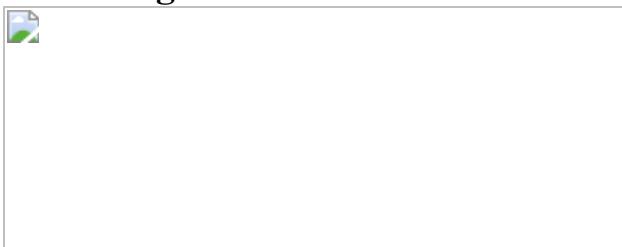


Fig. 4: A model of how vector computation builds the travelling direction signal in hΔB cells.

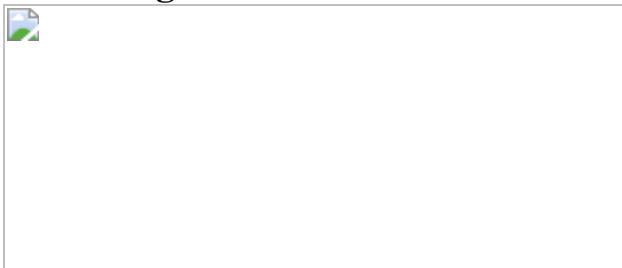
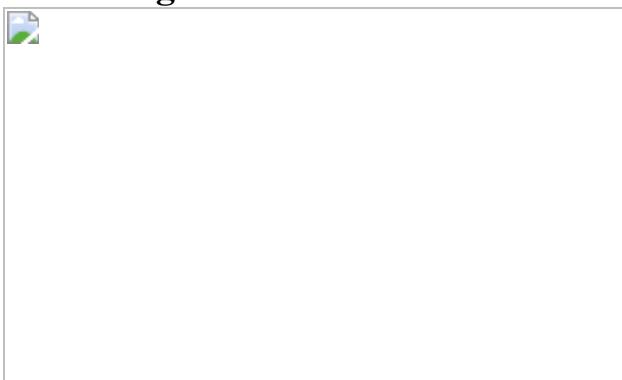


Fig. 5: Perturbations of neural activity induce changes in the travelling direction signal that are consistent with the vector sum model.



Data availability

Data for all of the main figures are available on Dropbox (<https://www.dropbox.com/sh/p8bqwavlsyl9ppv/AABz2-yda4Q3gukXqp8Ba2Gwa?dl=0>). Other data are available on request from the corresponding author.

Code availability

The analysis code has been deposited on GitHub (https://github.com/Cheng-Lyu/TravelingDirectionPaper_code).

References

1. 1. Zipser, D. & Andersen, R. A. A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature* **331**, 679–684 (1988).
2. 2. O’Keefe, J. An allocentric spatial model for the hippocampal cognitive map. *Hippocampus* **1**, 230–235 (1991).
3. 3. Touretzky, D. S., Redish, A. D. & Wan, H. S. Neural representation of space using sinusoidal arrays. *Neural Comput.* **5**, 869–884 (1993).
4. 4. Pouget, A. & Sejnowski, T. J. A neural model of the cortical representation of egocentric distance. *Cereb. Cortex* **4**, 314–329 (1994).
5. 5.

Salinas, E. & Abbott, L. F. Transfer of coded information from sensory to motor networks. *J. Neurosci.* **15**, 6461–6474 (1995).

6. 6.

Wittmann, T. & Schwegler, H. Path integration—a network model. *Biol. Cybern.* **73**, 569–575 (1995).

7. 7.

Stone, T. et al. An anatomically constrained model for path integration in the bee brain. *Curr. Biol.* **27**, 3069-3085.e11 (2017).

8. 8.

Seelig, J. D. & Jayaraman, V. Neural dynamics for landmark orientation and angular path integration. *Nature* **521**, 186–191 (2015).

9. 9.

Green, J., Vijayan, V., Mussells Pires, P., Adachi, A. & Maimon, G. A neural heading estimate is compared with an internal goal to guide oriented navigation. *Nat. Neurosci.* **22**, 1460–1468 (2019).

10. 10.

Giraldo, Y. M. et al. Sun navigation requires compass neurons in *Drosophila*. *Curr. Biol.* **28**, 2845-2852.e4 (2018).

11. 11.

Green, J. et al. A neural circuit architecture for angular integration in *Drosophila*. *Nature* **546**, 101–106 (2017).

12. 12.

Turner-Evans, D. et al. Angular velocity integration in a fly heading circuit. *eLife* **6**, e23496 (2017).

13. 13.

Shiozaki, H. M., Ohta, K. & Kazama, H. A multi-regional network encoding heading and steering maneuvers in *Drosophila*. *Neuron* **106**, 126–141 (2020).

14. 14.

Taube, J. S., Muller, R. U. & Ranck, J. B. Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *J. Neurosci.* **10**, 420–435 (1990).

15. 15.

Seeley, T. D. *Honeybee Democracy* (Princeton Univ. Press, 2010).

16. 16.

Wehner, R. *Desert Navigator* (Harvard Univ. Press, 2020).

17. 17.

Ardin, P. B., Mangan, M. & Webb, B. Ant homing ability is not diminished when traveling backwards. *Front. Behav. Neurosci.* **10**, 69 (2016).

18. 18.

Pfeffer, S. E. & Wittlinger, M. How to find home backwards? Navigation during rearward homing of *Cataglyphis fortis* desert ants. *J. Exp. Biol.* **219**, 2119–2126 (2016).

19. 19.

Riley, J. R. et al. Compensation for wind drift by bumble-bees. *Nature* **400**, 126–126 (1999).

20. 20.

Scheffer, L. K. et al. A connectome and analysis of the adult *Drosophila* central brain. *eLife* **9**, e57443 (2020).

21. 21.

Wolff, T., Iyer, N. A. & Rubin, G. M. Neuroarchitecture and neuroanatomy of the *Drosophila* central complex: a GAL4-based dissection of protocerebral bridge neurons and circuits. *J. Comp. Neurol.* **523**, 997–1037 (2015).

22. 22.

Kim, S. S., Rouault, H., Druckmann, S. & Jayaraman, V. Ring attractor dynamics in the *Drosophila* central brain. *Science* **356**, 849–853 (2017).

23. 23.

Cohn, R., Morantte, I. & Ruta, V. Coordinated and compartmentalized neuromodulation shapes sensory processing in *Drosophila*. *Cell* **163**, 1742–1755 (2015).

24. 24.

Maimon, G., Straw, A. D. & Dickinson, M. H. Active flight increases the gain of visual motion processing in *Drosophila*. *Nat. Neurosci.* **13**, 393–399 (2010).

25. 25.

Srinivasan, M. V., Zhang, S. W., Lehrer, M. & Collett, T. S. Honeybee navigation en route to the goal: visual flight control and odometry. *J. Exp. Biol.* **199**, 237–244 (1996).

26. 26.

Weir, P. T. & Dickinson, M. H. Functional divisions for visual processing in the central brain of flying *Drosophila*. *Proc. Natl Acad. Sci. USA* **112**, E5523–E5532 (2015).

27. 27.

Kim, A. J., Fenk, L. M., Lyu, C. & Maimon, G. Quantitative predictions orchestrate visual signaling in *Drosophila*. *Cell* **168**, 280–294.e12 (2017).

28. 28.

Hulse, B. K. et al. A connectome of the *Drosophila* central complex reveals network motifs suitable for flexible navigation and context-dependent action selection. *eLife* **10**, e66039 (2021).

29. 29.

Poody, C. & Edgar, L. Reversible alteration in the neuromuscular junctions of *Drosophila melanogaster* bearing a temperature-sensitive mutation, shibire. *J. Cell Biol.* **81**, 520–527 (1979).

30. 30.

Baines, R. A., Uhler, J. P., Thompson, A., Sweeney, S. T. & Bate, M. Altered electrical properties in *Drosophila* neurons developing without synaptic transmission. *J. Neurosci.* **21**, 1523–1531 (2001).

31. 31.

Govorunova, E. G., Sineshchekov, O. A., Janz, R., Liu, X. & Spudich, J. L. Natural light-gated anion channels: a family of microbial rhodopsins for advanced optogenetics. *Science* **349**, 647–650 (2015).

32. 32.

Mohammad, F. et al. Optogenetic inhibition of behavior with anion channelrhodopsins. *Nat. Methods* **14**, 271–274 (2017).

33. 33.

Wolff, T. & Rubin, G. M. Neuroarchitecture of the *Drosophila* central complex: a catalog of nodulus and asymmetrical body neurons and a

revision of the protocerebral bridge catalog. *J. Comp. Neurol.* **526**, 2585–2611 (2018).

34. 34.

Klapoetke, N. C. et al. Independent optical excitation of distinct neural populations. *Nat. Methods* **11**, 338–346 (2014).

35. 35.

Cei, A., et al Reversed theta sequences of hippocampal cell assemblies during backward travel. *Nat. Neurosci.* **17**, 719–724 (2014).

36. 36.

Maurer, A. P., Lester, A. W., Burke, S. N., Ferng, J. J. & Barnes, C. A. Back to the future: preserved hippocampal network activity during reverse ambulation. *J. Neurosci.* **34**, 15022–15031 (2014).

37. 37.

Bicanski, A. & Burgess, N. Neuronal vector coding in spatial cognition. *Nat. Rev. Neurosci.* **21**, 453–470 (2020).

38. 38.

Wang, C., Chen, X. & Knierim, J. J. Egocentric and allocentric representations of space in the rodent brain. *Curr. Opin. Neurobiol.* **60**, 12–20 (2020).

39. 39.

Vaswani, A. et al. Attention is all you need. Preprint at <https://arxiv.org/abs/1706.03762> (2017).

40. 40.

Lu, J. et al. Transforming representations of movement from body- to world-centric space. *Nature* <https://doi.org/10.1038/s41586-021->

[04191-x](#) (2021).

41. 41.

Dana, H. et al. Sensitive red protein calcium indicators for imaging neural activity. *eLife* **5**, e12727 (2016).

42. 42.

Lin, C. Y. et al. A comprehensive wiring diagram of the protocerebral bridge for visual information processing in the *Drosophila* brain. *Cell Rep.* **3**, 1739–1753 (2013).

43. 43.

Turner-Evans, D. et al. The neuroanatomical ultrastructure and function of a biological ring attractor. *Neuron* **108**, 145–163 (2020).

44. 44.

Nern, A., Pfeiffer, B. D. & Rubin, G. M. Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. *Proc. Natl Acad. Sci. USA* **112**, E2967–E2976 (2015).

45. 45.

Reiser, M. B. & Dickinson, M. H. A modular display system for insect behavioral neuroscience. *J. Neurosci. Methods* **167**, 127–139 (2008).

46. 46.

Maimon, G., Straw, A. D. & Dickinson, M. H. A simple vision-based algorithm for decision making in flying *Drosophila*. *Curr. Biol.* **18**, 464–470 (2008).

47. 47.

Fry, S. N., Rohrseitz, N., Straw, A. D. & Dickinson, M. H. Visual control of flight speed in *Drosophila melanogaster*. *J. Exp. Biol.* **212**, 1120–1130 (2009).

48. 48.

Leitch, K. J., Ponce, F. V., Dickson, W. B., van Breugel, F. & Dickinson, M. H. The long-distance flight behavior of *Drosophila* supports an agent-based model for wind-assisted dispersal in insects. *Proc. Natl Acad. Sci.* **118**, e2013342118 (2021).

49. 49.

Moore, R. J. D. et al. FicTrac: a visual method for tracking spherical motion and generating fictive animal paths. *J. Neurosci. Methods* **225**, 106–119 (2014).

50. 50.

Wu, M. et al. Visual projection neurons in the *Drosophila* lobula link feature detection to distinct behavioral programs. *eLife* **5**, e21022 (2016).

Acknowledgements

We thank the laboratories of V. Ruta, B. Dickson and L. Vosshall for fly stocks and members of the Maimon laboratory, especially P. Mussells-Pires, for helpful discussions; R. Wilson and J. Lu for sharing the hypothesis that the h Δ B cells may be the locus for the four-vector integration process rather than PFR cells, given anatomical considerations; A. Rizvi and S. Lewallen for helpful comments on the manuscript; J. Green and A. Adachi for recombinant stocks, and A. Adachi for initial anatomical characterization of some of the cell types and driver lines, which made this study progress faster than otherwise possible; V. Vijayan and J. Weisman for helping to optimize the csChrimson optogenetics approach during two-photon imaging; A. Kim for sharing the code for making optic-flow stimuli, developed off of the code initially given to us by P. Weir in the laboratory of

M. Dickinson; and I. Morantte from the laboratory of V. Ruta and P. Mussells-Pires for developing the genetic strategy and organizing the plasmids needed to generate the UAS-sytGCaMP7f fly line. Stocks obtained from the Bloomington Drosophila Stock Center (NIH P40OD018537) and the Vienna Drosophila Resource Center were used in this study. Research reported in this publication was supported by a Brain Initiative grant from the National Institute of Neurological Disorders and Stroke (R01NS104934) to G.M. and a Kavli Foundation seed grant to C.L. L.F.A. was supported by NSF NeuroNex Award DBI-1707398 and the Simons Collaboration for the Global Brain. G.M. is a Howard Hughes Medical Institute Investigator.

Author information

Affiliations

1. Laboratory of Integrative Brain Function and Howard Hughes Medical Institute, The Rockefeller University, New York, NY, USA

Cheng Lyu & Gaby Maimon

2. Mortimer B. Zuckerman Mind Brain Behavior Institute, Department of Neuroscience, Columbia University, New York, NY, USA

L. F. Abbott

Contributions

C.L. and G.M. conceived of the project. C.L. performed the experiments and analysed the data. C.L., G.M. and L.F.A. jointly interpreted the data and decided on new experiments. L.F.A. developed and implemented the formal models. C.L. wrote the initial draft of the paper, which was then edited by G.M. and L.F.A.

Corresponding author

Correspondence to [Gaby Maimon](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

[Extended Data Fig. 1 Characterizing the anatomy and physiology of hΔB cells, showing that sytGCaMP and RGEKO1a yield similar EPG phase estimates in the ellipsoid body, quantifying the EPG phase tracking of the closed loop dot, and evidence that the hΔB phase tracks the fly's traveling direction in walking flies.](#)

a, At least sixteen somas are labeled by the hΔB split-Gal4 line used in this paper. By comparison, the hemibrain connectome (v1.1) reports nineteen hΔB cells²⁰. **b**, GFP expression of the hΔB split Gal4 in the fan-shaped body. **c**, Same as panel **b**, but not showing the anti-nc82 neuropil stain. **d**, hΔB cells from hemibrain connectome v1.1²⁰. **e**, Top, hΔB GCaMP7f signal in a tethered, flying fly experiencing optic-flow (in the time window bracketed by the vertical dashed lines) with foci of expansion that simulate the following directions of travel: 180° (backward), -120°, -60°, 0° (forward), 60°, 120°. Bottom, Phase-nulled and averaged hΔB activity patterns in the fan-shaped body, calculated from the above [Ca²⁺] signals in the last 2.5 s of optic flow presentation. Population means with s.e.m. are

shown. **f**, Same as panel **e**, but with h Δ B sytGCaMP7f signal. Note that the single-bump structure in the sytGCaMP7f signal is clearer than the structure in the cytoplasmic GCaMP7f signal, which is consistent with sytGCaMP7f biasing GCaMP to axonal compartments of h Δ Bs. **g**, Probability distributions of the difference between the EPG phase and the bright dot's angular position, without and with optic flow. **h**, Circular standard deviation of the EPG phase – dot position distributions, without and with optic flow. Two-tailed unpaired t-test was performed. **i**, Correlations between the angular velocities of the EPG phase and the visual landmark position under different conditions. The first two columns use the same data as in panels **g** and **h**. The third and fourth columns use data from simultaneous GCaMP7f imaging of EPG cells and PFR cells in tethered-flying flies with a closed-loop dot. The fifth column use data from GCaMP6m imaging of EPG cells in tethered-walking flies with a closed-loop bar. Two-tailed one sample t-tests were performed against zero. P values are 1.7e-3, 1.3e-4, 5.3e-4, 1.2e-5 and 3.6e-4 comparing each column (from left to right) to zero, respectively. The relatively low, but significantly different from zero, r values show that the EPG phase tracks, even if poorly, the rotation of the landmark. The EPG phase measured in walking experiments tracks the closed-loop stimulus better than in tethered flight. See Main Text for possible technical reasons for why one would observe this difference. The fact that EPG-phase tracking of the closed loop dot is better when we co-imaged EPG cells and PFR cells compared to when we imaged EPG cells and h Δ Bs argues that the flies' genetic background (and thus how reliably flies perform tethered flight) can also quantitatively impact these measures. **j**, Angular-velocity correlations of the EPG phase and the visual landmark position under different conditions as a function of the time-lag between the two velocity signals. Same data as in panel **i**, but data with and without optic flow are lumped together. Correlation is highest at 290 ms, 260 ms and 375 ms for the three panels from left to right, respectively. Thus, we used time lags of 275 ms (mean of 290 and 260) and 375 ms for calculating the correlations in flight and walking experiments in panel **i**, respectively. **k**, Probability distribution of the angular position of the dot on the arena. Same data as in panels **g** and **h**, but data with and without optic flow are lumped together. We tested the uniformity of the distribution across angles using reduced χ^2 test. P value is > 0.995, meaning that we cannot reject the hypothesis that the dot position is not evenly distributed on

the arena. **I**, Circular standard deviation of the EPG phase minus the h Δ B phase distributions, without and with optic flow. Same data as in panels **g**, **h**. Two-tailed unpaired t-test was performed. P value equals 1.3e-6. **m**, Correlations between the EPG phase and the h Δ B phase. Same data as in panels **g**, **h** and **l**. Two-tailed unpaired t-test was performed. P value equals 3.9e-4. **n**, Data collected from tethered flies walking on a floating ball in complete darkness are shown in this panel and all subsequent panels in this figure. Sample time series of simultaneously imaged EPG and h Δ B Gal4 lines. Top two traces show [Ca²⁺] signals. Third trace shows the phase estimates of the two bumps. Bottom two traces show the forward velocity and sideslip velocity of the fly. Quasi-unidirectional walking bouts are labeled with walking directions indicated. **o**, Probability distribution of the difference between EPG phase and h Δ B phase from time segments where flies were walking in three different general directions (Methods). **p**, EPG – h Δ B phase as a function of the egocentric traveling direction. Gray: individual fly circular means. Black: population circular mean and s.e.m. The sign of EPG – h Δ B phase deviations seen here, in walking, are consistent with the signs observed in flight, for the same directions of backward-left and backward-right travel. Watson-Williams multi-sample tests were performed. P values are 1.6e-3 and 2.6e-6 comparing the 1st and 3rd columns (from left to right) to the 2nd column, respectively.

Extended Data Fig. 2 PFN_d and PFN_v activity bumps in the bridge are phase aligned with the EPG heading signal.

a, Sample trace, in tethered flight without optic flow, of simultaneously imaged GCaMP6m in EPG cells and jRGECO1a in PFN_d cells reveals that the activity bumps of these two cell classes are phase aligned in the bridge. **b**, Probability distribution of the EPG - PFN_d phase in tethered flight without optic flow. In this panel and throughout, the single fly data are in light gray and the population mean is in black. **c**, **d**, Same as panels **a**, **b**, but for GCaMP6m in EPG cells and jRGECO1a in PFN_v cells. **e**, Top three rows, sample trace of simultaneously imaged GCaMP6m in EPG cells and jRGECO1a in PFN_v cells in a tethered, flying fly experiencing optic flow (in the time window bracketed by the vertical dashed lines) with foci of expansion that simulate the following directions of travel: -120°, 0°

(forward), 120°. Bottom, circular-mean phase difference between EPG cells and PFN_v cells. **f**, Probability distribution of the EPG - PFN_v phase under three optic flow conditions. **g**, Circular mean of the EPG – PFN phase and s.e.m. under different visual stimulus conditions. Watson-Williams multi-sample tests, P>0.66 when comparing any experimental group with 0°. Note that we only collected a full EPG-PFN, dual-imaging data set with optic flow (moving dots) with PFN_v cells because, for reasons that are not fully clear, the jRGECO1a signal was too weak in PFN_d cells to properly estimate the PFN_d phase outside of the context of stationary dots (i.e., during optic flow). When imaging PFN_d cells with a split-Gal4 driver and with GCaMP rather than with jRGECO1a (e.g., Fig. 3j-l), the signal is much brighter.

Extended Data Fig. 3 Δ7 cells are poised to help create sinusoidally shaped activity bumps in PFN_d, PFN_y, and PFR cells in the protocerebral bridge.

Connectivity data are based on those in neuPrint²⁰, hemibrain:v1.1. **a**, Two Δ7 cells from neuPrint reveal a graded increase and decrease in dendritic density across the bridge. **b**, Synapse-number matrix for detected synapses from EPG cells to Δ7 cells in the protocerebral bridge. Each row represents one Δ7 cell. **c**, Same data as in panel **b**, but plotting each Δ7 cell separately. **d**, Phase-nulled EPG-to-Δ7 synapse # across the glomeruli of the bridge, averaged across all 42 Δ7 cells, based on the data in panel **c**. The anatomical input strength from EPG cells to Δ7 cells is sinusoidally modulated across the bridge. **e**, Transforming the EPG activity pattern across the bridge (blue) into a predicted Δ7 activity pattern (green, bottom row) based on the synaptic density profile in panel **c** (schematized in the middle). We first calculated the dot product between the EPG activity vector and each Δ7 cell's EPG-to-Δ7 synapse-number vector (panel **c**). Then, for each glomerulus, we averaged the dot-product-output for all of the Δ7 cells that have axonal terminals in that glomerulus, thus creating the predicted activity value for that glomerulus. (The size of each green square here schematizes the # of synapses from EPG cells to the Δ7 cell of that type in that column; the intensity of each Δ7 row indicates the expected

output strength of each $\Delta 7$ cell type, after being driven by the EPG signal above.) We plot the inverted, predicted activity output from $\Delta 7$ cells in the bottom row (green) because $\Delta 7$ cells are glutamatergic⁴³ and glutamatergic neurons in the *Drosophila* central nervous system typically inhibit their postsynaptic targets (via Glu-Cl channels). After inverting the $\Delta 7$ activity one can then imagine simply averaging the $\Delta 7$ predicted-activity row with the EPG activity—with some relative weighting for the $\Delta 7$ and EPG curves—to generate the net drive to the many downstream neurons that receive both EPG and $\Delta 7$ input²⁰, like PFN cells. Note that the EPG activity bumps are slightly narrower than the sinusoidal fits whereas the $\Delta 7$ activity bumps are slightly wider than the sinusoidal fits. **f**, Same as panel **e**, but using the phase-nulled, averaged EPG GCaMP activity pattern from a previous study¹¹. Note although the EPG bump is narrower in these data from walking flies than in panel **e** from flying flies, the shape of the predicted $\Delta 7$ output remains similar. **g**, Same as panel **e**, but starting with (imagined) EPG activity where there is only one active glomerulus on each side of the bridge. Note that the shape of the predicted $\Delta 7$ output remains similar to that in panels **e**, **f**. **h**, Measured, phase-nulled activity profiles from PFN_d , PFN_v and PFR cells. Thin lines: individual flies. Thick lines: population average. All three activity patterns conform well to their sinusoidal fits (gray dashed lines) (see [Methods](#) for goodness of fit). We hypothesize that the sinusoidal activity patterns in bridge columnar cells like PFN_d , PFN_v , PFR cells arises from the combined impact of EPG and $\Delta 7$ input. In other words, we posit that $\Delta 7$ cells ‘sinusoidalize’ the EPG bumps in the bridge—that is, they function to broaden and smoothen the EPG input to the bridge, to create two sinusoidally shaped bumps in their recipient cells, with these bumps often functioning as explicit, 2D vector signals in the fan-shaped body.

[**Extended Data Fig. 4 LNO1 and SpsP cells have \$\[\text{Ca}^{2+}\]\$ responses that are strongly tuned to the fly’s egocentric translation direction—in both walking and flying flies—with responses suggesting that these cells provide sign-inverting input to \$\text{PFN}_v\$ and \$\text{PFN}_d\$ cells, respectively.**](#)

Connectivity data and cell-type names are based on those in neuPrint²⁰, hemibrain:v1.1. **a**, LNO1 neurons are a class of cells (two total neurons per side, four per brain) that receive extensive synaptic input outside the central complex and provide extensive synaptic input to PFN_v cells in the noduli, with each PFN_v cell on average receive 131 synapses from LNO1s²⁰. **b**, Mean GCaMP signals in PFN_v and LNO1 cells in the nodulus as a function of the simulated traveling direction of the fly (via open-loop optic flow). Dotted rectangle indicates a repeated-data column, in this panel and throughout. **c**, Single-fly (colored circles) and population means \pm s.e.m. (black bars) of the average signal in the final 2.5 s of the optic flow epoch. Sinusoidal fits shown in this panel (Methods), and throughout. **d**, Each SpsP cell (two total neurons per side, four per brain) receives extensive synaptic input outside the central complex and provides extensive synaptic input to PFN_d cells on one side of the protocerebral bridge, with each PFN_d cell on average receive 56 synapses from SpsP cells²⁰. **e**, Same as panel **b**, but mean GCaMP signals in PFN_d and SpsP cells in the bridge as a function of the simulated traveling direction of the fly (via open-loop optic flow). A closed-loop bright dot was not present on the LED display when collecting the PFN_d data. **f**, Same as panel **c**, but averaging the bridge signal in panel **e**. **g**, Same as panel **b**, but analyzing the PFN_d signal in the noduli. A closed-loop bright dot was not present on the LED display. **h**, Same as panel **c**, but averaging the nodulus signal in panel **g**. **i**, The optic-flow-simulated egocentric traveling angle at which the activity of each cell type is strongest is depicted with a line at the associated angle. Note that the left-vs-right angular differences measured in the noduli are smaller, and closer to 90°, than the left-vs-right angular differences measured in the bridge. This difference might be a purposeful shift in optic-flow tuning related to the use of orthogonal and non-orthogonal PFN axes under different behavioral contexts (see [Supplementary Text](#)) and/or originate from differences in how SpsP cells in the bridge and LNO1 cells in the noduli balance optic-flow with proprioceptive/efference-copy inputs to generate their signals. **j**, Data collected from tethered flies walking on a floating ball in complete darkness are shown in this panel and all subsequent panels in this figure. Mean PFN_v GCaMP signals in the bridge as a function of the fly's forward speed. **k**, Right-minus-left PFN_v GCaMP signals in the bridge

as a function of the fly's sideslip speed. **I–m**, Same as panel **j** and **k**, but analyzing LNO1 signals in the nodulus. **n, o**, Same as panel **j** and **k**, but analyzing PFN_d signals in the bridge. **p–q**, Same as panel **j, k**, but analyzing SpsP signals in the bridge. In panel **b, e, g, j–q**, thin lines represent single-fly means and thick lines represent population means. Note that PFN_v and LNO1 cells have sign-inverted responses, and that PFN_d and SpsP cells have sign-inverted responses. The response signs to optic-flow simulating the fly's body translating forward and leftward (rightward) in flight are the same as the signs of responses to the fly walking forward and side-slipping leftward (rightward) when walking. Thus, these data are consistent with all these neurons being sensitive to the fly's egocentric translation direction, as assessed via optic flow (dominantly) in flight, and via proprioception or efference-copy (dominantly) in walking.

Extended Data Fig. 5 Multiple, functionally relevant ways of indexing angles across the protocerebral bridge.

Connectivity data and cell-type names are based on those in neuPrint²⁰, hemibrain:v1.1. **a**, The previously described mapping between EPG dendritic locations in the ellipsoid body and axonal-terminal locations in the bridge²¹. Numbers ordered based on the location of each EPG cell in the ellipsoid body. **b**, EPG cells divide the ellipsoid body into 16 wedges, each 22.5° wide. Each glomerulus in the bridge inherits its angle, in our analysis here, based on the EPG projection pattern shown in panel **a**. The angles of the outer two bridge glomeruli—which do not receive standard EPG input, but only EPGt input²⁰—were inferred to have angles equal to the middle two glomeruli (0° and 22.5°, respectively) based on how other cell types (e.g., PEN cells) innervate the bridge, as discussed in past work¹¹. This angular assignment maintains a 45° step size between adjacent glomeruli on each side of the bridge, which seems natural due to symmetry considerations. (Note that EPGt cells map from the ellipsoid body to the outer two glomeruli of the bridge with a small angular offset compared to the pattern set up by the EPG cells that target the central 16 glomeruli—as reported by other studies²⁸—a caveat that slightly complicates our angular assignments; however, EPGt cells receive extensive axonal input in the bridge that has the potential to align their output signals with the rest of the bridge system.)

Glomeruli are numbered 1 to 18 from left to right, to aid the comparisons made below. **c**, Two $\Delta 7$ cells from neuPrint (and past work²¹) reveal that the axonal terminals of each $\Delta 7$ cell are 8-glomeruli apart (#5→#13 for cell A and #2→#10→#18 for cell B). This anatomy argues that any two glomeruli 8 apart, such as #5 and #13, will experience $\Delta 7$ output of equal strength. Compelling physiological evidence for this statement is available in the $[Ca^{2+}]$ signals of the PEN2 (equivalently, PEN_b) columnar cell class in the bridge, which is a strong anatomical recipient of $\Delta 7$ synapses²⁰ and shows $[Ca^{2+}]$ activity across the bridge—clearly dissociable from the activity in EPG cells—with consistently equal signal strength at glomeruli spaced 8 apart, perfectly following the $\Delta 7$ anatomical prediction (orange trace in Fig. 3d and data points in Extended Data Fig. 2i from ref. 15). Note that in the EPG indexing, shown in panel **b**, glomeruli #5 and #13, as examples, have angular indices that are not identical, but differ by 22.5°. **d**, Angles assigned to each bridge glomerulus based on the $\Delta 7$ axonal anatomy. Because the $\Delta 7$ output anatomy requires that any two glomeruli 8 apart, across the whole bridge, have the same angular index assignment, this results in a situation where all neighboring glomeruli have angular assignments that are separated by 45°. Note that almost all neighboring glomeruli are separated by 45° in the EPG mapping as well, except that, critically, in the EPG mapping the middle two glomeruli are separated by only 22.5°. This discontinuity is not evident in the $\Delta 7$ output. To create an angular indexing of the bridge for $\Delta 7$ s that accommodates the anatomical constraints just described—i.e., one that incorporates an additional 22.5° in the bridge representation of angular space and thus ‘erases’ the EPG discontinuity—we shifted the angular index for each glomerulus on the left bridge leftward by 11.25° relative to the EPG indexing and we shifted the angular index for each glomerulus on the right bridge rightward by 11.25° relative to the EPG indexing. **e**, The EPG indexing in panels **a**, **b** predicts that EPG activity in the left bridge (#2→#9) will be left-shifted by 22.5° compared to EPG activity in the right bridge (#10→#17). Indeed, when we overlapped the left- and right-bridge EPG signals we found the two curves are detectably offset from each other. **f**, To quantify the data from panel **e**, for each imaging frame in which the fly was flying, we calculated the phase of the EPG bump in the left and right bridge separately (via a population-vector average) and took the difference of these two angles (black bars: population mean and s.e.m.). We then averaged this angular difference across all

analyzed frames for the same fly. For EPG cells, this angular difference should be -22.5° if it follows the EPG indexing in panel **b** and it should be 0° if the activity follows the $\Delta 7$ indexing in panel **d**. Across a population of 9 flies, we found the angular difference is close to -22.5° , but shifted toward 0° by 4.6° , consistent with the fact that the EPG signal itself receives strong anatomical input from the $\Delta 7$ s and thus could be modulated in its shape to follow the $\Delta 7$ indexing, in principle²⁰. It seems that the $\Delta 7$ feedback to EPG cells reshapes its signal, but incompletely. **g, h**, Same as panel **e**, but analyzing the PFN_d and PFN_v activity in the bridge. Because PFN cells only innervate the outer 8 glomeruli in each side of the bridge (unlike EPG cells, which innervate the inner 8), we compared glomeruli #1→#8 in the left bridge overlapped with glomeruli #11→#18 in the right bridge here (the middle two glomeruli contain no signal for PFN cells). **i**, Same as panel **f**, but analyzing the PFN_d and PFN_v activity in the bridge. Black bars: population mean and s.e.m. Note that because PFN_d and PFN_v cells innervate (and thus we can only analyze) the outer 8 glomeruli of the bridge, the angular difference in phase estimates between the left- and right-bridge activity should be $+67.5^\circ$ if it follows the EPG indexing (panel **b**) and $+90^\circ$ if it follows the $\Delta 7$ indexing (panel **d**). We found that the average angular difference in both PFN_d and PFN_v cells is intermediate between $+67.5^\circ$ and $+90^\circ$, consistent with PFNs receiving functional inputs from both EPG cells and $\Delta 7$ cells. We use the angular offsets measured in this panel as the basis for slightly adjusting the PFN_d and PFN_v angular indices in the bridge to an intermediate value between the EPG and $\Delta 7$ indexing options, described above. We believe that this approach represents the most careful way to combine the known anatomy and physiology to determine the azimuthal angle that each PFN cell signals with its activity in driving the $h\Delta B$ neurites in the fan-shaped body, which we analyze in the next figure. **j**, Angles assigned to each bridge glomerulus for PFN_d cells, based on the EPG indices from panel **b** and the physiologically determined adjustment required, based on the measurements in panel **i**. **k**, Same as panel **j**, but for PFN_v cells.

Extended Data Fig. 6 Computing the angular shift implemented by the PFN-to- $h\Delta B$ connections.

Connectivity data and cell-type names are based on those in neuPrint²⁰, hemibrain:v1.1. **a**, The anatomical angle of each PFN_v cell is indicated based on which glomerulus it innervates in the protocerebral bridge, using the indexing described in Extended Data Fig. 5k. **b**, Same as panel **a**, but for PFN_d cells, using the indexing described in Extended Data Fig. 5j. **c**, Synapse-number matrix for detected synapses from PFN_v cells to hΔB cells in the fan-shaped body. Note that the two stripes in the heatmap represent PFN_v cells synapsing onto the dendritic regions of hΔB cells. **d**, Same as panel **c**, but for synapses from PFN_d cells to hΔB cells. Note that two of the five stripes in the heatmap represent PFN_d cells synapsing onto the dendritic regions of hΔB cells, whereas the other, brighter, three stripes represent PFN_d cell synapsing onto the axons of hΔB cells. The average # of synapses that each hΔB compartment (axon vs. dendrite) receives from PFN cells is indicated on the bottom. **e**, Because hΔB cells are postsynaptic to both PFN_v and PFN_d cells that project to the fan-shaped body from both sides of the bridge (panels **c**, **d**), each hΔB cell can be assigned an anatomical angle in four potential ways. To calculate the angle for an hΔB cell through its connection with the left-bridge PFN_v cells, for example, we averaged the anatomical angles of all the left-bridge PFN_v cells that connect to the hΔB cell in question, weighted by the number of synapses from that PFN_v cell to the hΔB cell. **f**, The anatomical angle of each hΔB cell calculated based on its monosynaptic inputs from left-bridge PFN_vs using the method described in panel **e** and data in panel **c**. **g**, Same as panel **f**, but calculations were made with right-bridge PFN_v inputs to hΔB cells. **h**, Same as panel **f**, but calculations were made with left-bridge PFN_d inputs to hΔB cells, using only the synapses formed on the axonal terminals of hΔB cells. (We test the impact of this assumption—of complete functional dominance of PFN_d axonal synapses to hΔB cells—below.) **i**, Same as panel **f**, but calculations were made with right-bridge PFN_d inputs to hΔB cells, using only axonal synapses. **j**, For each hΔB cell, we calculated the angular difference between the mean left-bridge PFN_d input and the mean right-bridge PFN_d inputs (i.e., the difference between data points in panels **h** and **i**) and we plot a histogram of those values. **k–m**, Same as **j** for the cell types indicated. **n**, The anatomically predicted angles for the coordinate axes of

the four PFN vectors, as projected to the fan-shaped body and interpreted by h Δ B axons and dendrites, calculated by averaging the histogram values in panels **j–m**, respectively. **o**, Same as panel **n**, but including all synapses from PFN_d to h Δ B cells, not just the axonal ones as in panel **n**. We weigh dendritic and axonal synapses by PFN_d to h Δ B cells equally in the panel **e** calculation. Note that the angles between four coordinate-frame axes do not change very much when also including the dendritic synapses from PFN_d to h Δ B cells, likely because they are less numerous than the axonal ones and the impact of the dendritic angles also seem to cancel out in their net effect (compare panels **o** and **n**). **p**, Same as panel **n**, but using the EPG indexing from Extended Data Fig. 5 instead of the adjusted PFN_v and PFN_d indexing.

Note that the EPG indexing makes the front angle between the left- and right-bridge PFN_d axes smaller. The same is true for the back angle between the left- and right-bridge PFN_v axes. **q**, Same as panel **n**, but using the Δ 7 indexing from Extended Data Fig. 5 instead of the PFN_v and PFN_d indexing. Note that the Δ 7 indexing makes the front and back angles broader than 90°, when used in isolation. This analysis suggests that EPG and Δ 7 inputs to PFNs are perfectly weighted to create axes that are orthogonal in our experiments in flying flies and also raise the possibility that orthogonality of this 4-vector system can be dynamically modulated via changing the weights of EPG and Δ 7 inputs to PFNs (see [Supplementary Text](#)).

Extended Data Fig. 7 PFR neurons track a variable similar to allocentric traveling direction in walking and flying flies.

a, Schematics of two example EPG cells, two example PFR cells and two example h Δ B cells, which are the anatomically dominant input to PFRs. **b**, Sample GCaMP7f frames of the EPG bump in the ellipsoid body and the PFR bump in the fan-shaped body. **c**, Top, EPG (blue) and PFR (purple) GCaMP7f signal in a tethered, flying fly experiencing optic-flow (in the time window bracketed by the vertical dashed lines) with foci of expansion that simulate the following directions of travel: 180° (backward), −120°, −60°, 0° (forward), 60°, 120°, 180° (backward; repeated data). Third row, EPG and PFR phases extracted from the above [Ca²⁺] signals. Fourth row,

circular-mean phase difference between EPG cells and PFR cells. Bottom two rows, average of left-minus-right and left-plus-right wingbeat amplitude. Single fly means: light gray. Population means: black. Dotted rectangle indicates a repeated-data column. **d**, EPG – PFR phase as a function of the egocentric traveling direction simulated by the optic flow, at three different speeds. Circular means were calculated in the last 2.5 s of optic flow presentation. Gray: individual fly circular means. Black: population circular mean and s.e.m. Dotted rectangle indicates a repeated-data column. (See [Methods](#) for how we calculate the optic flow speed.) Note that the data points deviate slightly from the unity line in a manner that means that the PFR phase is slightly shifted away from the traveling direction indicated by the optic flow and toward a frontal heading direction. The h Δ B data in Fig. [1h](#) does not show this deviation from unity. We performed two-tailed one-sample t-tests against the diagonal line for data points in the $\pm 60^\circ$ and $\pm 120^\circ$ columns for the 35 cm/s data from PFR cells here and the 35 cm/s data from h Δ B cells in Fig. [1h](#). For the PFR results on the left of panel **d**, P values are 4.7e-5, 4.7e-5, 5.4e-4 and 9.1e-3 for the -120° , -60° , $+60^\circ$ and $+120^\circ$ columns, respectively. For the h Δ B results in Fig. [1h](#), P values are 0.39, 0.88, 0.058 and 0.44 for the -120° , -60° , $+60^\circ$ and $+120^\circ$ columns, respectively. **e**, PFR phase as a function of the inferred allocentric traveling direction, calculated by assuming that the EPG phase indicates allocentric heading direction and adding to this angle, at every sample point, the optic-flow angle. Gray: individual fly means. Black: population mean. In panels **d** and **e**, data from the middle column (35cm/s) were the same as in panel **c**. **f**, Tethered, walking, $[Ca^{2+}]$ -imaging setup with a bright blue bar that rotates in closed loop with the fly's turns. **g**, Sample time series of simultaneously imaged EPG and PFR bumps in a tethered, walking fly. Top two traces show $[Ca^{2+}]$ signals. Third trace shows the phase estimates of the two bumps. Bottom trace shows the forward speed of the fly. **h**, Probability distributions of the EPG – PFR phase in walking and standing flies. Thin lines: single flies. Thick line: population mean. **i**, Circular mean of the EPG – PFR phase in walking and standing flies. Watson-Williams multi-sample tests, $P>0.63$ when comparing any experimental group with 0° . Gray dots: single fly values. Black bars: population means \pm s.e.m. **j**, Same as panel **i**, but plotting circular standard deviation. Two-tailed unpaired t-tests were performed. P value equals 0.042. **k**, Tethered-walking setup where we used a 617 nm LED focused on

the center of the fly's head to optogenetically trigger backward walking via activation of LC16 visual neurons expressing CsChrimson⁵⁰ (Methods). **I**, An example 2D trajectory of optogenetically triggered backward walking. An arrow is shown every ~0.1 seconds. Red arrows indicate backward walking during the red-light pulse; blue arrows indicate the 1.2 s before the red light turned on. **m**, Left, time series of EPG (blue) and PFR (purple) bumps and phase-estimates from the trajectory in panel **I**. Right, time series of forward velocity, sideslip velocity and the difference between the PFR and EPG phase in the trajectory shown in panel **I**. The $\Delta F/F$ heatmap range is more compressed here than in other plots because the PFR signal strength typically dips when the fly initiates backward walking (a phenomenon whose mechanism we have not yet explored). Nevertheless, clear moments where the PFR phase separates from the EPG phase are evident, even after the PFR signal strength has recovered, in this sample trace (and in others). **n**, Time series of the mean forward velocity, mean sideslip velocity and the circular mean of the difference between the PFR and EPG phase during backward walking, grouped by optogenetic trials in which the fly walked to the back left (left panel) or to the back right (right panel). The sign of PFR-EPG phase deviations seen here, in walking, are consistent with the signs observed in flight, for the same directions of backward-left and backward-right travel. Thin lines and gray dots: individual trials. Thick line and black dot: population mean (circular mean for bottom row). **o**, Circular mean and s.e.m. of the peak EPG – PFR phase during triggered left-backward and right-backward walking bouts (0.6 s to 1.4 s after the dashed lines in panel **n**). Watson-Williams multi-sample tests were performed and P value equals 1.6e-6.

Extended Data Fig. 8 Response-tuning in PFN, PFR and hAB neurons to the translation speed indicated by our optic-flow stimuli.

a, Top row: same optic-flow tuning curves as in panel **b**, plotted twice (left and right). Bottom row: phase-nulled PFN_d GCaMP activity across the bridge, averaged in the final 2.5 s of the optic flow epoch. We show responses to optic flow simulating traveling backward at four different speeds (left) and responses to optic flow simulating forward travel at four

different speeds (right). The mapping between bridge $[Ca^{2+}]$ signals and data points in the plots in subsequent panels is indicated (arrows) for a few example points, using measurements from left-bridge PFN_d cells, as an example. How we calculate the mean and amplitude of each bump is schematized. **b**, The population-averaged amplitude of the phase-nulled left-bridge PFN_d $[Ca^{2+}]$ activity in the final 2.5 s of the optic flow epoch, plotted as a function of the egocentric traveling direction simulated by the optic flow. The translational speed of optic flow increases across the four columns, from left to right. Gray lines: sinusoidal fits. S.e.m. are shown in this panel and throughout. **c**, Same as panel **b**, but analyzing the right-bridge PFN_d activity. **d**, Same as panel **b**, but analyzing the left-bridge PFN_v activity. **e**, Same as panel **b**, but analyzing the right-bridge PFN_v activity. **f**, Same as panel **b**, but analyzing the h Δ B activity in the fan-shaped body. **g**, Same as panel **b**, but analyzing the PFR activity in the fan-shaped body. **h**, Same as panel **b**, but analyzing the h Δ B activity in the fan-shaped body in non-flying flies. **i**, Same as panel **f**, but with a more zoomed-in y-axis. **j**, Amplitude of the four sinusoids in panel **b** to indicate how PFN_d responses, overall, scale with optic-flow translation speed. **k–m**, Same as panel **j**, but for the plots and cell type shown to the left. Note that the amplitudes of the PFN sinusoidal activity patterns are not only scaled by the traveling direction angle (panels **b–e**), but also by traveling speed (panels **j–m**). These plots make sense as a way to quantify the amplitude of sinusoidally modulated responses, like those of PFNs, but we also show, for completeness, the results of the same analysis for h Δ B and PFR cells, where this way of quantifying forward-speed tuning makes less sense. **n**, Mean of the four sinusoids in panel **f** to indicate how h Δ B responses, overall, scale with optic-flow translation speed. **o**, Same as panel **n**, but for the PFR plots shown to the left. **p**, Same as panel **n**, but for the h Δ B plots shown to the left. Note that response-scaling with speed in h Δ B and PFR cells was not consistent across all traveling directions (panels **f–h**). The fact that the speed tuning of h Δ B cells remains nonuniform across traveling directions in non-flying flies (panel **h**) suggests that this nonuniform tuning is not entirely due to an efference copy/proprioceptive signal being mismatched with backward optic-flow directions in tethered flight, though the interpretation of this nonuniform tuning will need to be resolved in future work. **q**, Same as panel **n**, but for the h Δ B plots shown to the left. **r**, Same

as panel **b**, but analyzing the mean (rather than the amplitude) of the left-bridge PFN_d [Ca²⁺] activity patterns. Gray lines: same sinusoidal fits from panel **b** with a vertical offset and a scale factor that is constant across all four speeds. The fact that our amplitude fits from panel **b** also fit the mean responses shown here well supports the hypothesis that the heading input and the optic-flow input to PFN cells are integrated multiplicatively (see [Methods](#)). **s–x**, Same as panel **r**, but analyzing the cell type indicated on the left side of the figure, for each row. See [Methods](#) for how the optic-flow speed was calculated.

Extended Data Fig. 9 The neural circuit described in this paper implements an egocentric-to-allocentric coordinate transformation.

a, Schematic of the computation implemented in the *Drosophila* central complex. Traveling-direction signals referenced to the body axis (i.e., optic flow signals in SpsP and LNO1 cells, which indicate the egocentric traveling angle, green) are converted into traveling-angle signals referenced to cues in the world (i.e., the hΔB bump position, which indicates allocentric traveling angle, red). **b**, Schematic of a very similar computation hypothesized to take place in monkey parietal cortex.

Extended Data Fig. 10 A traveling-direction signal computed via optic flow is robust to changes in the yaw angle of the fly's head.

a, A fly flying straight with the head aligned to the body axis. EPG and hΔB signals are aligned in the ellipsoid body and fan-shaped body, respectively. **b**, A fly flying straight forward with the head rotated 20° to the right. The EPG bump—assuming the EPG bump position tracks the fly's head (rather than body) direction—will rotate 20° counterclockwise. The hΔB bump, however, will remain pointing in the same allocentric traveling direction because the net effect of the EPG bump rotating 20° in one direction and the ego-motion signal from optic flow (not represented in the diagram) rotating 20° in the opposite direction is that the PFR/hΔB bump stably indicates the same traveling direction throughout.

Supplementary information

Supplementary Information

This file contains Supplementary Text and additional references.

Reporting Summary

Supplementary Video 1

A dynamic representation of the analytical model. (Top left) A fly walks in a circle while facing north/up the whole time. The instantaneous direction of travel (red vector) is decomposed into components (black vectors). (Top middle) The same trajectory decomposed into components rotated by 45°. The two dotted lines represent four axes pointing at ±45° and ±135°. Only the two projections with positive components are shown on any given frame (black vectors). (Top right) Four sets of central complex neurons are drawn: left- and right-bridge SpsPs and left- and right-nodulus LNO1s. The green saturation level represents the activity of each cell class, which is negatively and linearly correlated with the projection length of the fly's traveling vector onto the ±45° and ±135° axes. (Bottom right) The two sinusoidally shaped bumps of PFNd cells (brown, top) and the two sinusoidally shaped bumps of PFNv cells (orange, bottom) in the bridge are shown. PFNs receive sign-inverting drive from SpsPs in the bridge and/or LNO1s in the noduli. These inputs induce modulations of the PFN bump amplitudes, with each PFN class' activity showing a positive, linear correlation to the projection length of the fly's traveling vector onto one of the ±45° and ±135° axes. (Bottom middle) PFN sinusoids shown overlaid in the fan-shaped body, phase shifted based on how PFNs anatomically project from the bridge to the fan-shaped body. hΔB neurons sum the PFN sinusoids to generate a new sinusoid whose peak represents the fly's traveling direction (red curve). (Bottom left) The red dot position represents the peak position of the hΔB sinusoid. EPG neurons in the donut-shaped ellipsoid body express a bump of activity whose position tracks the fly's heading (blue dot). Because the fly is always oriented in the same direction, the blue bump is stationary

Supplementary Video 2

Video of simultaneously imaged EPG and PFR signals in a flying fly before, during and after presentation of an optic flow stimulus that simulates forward travel

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Lyu, C., Abbott, L.F. & Maimon, G. Building an allocentric travelling direction signal via vector computation. *Nature* **601**, 92–97 (2022).
<https://doi.org/10.1038/s41586-021-04067-0>

- Received: 12 October 2020
- Accepted: 28 September 2021
- Published: 15 December 2021
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04067-0>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

Further reading

- [Transforming representations of movement from body- to world-centric space](#)
 - Jenny Lu
 - Amir H. Behbahani
 - Rachel I. Wilson

Nature (2021)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04067-0>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 15 December 2021](#)

Transforming representations of movement from body- to world-centric space

- [Jenny Lu](#)¹,
- [Amir H. Behbahani](#)² na1,
- [Lydia Hamburg](#)³ na1,
- [Elena A. Westeinde](#)¹ na1,
- [Paul M. Dawson](#)¹,
- [Cheng Lyu](#)⁴,
- [Gaby Maimon](#) [ORCID: orcid.org/0000-0003-1219-5856](#)⁴,
- [Michael H. Dickinson](#) [ORCID: orcid.org/0000-0002-8587-9936](#)²,
- [Shaul Druckmann](#) [ORCID: orcid.org/0000-0003-0068-3377](#)³ &
- [Rachel I. Wilson](#) [ORCID: orcid.org/0000-0001-8573-9266](#)¹

[Nature](#) volume **601**, pages 98–104 (2022)

- 2936 Accesses
- 4 Citations
- 56 Altmetric
- [Metrics details](#)

Subjects

- [Network models](#)

- [Sensorimotor processing](#)

Abstract

When an animal moves through the world, its brain receives a stream of information about the body's translational velocity from motor commands and sensory feedback signals. These incoming signals are referenced to the body, but ultimately, they must be transformed into world-centric coordinates for navigation^{1,2}. Here we show that this computation occurs in the fan-shaped body in the brain of *Drosophila melanogaster*. We identify two cell types, PFNd and PFNv^{3,4,5}, that conjunctively encode translational velocity and heading as a fly walks. In these cells, velocity signals are acquired from locomotor brain regions⁶ and are multiplied with heading signals from the compass system. PFNd neurons prefer forward–ipsilateral movement, whereas PFNv neurons prefer backward–contralateral movement, and perturbing PFNd neurons disrupts idiothetic path integration in walking flies⁷. Downstream, PFNd and PFNv neurons converge onto hΔB neurons, with a connectivity pattern that pools together heading and translation direction combinations corresponding to the same movement in world-centric space. This network motif effectively performs a rotation of the brain's representation of body-centric translational velocity according to the current heading direction. Consistent with our predictions, we observe that hΔB neurons form a representation of translational velocity in world-centric coordinates. By integrating this representation over time, it should be possible for the brain to form a working memory of the path travelled through the environment^{8,9,10}.

This is a preview of subscription content

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Fig. 1: PFN neurons that encode heading and translational velocity.

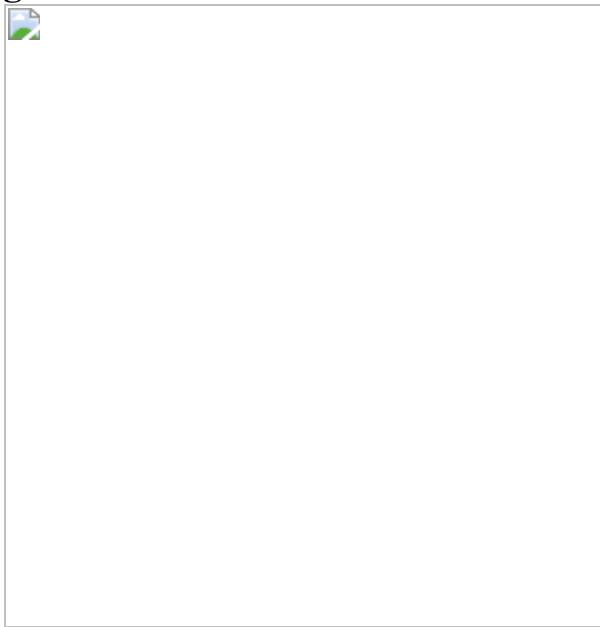


Fig. 2: Velocity tuning in PFNd neurons from graded release of inhibition.



Fig. 3: Behaviour and connections downstream from PFN neurons.

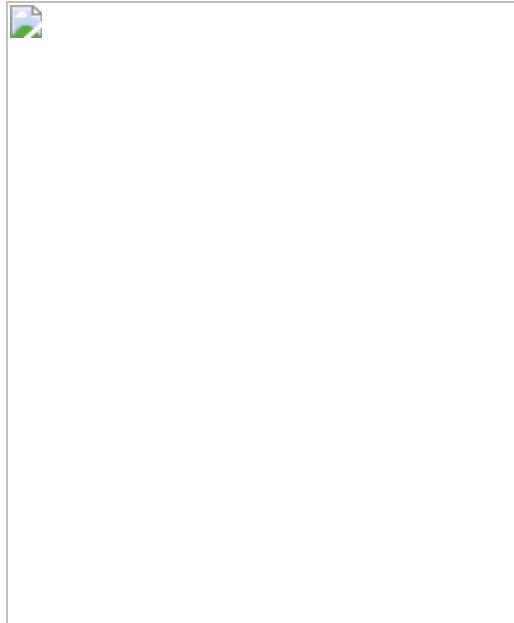
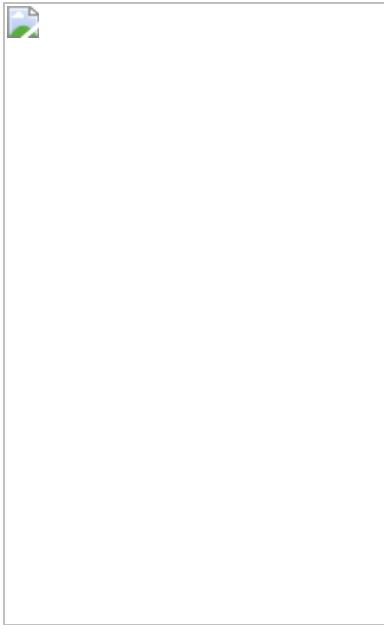


Fig. 4: From heading and body-centric velocity to world-centric velocity.



Data availability

The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Code availability

Code for implementing the computational model is available at
<https://github.com/druckmann-lab/Translational-velocity-and-heading-model>.

References

1. 1.

Bicanski, A. & Burgess, N. Neuronal vector coding in spatial cognition. *Nat. Rev. Neurosci.* **21**, 453–470 (2020).

2. 2.

Wang, C., Chen, X. & Knierim, J. J. Egocentric and allocentric representations of space in the rodent brain. *Curr. Opin. Neurobiol.* **60**, 12–20 (2020).

3. 3.

Wolff, T., Iyer, N. A. & Rubin, G. M. Neuroarchitecture and neuroanatomy of the *Drosophila* central complex: A GAL4-based dissection of protocerebral bridge neurons and circuits. *J. Comp. Neurol.* **523**, 997–1037 (2015).

4. 4.

Wolff, T. & Rubin, G. M. Neuroarchitecture of the *Drosophila* central complex: a catalog of nodulus and asymmetrical body neurons and a revision of the protocerebral bridge catalog. *J. Comp. Neurol.* **526**, 2585–2611 (2018).

5. 5.

Hulse, B. K. et al. A connectome of the *Drosophila* central complex reveals network motifs suitable for flexible navigation and context-dependent action selection. *eLife* **10**, e66039 (2021).

6. 6.

Namiki, S., Dickinson, M. H., Wong, A. M., Korff, W. & Card, G. M. The functional organization of descending sensory-motor pathways in *Drosophila*. *eLife* **7**, e34272 (2018).

7. 7.

Behbahani, A. H., Palmer, E. H., Corfas, R. A. & Dickinson, M. H. *Drosophila* re-zero their path integrator at the center of a fictive food patch. *Curr. Biol.* **31**, 4534–4546.e5 (2021).

8. 8.

Kim, I. S. & Dickinson, M. H. Idiothetic path integration in the fruit fly *Drosophila melanogaster*. *Curr. Biol.* **27**, 2227–2238 e3 (2017).

9. 9.

Müller, M. & Wehner, R. Path integration in desert ants, *Cataglyphis fortis*. *Proc. Natl Acad. Sci. USA* **85**, 5287–5290 (1988).

10. 10.

Esch, H. & Burns, J. Distance estimation by foraging honeybees. *J. Exp. Biol.* **199**, 155–162 (1996).

11. 11.

Ronacher, B. Path integration as the basic navigation mechanism of the desert ant *Cataglyphis fortis* (Forel, 1902) (Hymenoptera: Formicidae). *Myrmecol. News* **11**, 53–62 (2008).

12. 12.

Corfas, R. A., Sharma, T. & Dickinson, M. H. Diverse food-sensing neurons trigger idiothetic local search in *Drosophila*. *Curr. Biol.* **29**, 1660–1668.e4 (2019).

13. 13.

Ronacher, B. D. & Wehner, R. Desert ants *Cataglyphis fortis* use self-induced optic flow to measure distance travelled. *J. Comp. Physiol. A* **177**, 21–27 (1995).

14. 14.

Schöne, H. Optokinetic speed control and estimation of travel distance in walking honeybees. *J. Comp. Physiol. A* **179**, 587–592 (1996).

15. 15.

Wittlinger, M., Wehner, R. & Wolf, H. The ant odometer: stepping on stilts and stumps. *Science* **312**, 1965–1967 (2006).

16. 16.

Tuthill, J. C. & Wilson, R. I. Mechanosensation and adaptive motor control in insects. *Curr. Biol.* **26**, R1022–R1038 (2016).

17. 17.

Seelig, J. D. & Jayaraman, V. Neural dynamics for landmark orientation and angular path integration. *Nature* **521**, 186–191 (2015).

18. 18.

Turner-Evans, D. et al. Angular velocity integration in a fly heading circuit. *eLife* **6**, e23496 (2017).

19. 19.

Green, J. et al. A neural circuit architecture for angular integration in *Drosophila*. *Nature* **546**, 101–106 (2017).

20. 20.

Stone, T. et al. An anatomically constrained model for path integration in the bee brain. *Curr. Biol.* **27**, 3069–3085.e11 (2017).

21. 21.

Shiozaki, H. M., Ohta, K. & Kazama, H. A multi-regional network encoding heading and steering maneuvers in *Drosophila*. *Neuron* **106**, 126–141.e5 (2020).

22. 22.

Dana, H. et al. High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nat. Methods* **16**, 649–657 (2019).

23. 23.

Reiser, M. B. & Dickinson, M. H. A modular display system for insect behavioral neuroscience. *J. Neurosci. Methods* **167**, 127–139 (2008).

24. 24.

Scheffer, L. K. et al. A connectome and analysis of the adult *Drosophila* central brain. *eLife* **9**, e57443 (2020).

25. 25.

Zheng, Z. et al. A complete electron microscopy volume of the brain of adult *Drosophila melanogaster*. *Cell* **174**, 730–743 (2018).

26. 26.

Eckstein, N. et al. Neurotransmitter classification from electron microscopy images at synaptic sites in *Drosophila*. Preprint at <https://doi.org/10.1101/2020.06.12.148775> (2020).

27. 27.

Liu, W. W. & Wilson, R. I. Glutamate is an inhibitory neurotransmitter in the *Drosophila* olfactory system. *Proc. Natl Acad. Sci. USA* **110**, 10294–10299 (2013).

28. 28.

Lacin, H. et al. Neurotransmitter identity is acquired in a lineage-restricted manner in the *Drosophila* CNS. *eLife* **8**, e43701 (2019).

29. 29.

Hampel, S., Franconville, R., Simpson, J. H. & Seeds, A. M. A neural command circuit for grooming movement control. *eLife* **4**, e08758 (2015).

30. 30.

DeAngelis, B. D., Zavatone-Veth, J. A. & Clark, D. A. The manifold structure of limb coordination in walking *Drosophila*. *eLife* **8**, e46409 (2019).

31. 31.

Rayshubskiy, A. et al. Neural circuit mechanisms for steering control in walking *Drosophila*. Preprint at <https://doi.org/10.1101/2020.04.04.024703> (2020).

32. 32.

Wittmann, T. & Schwegler, H. Path integration—a network model. *Biol. Cybern.* **73**, 569–575 (1995).

33. 33.

Homberg, U., Heinze, S., Pfeiffer, K., Kinoshita, M. & el Jundi, B. Central neural coding of sky polarization in insects. *Philos. Trans. R. Soc. Lond. B* **366**, 680–687 (2011).

34. 34.

Vinepinsky, E. et al. Representation of edges, head direction, and swimming kinematics in the brain of freely-navigating fish. *Sci. Rep.* **10**, 14762 (2020).

35. 35.

Wang, C. et al. Egocentric coding of external items in the lateral entorhinal cortex. *Science* **362**, 945–949 (2018).

36. 36.

Solstad, T., Boccara, C. N., Kropff, E., Moser, M.-B. & Moser, E. I. Representation of geometric borders in the entorhinal cortex. *Science* **322**, 1865–1868 (2008).

37. 37.

Savelli, F., Yoganarasimha, D. & Knierim, J. J. Influence of boundary removal on the spatial representations of the medial entorhinal cortex. *Hippocampus* **18**, 1270–1282 (2008).

38. 38.

Deshmukh, S. S. & Knierim, J. J. Influence of local objects on hippocampal representations: landmark vectors and memory. *Hippocampus* **23**, 253–267 (2013).

39. 39.

Byrne, P., Becker, S. & Burgess, N. Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychol. Rev.* **114**, 340–375 (2007).

40. 40.

Bicanski, A. & Burgess, N. A neural-level model of spatial memory and imagery. *eLife* **7**, e33752 (2018).

41. 41.

Lyu, C., Abbott, L. F. & Maimon, G. Building an allocentric travelling direction signal via vector computation. *Nature* <https://doi.org/10.1038/s41586-021-04067-0> (2021).

42. 42.

Srinivasan, M., Zhang, S., Lehrer, M. & Collett, T. Honeybee navigation en route to the goal: visual flight control and odometry. *J. Exp. Biol.* **199**, 237–244 (1996).

43. 43.

Wilson, R. I. & Laurent, G. Role of GABAergic inhibition in shaping odor-evoked spatiotemporal patterns in the *Drosophila* antennal lobe. *J. Neurosci.* **25**, 9069–9079 (2005).

44. 44.

Pfeiffer, B. et al. Refinement of tools for targeted gene expression in *Drosophila*. *Genetics* **186**, 735–755 (2010).

45. 45.

Tirian, L. & Dickson, B. J. The VT GAL4, LexA, and split-GAL4 driver line collections for targeted expression in the *Drosophila* nervous system. Preprint at <https://doi.org/10.1101/198648> (2017).

46. 46.

Nern, A., Pfeiffer, B. D. & Rubin, G. M. Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. *Proc. Natl Acad. Sci. USA* **112**, E2967–E2976 (2015).

47. 47.

Miyamoto, T., Slone, J., Song, X. & Amrein, H. A fructose receptor functions as a nutrient sensor in the *Drosophila* brain. *Cell* **151**, 1113–1125 (2012).

48. 48.

Klapoetke, N. C. et al. Independent optical excitation of distinct neural populations. *Nat. Methods* **11**, 338–346 (2014).

49. 49.

Pfeiffer, B. D., Truman, J. W. & Rubin, G. M. Using translational enhancers to increase transgene expression in *Drosophila*. *Proc. Natl Acad. Sci. USA* **109**, 6626–6631 (2012).

50. 50.

von Reyn, C. R. et al. A spike-timing mechanism for action selection. *Nat. Neurosci.* **17**, 962 (2014).

51. 51.

Pologruto, T. A., Sabatini, B. L. & Svoboda, K. ScanImage: flexible software for operating laser scanning microscopes. *Biomed. Eng. Online* **2**, 13 (2003).

52. 52.

Gouwens, N. W. & Wilson, R. I. Signal propagation in *Drosophila* central neurons. *J. Neurosci.* **29**, 6239–6249 (2009).

53. 53.

Moore, R. J. et al. FicTrac: a visual method for tracking spherical motion and generating fictive animal paths. *J. Neurosci. Methods* **225**, 106–119 (2014).

54. 54.

Morimoto, M. M. et al. Spatial readout of visual looming in the central brain of *Drosophila*. *eLife* **9**, e57685 (2020).

55. 55.

Pnevmatikakis, E. A. & Giovannucci, A. NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. *J. Neurosci. Methods* **291**, 83–94 (2017).

56. 56.

Kempter, R., Leibold, C., Buzsáki, G., Diba, K. & Schmidt, R. Quantifying circular–linear associations: hippocampal phase precession. *J. Neurosci. Methods* **207**, 113–124 (2012).

57. 57.

Clements, J. et al. neuPrint: Analysis Tools for EM Connectomics. Preprint at <https://doi.org/10.1101/2020.01.16.909465> (2020).

58. 58.

Bates, A. S. et al. The natverse, a versatile toolbox for combining and analysing neuroanatomical data. *eLife* **9**, e53350 (2020).

59. 59.

Tobin, W. F., Wilson, R. I. & Lee, W.-C. A. Wiring variations that enable and constrain neural computation in a sensory microcircuit. *eLife* **6**, e24838 (2017).

Acknowledgements

This study benefited from the public release of the hemibrain connectome by the FlyEM Team at Janelia. We thank I. S. Haber and A. A. Li for tracing and annotation in the full adult female brain dataset (FAFB)²⁵; N. Eckstein, A. S. Bates, J. Funke and G. X. E. Jefferis for neurotransmitter predictions based on those data; J. Omoto for assistance with behavioural experiments; W. B. Dickson for sharing modified FicTrac software and machining help; T. Wolff, G. M. Rubin, V. Jayaraman, G. Card, B. D. Pfeiffer, D. J. Anderson and H. Amrein for providing fly stocks; H. H. Yang, M. A. Basnak, M. J. Marquis, Y. E. Fisher, T. Okubo, A. Rayshubskiy, C. D. Harvey, B. L. de Bivort, J. Drugowitsch, B. el Jundi, C. Pehlevan, J. A. Assad and the Wilson laboratory for discussions. This work was supported by the Harvard Medical School Neurobiology Imaging Facility (NINDS P30 #NS072030), the HMS Research Computing Group O2 cluster, and the HMS Research Instrumentation Core Facility. This study was supported by NIH grants T32 GM007753, F30 DC017698 (to J.L.), R01 EB028171 (to S.D.), and U19 NS104655 (to M.H.D., S.D. and R.I.W.). R.I.W. and G.M. are HHMI Investigators.

Author information

Author notes

1. These authors contributed equally: Amir H. Behbahani, Lydia Hamburg, Elena A. Westeinde

Affiliations

1. Department of Neurobiology and Howard Hughes Medical Institute,
Harvard Medical School, Boston, MA, USA

Jenny Lu, Elena A. Westeinde, Paul M. Dawson & Rachel I. Wilson

2. Division of Biology and Biological Engineering, California Institute of
Technology, Pasadena, CA, USA

Amir H. Behbahani & Michael H. Dickinson

3. Department of Neurobiology, Stanford University, Stanford, CA, USA

Lydia Hamburg & Shaul Druckmann

4. Laboratory of Integrative Brain Function and Howard Hughes Medical
Institute, The Rockefeller University, New York, NY, USA

Cheng Lyu & Gaby Maimon

Contributions

J.L. and R.I.W. conceived the project and coordinated the work. J.L. designed and performed imaging experiments and analyses. A.H.B. and M.H.D. designed and performed behavioural experiments and analyses; M.H.D. also provided hardware and software support. L.H. and S.D. designed, implemented and analysed the computational model. E.A.W. designed and performed electrophysiological experiments and analyses. P.M.D. performed MCFO experiments. C.L. and G.M. provided the h Δ B split-Gal4 line prior to publication. J.L. and R.I.W. analysed data and wrote the paper with input from all authors.

Corresponding author

Correspondence to [Rachel I. Wilson](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

[Extended Data Fig. 1 Walking statistics on a spherical treadmill.](#)

a. Distribution of forward × lateral, forward × rotational, and lateral × rotational velocities. Shown along each axis is the marginal distribution (gray lines on top right of each heatmap denote scale for the marginal distribution). Data are pooled across n=27 flies. We used the velocities recorded at the camera sampling rate (50 Hz) prior to down-sampling to volumetric calcium imaging rate. b. An example walking bout (30 s). Shown are the fly's forward, lateral, and rotational velocity as well as its heading (based on the position of the visual cue shown in closed loop; note that we used a visual closed loop gain of 0.8×, meaning that the landmark is displaced by an azimuthal angle equal to 0.8× the ball's yaw displacement). c. Fictive trajectory of the fly in 2D space based on the walking parameters in the example bout shown in b. The dotted line shows the calculated trajectory using only the forward velocity and the heading of the fly, ignoring the lateral velocity. The solid line shows the calculated trajectory using the forward velocity, lateral velocity, and heading of the fly. Note that the dotted line underestimates the curvature of the fly's path.

Extended Data Fig. 2 PFNd tuning properties.

a. Circular correlation between bump and cue position for PFNd (n=16 flies) and EPG neurons (n=5 flies). Note that PFNd bump position is not as correlated with heading as EPG activity is. This is because PFNd neurons conjunctively encode velocity and heading, whereas EPG neurons encode only heading. For example, when the fly walks forward right, the PFNd bump on the left diminishes in amplitude, and *vice versa*. When the left and right bumps have different amplitudes, this diminishes the accuracy of our estimate of the bump position. Moreover, when the fly steps backward, both PFNd bumps diminish in amplitude, which again makes it difficult to accurately estimate bump position. b. Normalized PFNd PB bump amplitude versus forward velocity (left), and lateral velocity (right). Gray lines are individual flies and the black line is the mean across flies (n=16 flies). Data from the right and left PB are combined, and lateral velocity is computed in the ipsilateral direction (so that, for PFNd.L neurons, leftward lateral velocity is positive and rightward lateral velocity is negative). The red line shows the linear fit to the mean line, with the fitted equation below each plot. c. Computation of preferred translational direction angle using the linear regression slopes for forward and lateral velocity. We used the ratio of the slopes of the linear fits to lateral and forward velocity to calculate the angle of preferred translational direction. d. PFNd data from Fig. 1g, re-plotted in polar coordinates. Here, normalized bump amplitude is displayed as a function of body-centric translation direction and binned by speed. e. Normalized PFNd bump amplitude versus velocity in the preferred translational direction (v^p). Data from the right and left PB are combined and binned by the fly's velocity orthogonal to the preferred translational direction (see schematic at right). Shown is the mean across flies (n=16 flies). Note that a positive value in the orthogonal axis is in the ipsilateral direction. Whereas there is a significant effect of velocity in the preferred direction (2-way ANCOVA, $P<10^{-10}$), there is no significant effect of velocity in the orthogonal direction ($p=0.97$). f. Normalized PFNd bump amplitude versus lateral velocity in the ipsilateral direction. Data from the right and left PB are combined, binned by ipsilateral rotational velocity, and averaged across flies (n=16 flies). Whereas there is a significant effect of lateral velocity (2-way ANCOVA, $P<10^{-10}$), there is no significant effect of rotational velocity ($p=0.59$). This analysis shows that

there is little or no systematic relationship between PFNd activity and rotational velocity once we account for the effect of lateral velocity. Note that, because rotational and lateral velocity are correlated, rotational velocity bins are asymmetrically populated. g. Circular correlation between bump and cue position for PFNd neurons when the fly walks in darkness ($n=7$ flies). h. Normalized bump amplitude versus lateral velocity in the ipsilateral direction, binned and color-coded by forward velocity, for PFNd neurons when the fly walks in darkness ($n=7$ flies). Lateral velocity is measured in the ipsilateral direction, and data from the right and left PB are combined and then averaged across flies. Both forward and lateral velocity have a significant effect (2-way ANCOVA, $P<10^{-10}$ and $P<10^{-5}$).

Extended Data Fig. 3 PFNv tuning properties.

a. Circular correlation between bump and cue position for EPG ($n=5$ flies, reproduced from Extended Data Fig. [2a](#)) and PFNv neurons ($n=11$ flies). Note that PFNv bump position is not as correlated with heading as EPG activity is. This is because PFNv neurons conjunctively encode velocity and heading, whereas EPG neurons encode only heading. In particular, PFNv bump amplitude is generally quite low when the fly is walking forward. b. Normalized PFNv PB bump amplitude versus forward velocity (left), and lateral velocity (right). Gray lines correspond to individual flies and the black line corresponds to the mean across flies ($n=11$ flies). Data for the right and left PB are combined, and lateral velocity is computed in the ipsilateral direction. The blue line shows the linear fit to the mean line, with the fitted equation below each plot. c. Computation of preferred translational direction angle using the linear regression slopes for forward and lateral velocity. We used the ratio of the slopes of the linear fits to lateral and forward velocity to calculate the angle of preferred translational direction. d. PFNv data from Fig. [1g](#), re-plotted in polar coordinates. Here, normalized bump amplitude is displayed as a function of body-centric translation direction and binned by speed. e. Normalized PFNv bump amplitude versus velocity along the angle of preferred translational direction (v^p). Data are combined between the right and left PB and binned by the velocity along the angle of translational movement orthogonal to the preferred direction (see schematic at right). Shown is the mean across flies ($n=11$ flies). The orthogonal directions for the right and left PFNv

population are shown (right); note that a positive value in the orthogonal axis remains in the contralateral direction for the given right/left population. Whereas there is a significant effect of velocity in the preferred direction (2-way ANCOVA, $P<10^{-10}$), there is no significant effect of velocity in the orthogonal direction ($p=0.30$). f. Normalized PFNv bump amplitude versus lateral velocity in the ipsilateral direction. Data for the right and left PB are combined, binned by the ipsilateral rotational velocity, and averaged across flies ($n=11$ flies). For this cell type, both lateral and rotational velocity have significant effects (2-way ANCOVA, $P<10^{-10}$ and $P<0.005$). Note that, because rotational and lateral velocity is correlated, rotational velocity bins are asymmetrically populated. g. Circular correlation between bump and cue position for PFNv neurons when the fly walks in darkness ($n=4$ flies). h. Normalized bump amplitude versus lateral velocity in the ipsilateral direction, binned and color-coded by forward velocity, for PFNv neurons when the fly walks in darkness ($n=4$ flies). Lateral velocity is measured in the ipsilateral direction, and data from the right and left PB are combined and then averaged across flies. Both forward and lateral velocity have a significant effect (2-way ANCOVA, $p<10^{-7}$ for each factor).

Extended Data Fig. 4 Interaction between heading and velocity tuning in PFNd neurons.

a. Firing rate versus v^p for all PFNd recordings. Data are divided into bins based on the proximity of the fly's heading to the neuron's preferred heading. Three of these cells are shown in Fig. 2b. b. Linear fits for one example cell. c. Fitted slope values (reproduced from Fig. 2b) and y-intercept values for all cells ($n=14$ cells in 9 flies). Horizontal lines indicate mean values. For both parameters, there is a statistically significant effect of heading (2-way paired t-tests, Bonferroni-corrected p values). However, the effect of heading on the slope is relatively large and consistent, as compared to the effect on the y-intercept, which is smaller and less consistent. This implies that the effect of heading (θ) on the cell's firing rate (f) is largely multiplicative, i.e., it controls the slope of the relationship between f and v^p , as in $f \propto (\cos(\theta - \theta^p) + a) v^p + b$ where θ^p , a , and b are constants. In our computational model (Fig. 4a-d), we use this same relationship, with $\theta^p=0$, $a=1$, $b=0$.

Extended Data Fig. 5 Connectomics analysis of inputs to PFNd and PFNv neurons.

a. Distribution of input synapses onto PFNd neurons from the hemibrain connectome²⁴, grouped by cell type. Shown are the top ten cell type inputs onto PFNd neurons; all other identified cell types are grouped into “Other.” Collectively, the distribution shown comprises 94.2% of all input synapses onto PFNd neurons. Numbers indicate the percentage of synapses contributed by each input cell type. Note that $\Delta 7$ neurons and FB3A/4C/4M neurons are major inputs to PFNd neurons, but we did not screen these neurons as part of our search for the origin of body-centric velocity signals in PFNd neurons, for the following reasons: $\Delta 7$ neurons: $\Delta 7$ population activity is known to encode the fly’s heading direction, reflecting the strong input to $\Delta 7$ neurons from EPG neurons. It has been proposed that the function of $\Delta 7$ neurons is to reshape the heading bump into a cosine-shaped activity profile^{5,41}. Thus, much of the “compass input” that we refer to in our study as originating from EPG neurons is probably due to the combined action of EPG neurons (which constitute the primary computational map of the compass system) and $\Delta 7$ neurons (which reshape and reinforce the compass system output). FB3A/4C/4M neurons: These neurons are FB tangential cells, meaning their axons run across the entire horizontal extent of the FB, perpendicular to PFNd dendrites⁵. Like other FB tangential cells, these neurons receive input from outside the central complex and they synapse onto a variety of cell types in the FB. There is evidence that FB tangential cells encode information about context, behavioral state, and internal physiological needs, including the need for sleep⁵. b. Input connectivity matrix for PFNd neurons, shown for the top ten input cell types. Connections comprising 3 or fewer synapses are not shown. Note that the cell types that provide major unilateral input to PFNd neurons are LNO2, IbSpsP, EPG, SpsP, and LNO1. c. Same as (a) but for PFNv neurons. Collectively, the distribution shown comprises 93.1% of all input synapses onto PFNv neurons.

Extended Data Fig. 6 LNO2 and h Δ B split-Gal4 line characterization.

a. GFP expression driven by the LNO2 split-Gal4 line: +; Mi{Trojan-p65AD.2}VGlut[MI04979-Tp65AD.2]; P{VT008681-Gal4.DBD}attP2. Shown is a coronal projection of a confocal stack through the anterior half of the brain. GFP staining is shown in green, and neuropil staining (nc82) is shown in magenta. The scale bar is 50 μ m. Note that, in addition to targeting LNO2 neurons in the LAL, there are some cells labeled in the superior brain which are not LNO2 cells. The observation that this VGlut-split-Gal4 construct drives expression in LNO2 neurons is evidence in support of the conclusion that LNO2 neurons are glutamatergic. b. Same as (a) but for individual optical slices. Shown are the location of the LNO2 cell bodies (left, arrows), neurites in the LAL (middle, arrows), and neurites in NO2 (right, arrows). Scale bars are 50 μ m. c. Skeleton of LNO2 neuron from the hemibrain dataset. Overlaid are the anatomical boundaries of the LAL and the NO (divided into subunits NO1, NO2, and NO3). The black sphere denotes the position of the cell body. There is one LNO2 neuron per hemisphere. d. MCFO labeling of a single LNO2 neuron from the LNO2-split Gal4 line. Scale bar is 50 μ m. e. On occasion, the LNO2 split-Gal4 line shows expression in NO3. Shown is an MCFO sample from the LNO2-split Gal4 line that labels this additional neuron in NO3 (arrow). Given that two channels (green and red) label the LNO2 on the ipsilateral side, whereas only one channel (red) shows the NO3-innervating neuron, this neuron appears to be a distinct neuron from LNO2. Scale bar is 50 μ m. f. Skeletons of two h Δ B neurons from the hemibrain dataset. Overlaid are the anatomical boundaries of the FB. Spheres denote soma positions. g. MCFO labeling of two h Δ B neurons from the h Δ B split Gal4 line +; P{R72B05-p65.AD}attP40; P{VT055827-Gal4.DBD}attP2. Scale bar is 20 μ m.

Extended Data Fig. 7 SpsP, LNO2, IbSpsP, and LNO1 physiology.

a. Schematic of SpsP and LNO2 input onto a single PFNd neuron. PFNd neurons have dendrites in the PB on the side ipsilateral to their soma, and dendrites in the NO on the side contralateral to their soma. As a result, PFNd neurons receive input from ipsilateral SpsP neurons and the contralateral LNO2 neuron. Thus, although SpsP and LNO2 neurons have opposite velocity preferences (Fig. 2c), they have congruent effects on PFNd neurons. b. SpsP and LNO2 activity as a fly walks in closed loop

with a visual cue. c. SpsP and LNO2 $\Delta F/F$ versus lateral velocity in the ipsilateral direction. Data for the right and left PB are combined, binned by the ipsilateral rotational velocity, and averaged across flies (n=8 flies for SPS, 4 flies for LNO2). Because rotational and lateral velocity are correlated, rotational velocity bins are asymmetrically populated. There is a significant effect of lateral velocity (2-way ANCOVA, $P<10^{-10}$ for both SpsP and LNO2) but not rotational velocity ($p=0.59$ for SpsP, $p=0.14$ for LNO2). Note however that SpsP activity increases when rotational *speed* is high, for both ipsi- and contralateral rotations. d. Control experiments for SpsP optogenetic activation. There is little effect of light in PFNd recordings from flies where an empty split-Gal4 line is combined with UAS-CsChrimson (n=3) or in flies with UAS-CsChrimson expressed under SpsP split-Gal4 control (ss52267) but reared in the absence of all-trans-retinal (ATR; n=3). We consistently see strong inhibition in flies that express UAS-CsChrimson under SpsP split-Gal4 control (ss52267) and that are raised on culture media containing ATR (n=9, reproduced from Fig. 2d). PFNd recordings were performed in TTX to isolate monosynaptic responses (see [Methods](#)). e. Each IbSpsP neuron receives input from the inferior bridge (IB) and SPS, and projects to a few adjacent PB glomeruli. f. Circular correlation between visual cue position and IbSpsP bump position (n=8 flies). Shown for comparison is the circular correlation for EPG neurons (n=5 flies), reproduced from Extended Data Fig. 2a. g. IbSpsP population activity in the PB as a fly walks in closed loop with a visual cue. h. Normalized IbSpsP bump amplitude versus forward velocity. Data are binned by lateral velocity in the ipsilateral direction, combined for the right and left PB, and averaged across flies (n=8 flies). There is a significant effect of lateral velocity ($P<0.01$) but not forward velocity ($p=0.65$, 2-way ANCOVA). i. Normalized IbSpsP bump amplitude in the PB, versus body-centric translational direction. Data are binned by speed. Lateral velocity is expressed in the direction ipsilateral to the imaged PB, allowing us to combine data from the right and left PB before averaging across flies (n=8 flies). j. Each LNO1 neuron receives input from the LAL and synapses onto PFNv and PFNd dendrites in the NO. k. LNO1 activity as a fly walks in closed loop with a visual cue. We used jGCaMP7s in these experiments (rather than jGCaMP7f) because LNO1 fluorescence was dim with jGCaMP7f. l. LNO1 activity versus forward velocity. Data for the left and right NO are combined, binned by lateral velocity in the ipsilateral

direction, and averaged across flies ($n=8$ flies). LNO1 activity decreases slightly with ipsilateral backward movement. There is a significant effect of both forward velocity ($P<10^{-10}$) and lateral velocity ($P<0.01$, 2-way ANCOVAs).

Extended Data Fig. 8 PFN→hΔB connectivity.

a. Schematized projections of the PFNd and PFNv populations, from the hemibrain connectome. Gray numbers denote PB glomeruli³. Note that the mapping from PB glomeruli to FB horizontal locations is the same for PFNd (red) and PFNv (blue). For each cell type, each half of the PB contains a complete heading map (black arrows) which is projected onto the full horizontal axis of the FB. b. Top: PFN→hΔB connection matrix from the hemibrain connectome, reproduced from Fig. 3g. Note that, for a given hΔB neuron, PFN projections from the left and right PB are horizontally shifted, corresponding to the morphologies in (a). Bottom: Permuted PFN→hΔB connection matrix. Here, the shifts between left and right PFN matrices are eliminated. We used this permuted connection matrix in Fig. 4d (“left-right shift eliminated”).

Extended Data Fig. 9 Model performance as a function of relative synaptic weight.

a. hΔB dendrites receive PFNd and PFNv inputs at their dendrites. By contrast, hΔB axon terminals receive PFNd inputs but no PFNv input. In the bar plot at right, each bar represents one hΔB neuron in the hemibrain connectome ($n = 19$ neurons). The computational model in Fig. 4a–d assigns an equal weight to all synapses, meaning that all connections are simply weighted by the number of synapses they contain, regardless of whether they are axo-dendritic or axo-axonic connections. b. To determine if the model might perform better if we treated these connections differently, we systematically varied the weight of PFN synapses onto hΔB dendrites versus axons, and we used the population vector average of hΔB activity to decode the fly’s simulated movement. Grayscale heatmap shows the error in translational direction encoding (left) and speed encoding (right), with lower values indicating more accurate encoding. Note that we

obtain the best translation direction encoding if we apply equal weight to axo-dendritic or axo-axonic connections (as we do in Fig. 4a–d). Speed encoding improves if we minimize the weight at the synapses onto h Δ B axons; this is because this reduces the contribution of PFNd inputs (relative to PFNv), and so it tends to reduce the disproportionate gain when the fly is walking in the preferred direction φ^p of the PFNd population (Fig. 4c). We do not know whether axo-dendritic and axo-axonic connections are actually weighted equally in the real network, but the fact that we observe good encoding of φ in the h Δ B population (Fig. 4h) suggests that these connections carry similar weight, at least as measured with jGCaMP7f. c. We also systematically varied the weight of PFNd and PFNv synapses. We obtain the best translation direction encoding if we apply equal weight to PFNd and PFNv connections (as we do in Fig. 4a–d). Speed encoding improves if we reduce PFNd weights, again because this reduces the disproportionate gain when the fly is walking in the preferred direction φ^p of the PFNd population (Fig. 4c).

Extended Data Fig. 10 h Δ B bump deviations.

- a. h Δ B $\Delta F/F$ in each FB column as a fly walks in closed loop with a visual cue. When the fly steps laterally (indicated by the arrowhead), the bump deviates from the cue. b. Histograms showing the difference between cue position and bump position, mean-centered in each experiment, and binned by translation direction; n=4 flies for h Δ B, 16 flies for PFNd, and 11 flies for PFNv, # = relatively poor correlation between cue and bump; these experiments are omitted from panel c. At more lateral translation angles, the h Δ B bump deviates away from where it would be when the fly is walking forward. c. Mean difference between cue position and bump position. Each set of connected symbols is one experiment. For h Δ B neurons (n=4 flies), we found the shift was significant when comparing left translation-heading deviations to centered translation-heading deviations ($P=0.0013$, 2-sided paired-sample t-test with Bonferroni-corrected $\alpha = 0.0167$, CI = $[-0.460, -0.191]$ radians) and when comparing right translation-heading deviations to centered translation-heading deviations ($P=0.0115$, $\alpha = 0.0167$, CI = $[-0.698, -0.0473]$ radians). For PFNd neurons (n=16 flies), the shift is not significant when comparing left translation-heading deviations to centered translation-heading deviations ($P=0.0215$, 2-sided paired-sample t-test with

Bonferroni-corrected $\alpha = 0.0167$, CI = [-0.180, 0.0044] radians) or when comparing right translation-heading deviations to centered translation-heading deviations ($P=0.4790$, $\alpha = 0.0167$, CI = [-0.0467, 0.0812] radians). For PFNv neurons ($n=9$ flies; 2 flies were excluded from our analysis), this shift is significant when comparing left translation-heading deviations to centered translation-heading deviations ($P=0.0011$, 2-sided paired-sample t-test with Bonferroni-corrected $\alpha = 0.0167$, CI = [0.0544, 0.222] radians) but not significant when comparing right translation-heading deviations to centered translation-heading deviations ($P=0.0313$, $\alpha = 0.0167$, CI = [-0.0135, 0.1848] radians); note that the shift is opposite to h Δ B neurons. d. Same as Fig. 4f-g but color-coded by fly ($n=28$ epochs in 10 flies for h Δ B, $n=22$ epochs in 6 flies for EPG). e. Maximum bump deviation versus φ , measured in all epochs ≥ 300 ms when the φ was consistent over the epoch. Within each fly, epochs are binned by φ and then averaged (\circ) before averaging across flies (\bullet). For h Δ B neurons, the data are close to the identity line (purple); while for EPG neurons, the data are close to the zero line (gold). $n=10$ flies for h Δ B, $n=10$ flies for EPG. f. Normalized h Δ B bump amplitude versus φ , binned by speed ($n=11$ flies).

Supplementary information

Reporting Summary

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Lu, J., Behbahani, A.H., Hamburg, L. *et al.* Transforming representations of movement from body- to world-centric space. *Nature* **601**, 98–104 (2022). <https://doi.org/10.1038/s41586-021-04191-x>

- Received: 22 December 2020
- Accepted: 28 October 2021
- Published: 15 December 2021
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04191-x>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

Further reading

- [**Building an allocentric travelling direction signal via vector computation**](#)
 - Cheng Lyu
 - L. F. Abbott
 - Gaby Maimon

Nature (2021)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04191-x>

- Article
- [Published: 01 December 2021](#)

Local circuit amplification of spatial selectivity in the hippocampus

- [Tristan Geiller](#) [ORCID: orcid.org/0000-0001-8957-7786^{1,2}](#),
- [Sadra Sadeh](#) [ORCID: orcid.org/0000-0001-8159-5461³](#),
- [Sebastian V. Rolotti^{1,2}](#),
- [Heike Blockus^{1,2}](#),
- [Bert Vancura^{1,2}](#),
- [Adrian Negrean^{1,2}](#),
- [Andrew J. Murray⁴](#),
- [Balázs Rózsa⁵](#),
- [Franck Polleux](#) [ORCID: orcid.org/0000-0003-4313-0481^{1,2,6}](#),
- [Claudia Clopath](#) [ORCID: orcid.org/0000-0003-4507-8648³](#) &
- [Attila Losonczy](#) [ORCID: orcid.org/0000-0002-7064-0252^{1,2,6}](#)

[Nature](#) volume 601, pages 105–109 (2022)

- 8935 Accesses
- 96 Altmetric
- [Metrics details](#)

Subjects

- [Hippocampus](#)
- [Long-term potentiation](#)

- [Network models](#)
- [Neural circuits](#)

Abstract

Local circuit architecture facilitates the emergence of feature selectivity in the cerebral cortex¹. In the hippocampus, it remains unknown whether local computations supported by specific connectivity motifs² regulate the spatial receptive fields of pyramidal cells³. Here we developed an *in vivo* electroporation method for monosynaptic retrograde tracing⁴ and optogenetics manipulation at single-cell resolution to interrogate the dynamic interaction of place cells with their microcircuitry during navigation. We found a local circuit mechanism in CA1 whereby the spatial tuning of an individual place cell can propagate to a functionally recurrent subnetwork⁵ to which it belongs. The emergence of place fields in individual neurons led to the development of inverse selectivity in a subset of their presynaptic interneurons, and recruited functionally coupled place cells at that location. Thus, the spatial selectivity of single CA1 neurons is amplified through local circuit plasticity to enable effective multi-neuronal representations that can flexibly scale environmental features locally without degrading the feedforward input structure.

This is a preview of subscription content

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.
VAT will be added later in the checkout.
Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Fig. 1: In vivo single-cell electroporation and monosynaptic rabies tracing in hippocampal region CA1.

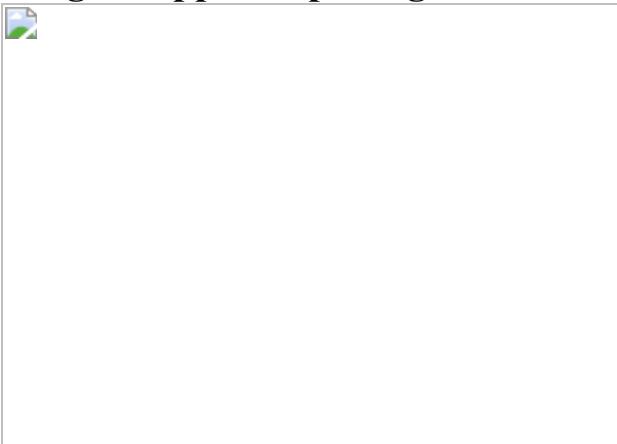


Fig. 2: Interneurons presynaptic to a place cell exhibit inverse spatial selectivity.



Fig. 3: Optogenetic place field induction in single pyramidal cells reorganizes interneuron networks.

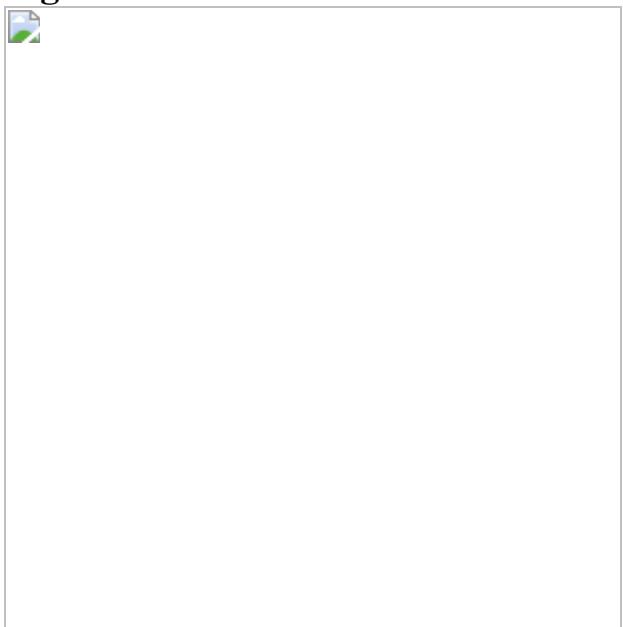
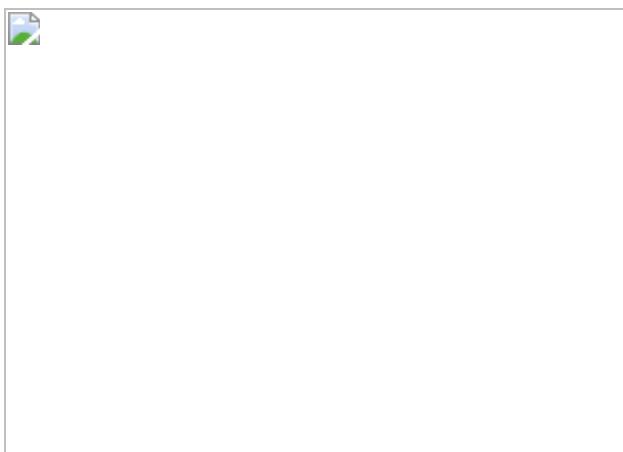


Fig. 4: Recruitment of local pyramidal cells during place field induction is consistent with a subnetwork architecture.



Data availability

All data are available from the corresponding authors upon reasonable request. [Source data](#) are provided with this paper.

Code availability

All custom codes are available from the corresponding authors upon reasonable request.

References

1. 1.

Miller, K. D. Canonical computations of cerebral cortex. *Curr. Opin. Neurobiol.* **37**, 75–84 (2016).

2. 2.

Jiang, X. et al. Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* **350**, aac9462 (2015).

3. 3.

O'Keefe, J. & Dostrovsky, J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* **34**, 171–175 (1971).

4. 4.

Wertz, A. et al. Single-cell-initiated monosynaptic tracing reveals layer-specific cortical network modules. *Science* **349**, 70–74 (2015).

5. 5.

Najafi, F. et al. Excitatory and inhibitory subnetworks are equally selective during decision-making and emerge simultaneously during

learning. *Neuron* **105**, 165–179 (2020).

6. 6.

Eichenbaum, H. A cortical–hippocampal system for declarative memory. *Nat. Rev. Neurosci.* **1**, 41–50 (2000).

7. 7.

Epsztein, J., Brecht, M. & Lee, A. K. Intracellular determinants of hippocampal CA1 place and silent cell activity in a novel environment. *Neuron* **70**, 109–120 (2011).

8. 8.

Stachenfeld, K. L., Botvinick, M. M. & Gershman, S. J. The hippocampus as a predictive map. *Nat. Neurosci.* **20**, 1643–1653 (2017).

9. 9.

Rossi, L. F., Harris, K. D. & Carandini, M. Spatial connectivity matches direction selectivity in visual cortex. *Nature* **588**, 648–652 (2020).

10. 10.

Sadeh, S. & Clopath, C. Inhibitory stabilization and cortical computation. *Nat. Rev. Neurosci.* **22**, 21–37 (2021).

11. 11.

Houweling, A. R. & Brecht, M. Behavioural report of single neuron stimulation in somatosensory cortex. *Nature* **451**, 65–68 (2008).

12. 12.

Cossell, L. et al. Functional organization of excitatory synaptic strength in primary visual cortex. *Nature* **518**, 399–403 (2015).

13. 13.

Kesner, R. P. & Rolls, E. T. A computational theory of hippocampal function, and tests of the theory: new developments. *Neurosci. Biobehav. Rev.* **48**, 92–147 (2015).

14. 14.

Reardon, T. R. et al. Rabies virus CVS-N2cδG strain enhances retrograde synaptic transfer and neuronal viability. *Neuron* **89**, 711–724 (2016).

15. 15.

Geiller, T., Fattah, M., Choi, J. S. & Royer, S. Place cells are more strongly tied to landmarks in deep than in superficial CA1. *Nat. Commun.* **8**, 14531 (2017).

16. 16.

Geiller, T. et al. Large-scale 3D two-photon imaging of molecularly identified CA1 interneuron dynamics in behaving mice. *Neuron* **108**, 968–983 (2020).

17. 17.

Bittner, K. C. et al. Conjunctive input processing drives feature selectivity in hippocampal CA1 neurons. *Nat. Neurosci.* **18**, 1133–1142 (2015).

18. 18.

Kaufman, A. M., Geiller, T. & Losonczy, A. A role for the locus coeruleus in hippocampal CA1 place cell reorganization during spatial reward learning. *Neuron* **105**, 1018–1026 (2020).

19. 19.

English, D. F. et al. Pyramidal cell–interneuron circuit architecture and dynamics in hippocampal networks. *Neuron* **96**, 505–520 (2017).

20. 20.

Yang, S. et al. Interlamellar CA1 network in the hippocampus. *Proc. Natl Acad. Sci. USA* **111**, 12919–12924 (2014).

21. 21.

Grienberger, C., Milstein, A. D., Bittner, K. C., Romani, S. & Magee, J. C. Inhibitory suppression of heterogeneously tuned excitation enhances spatial coding in CA1 place cells. *Nat. Neurosci.* **20**, 417–426 (2017).

22. 22.

Guzman, S. J., Schlögl, A., Frotscher, M. & Jonas, P. Synaptic mechanisms of pattern completion in the hippocampal CA3 network. *Science* **353**, 1117–1123 (2016).

23. 23.

Draguhn, A., Traub, R. D., Schmitz, D. & Jefferys, J. G. R. Electrical coupling underlies high-frequency oscillations in the hippocampus in vitro. *Nature* **394**, 189–192 (1998).

24. 24.

Hefft, S. & Jonas, P. Asynchronous GABA release generates long-lasting inhibition at a hippocampal interneuron-principal neuron synapse. *Nat. Neurosci.* **8**, 1319–1328 (2005).

25. 25.

Xu, H. T. et al. Distinct lineage-dependent structural and functional organization of the hippocampus. *Cell* **157**, 1552–1564 (2014).

26. 26.

Picardo, M. A. et al. Pioneer GABA cells comprise a subpopulation of hub neurons in the developing hippocampus. *Neuron* **71**, 695–709 (2011).

27. 27.

Caroni, P., Donato, F. & Muller, D. Structural plasticity upon learning: Regulation and functions. *Nat. Rev. Neurosci.* **13**, 478–490 (2012).

28. 28.

Magee, J. C. & Grienberger, C. Synaptic plasticity forms and functions. *Annu. Rev. Neurosci.* **43**, 95–117 (2020).

29. 29.

Mehta, M. R. From synaptic plasticity to spatial maps and sequence learning. *Hippocampus* **25**, 756–762 (2015).

30. 30.

Kaifosh, P., Zaremba, J. D., Danielson, N. B. & Losonczy, A. SIMA: Python software for analysis of dynamic fluorescence imaging data. *Front. Neuroinform.* **8**, 80 (2014).

31. 31.

Keemink, S. W. et al. FISSA: a neuropil decontamination toolbox for calcium imaging signals. *Sci. Rep.* **8**, 3493 (2018).

32. 32.

Bezaire, M. J. & Soltesz, I. Quantitative assessment of CA1 local circuits: knowledge base for interneuron-pyramidal cell connectivity. *Hippocampus* **23**, 751–785 (2013).

Acknowledgements

A.L. is supported by the National Institute of Mental Health (NIMH) 1R01MH124047 and 1R01MH124867; the National Institute of Neurological Disorders and Stroke (NINDS) 1U19NS104590 and 1U01NS115530; and the Kavli Foundation. B.V. is supported by (NIH) T32GN007367 and (NIMH) F30MH125628. A.J.M. is supported by the Gatsby Charitable Foundation (GAT3361) and the Wellcome Trust (090843/F/09/Z). S.V.R. is supported by (NIMH) F31MH117892. F.P. is supported by (NINDS) R01NS067557 and R21NS109753-442 01A1. H.B. is supported by (NINDS) K99NS115984-01. C.C. is supported by the Biotechnology and Biological Sciences Research Council (BB/N013956/1 and BB/N019008/1), the Wellcome Trust (200790/Z/16/Z), the Simons Foundation (564408) and the EPSRC (EP/R035806/1). We thank the Zuckerman Institute's Cellular Imaging platform for instrument use and technical advice. Imaging was performed with support from the Zuckerman Institute's Cellular Imaging platform and the NIH (1S10OD023587-01). We thank S. Siegelbaum and members of the Losonczy laboratory for comments on the manuscript.

Author information

Affiliations

1. Department of Neuroscience, Columbia University, New York, NY, USA

Tristan Geiller, Sebastian V. Rolotti, Heike Blockus, Bert Vancura, Adrian Negrean, Franck Polleux & Attila Losonczy

2. Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA

Tristan Geiller, Sebastian V. Rolotti, Heike Blockus, Bert Vancura, Adrian Negrean, Franck Polleux & Attila Losonczy

3. Bioengineering Department, Imperial College London, London, UK

Sadra Sadeh & Claudia Clopath

4. Sainsbury Wellcome Centre, University College London, London, UK

Andrew J. Murray

5. Institute of Experimental Medicine, Budapest, Hungary

Balázs Rózsa

6. The Kavli Institute for Brain Science, Columbia University, New York, NY, USA

Franck Polleux & Attila Losonczy

Contributions

T.G. and A.L. conceived the study and wrote the manuscript. T.G. performed all of the experiments and analysed the data. T.G. and B.V. performed immunohistochemistry and tissue clearing. T.G., S.V.R. and A.N. developed the optogenetics induction and electroporation protocols. B.R. supported AOD imaging-related software development. H.B., A.J.M. and F.P. produced viral and plasmid reagents. S.S. and C.C. developed and implemented the computational model with input from T.G.

Corresponding authors

Correspondence to [Tristan Geiller](#) or [Attila Losonczy](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Rosa Cossart and the other, anonymous, reviewers for their contribution to the peer review of this work.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Anatomical location of presynaptic neurons targeting a single CA1 pyramidal cell.

See Supplementary Table 1. **a**, Representative coronal slice of the dorsal CA1 hippocampus with the starter pyramidal cell expressing the fluorophore Venus (green), TVA receptor and glycoprotein G, after electroporation. **b**, Coronal slice of the hippocampus 14 days after rabies injection. Neurons in red expressing tdTomato are presynaptic to the starter cell. **c–e**, Presynaptic neurons can be found in the entorhinal cortex, medial septum and supramammillary nucleus (**a** to **d**, blue is DAPI) **f**, In vivo two-photon images of a starter neuron (green) and presynaptic neurons (red). **g**, Post hoc immunohistochemistry labelling of the same tissue reveals that the HA tag fused with the TVA receptor is uniquely expressed in the starter neuron, indicating that rabies tracing is restricted to this individual cell. Scale bars are 50 μ m. **h**, Lateral distribution of the presynaptic interneurons (red) and unlabelled interneurons (grey) calculated on in vivo two-photon Z-stacks ($n = 7$ mice). Coordinates (0, 0) indicate the location of the starter neuron. **i**, Same, but for depth distributions. S.O: stratum oriens, S.P: stratum pyramidale, S.R: stratum radiatum. **j**, Strategy to generate VGAT-EYFP mice in which EYFP is expressed in all inhibitory interneurons. **k**, Schematic of the experiment. A starter cell is electroporated in a VGAT-EYFP mouse, followed by injection of a RABV-tdTomato. As a result, presynaptic interneurons will co-express EYFP and tdTomato and presynaptic pyramidal cells will express only tdTomato. **l**, Representative confocal images of the starter cell (left), presynaptic and unlabelled interneurons (middle) and presynaptic pyramidal cells (right). Scale bars are 50 μ m. **m**, Quantification for 4 mice across the ipsilateral CA1.

[Source data](#)

Extended Data Fig. 2 Spontaneous place field formation is not associated with a detectable decrease in the level of presynaptic inhibition.

See Supplementary Table 2. **a**, Representative trace of the starter neuron's fluorescence activity during navigation. The first transient (pink) corresponds to the spontaneous formation of a place field, as shown in the fluorescence heat map (bottom). Fluorescence amplitude of the calcium transient during field formation is significantly higher than all other subsequent events ($n = 11$ mice, paired t-test, $P = 0.008$). **b**, Lap-average ($n = 11$ networks) activity (mean \pm s.e.m.) of the presynaptic (red) and unlabelled (grey) interneurons centred around the onset lap of field formation (starter, blue) from. **c**, Inhibition levels in both populations remained relatively constant before and after formation. All groups $n = 11$, One-way ANOVAs: starter, $P = 0.0004$ (post hoc Tukey's tests with P-values adjusted for multiple comparisons: all $P < 0.05$); presynaptic: $P = 0.32$; unlabelled $P = 0.68$ **d**, Average tuning curve (mean \pm s.e.m., all $n = 11$ networks) centred around the starter's place field for the presynaptic and unlabelled interneurons at three different time points during field formation, showing no immediate spatial reconfiguration of their responses. **e**, Same analysis using population-vector correlation before and at lap formation onset for the presynaptic interneurons. **f**, Distribution of in-field selectivity index (IFS) for presynaptic interneurons before and during the lap of field formation, showing no change in spatial selectivity at the field's location ($n = 199$ from 11 mice). **g**, Distribution of the IFS difference ($n = 199$ from 11 mice) compared to a shuffle distribution in which the location of the starter's place field is randomized on the belt.

[Source data](#)

Extended Data Fig. 3 Presynaptic interneuron spatial responses are not spatially selective when the starter is inactive and do not immediately reconfigure after spontaneous field formation.

See Supplementary Table 2. **a**, Normalized average tuning curves of the starter neurons (blue), their presynaptic partners (red) and unlabelled

interneurons (grey), centred around the middle of treadmill. Thick line represents the average for $n = 14$ mice and shaded area the s.e.m. **b**, Box plots of IFS values for all 14 mice, averaged at the network level (paired t-test, $P = 0.32$). **c**, In-field selectivity (IFS) index for all presynaptic ($n = 223$) and unlabelled ($n = 1730$) interneurons from $n = 14$ mice, $P = 0.19$ (Kolmogorov-Smirnov two sample test). Negative IFS indicates negative selectivity in the starter's place field. Insets (mean \pm s.e.m.), $P = 0.42$ (t-test). **d**, IFS values were computed in **b** and **c** for a virtual place field in the middle of the treadmill. Here, each point represents the t-test's P-values for IFS values of presynaptic vs. unlabelled interneurons while iteratively moving the location of the virtual field along the belt and recomputing the IFS at each location. This analysis shows that there is no difference in spatial selectivity anywhere on the belt when the starter cell has no place field. **e**, Experimental timeline: mice were imaged twice a day. Between each imaging session, they were allowed to rest in their home cage for one hour (also see [Methods](#)). In $n = 4$ mice, we tracked the spontaneous emergence of a place field in the starter neuron and its persistence in a later session. **f**, Representative heat map activity for a starter cell as a function of lap (y-axis) and position (x-axis) on the belt. Field creation occurred in the first session of the day at lap 4 (white arrow) and persisted after rest in a later session at the same location. **g**, Session-average tuning curve for the starter cell shown in **f** and 6 of its presynaptic interneurons, reconfiguring their response and developing anti-selectivity around the starter's place field (dashed line) in the later session. **h**, Cell-by-cell correlation coefficients between the spatial response in the first session when the field emerged (creation) and a later session (stable) for the presynaptic ($n=81$) and unlabelled ($n=267$) neurons from 4 mice, $P = 0.04$ (unpaired t-test). **i**, Same analysis but for network averages ($n = 4$ mice), $P = 0.26$ (Paired t-test pre. vs unlab). **j**, Difference between the presynaptic and unlabelled interneurons average activity centred around the starter's place field (grey), for both creation (top) and stable field session (bottom). In purple, P-values between the two distributions as a function of position on the belt. Purple shaded area indicates positions where $P < 0.05$. Notice the dip in activity in the stable session indicating the development of anti-selectivity in the presynaptic ensemble when the starter cell has an already established place field. All box plots represent median (central line) and interquartile range

(25th and 75th percentile); whiskers extend to the most-extreme data points (excluding outliers).

[Source data](#)

Extended Data Fig. 4 Photostimulation of a single pyramidal cell increases interneuron activity.

a, Left: Peri-stimulus time histogram (mean \pm s.e.m.) centred around the onset of the LED stimulations for all interneurons (green, $n = 2613$ from 6 mice) and a shuffle trace in which LED onsets were randomly shuffled in time in each session (grey, same n). Right: Quantification of increased activity (data, $P < 10^{-10}$; shuffle, $P = 0.12$, one-sample t-tests). Data vs shuffle, $P < 10^{-10}$ (paired t-test). **b**, Same analysis as **a** but all traces are averaged ($n = 14$ sessions in 6 mice, mean \pm s.e.m.) for a given session (data, $P=0.002$; shuffle, $P=0.23$, one-sample t-tests). Data vs shuffle, $P=0.003$ (paired t-test). **c**, Difference in IFS between the PRE and POST session as a function of increased $\Delta F/F$ during optogenetics stimulations ((+), $n=1208$, $P < 10^{-7}$; (-), $n=1157$, $P=0.12$; Pearson's R, $n = 6$ mice). **d**, Same as **c** but for the IFS in PRE only ((+), $n=1208$, $P=0.00012$; (-), $n=1190$, $P=0.15$; Pearson's R, $n = 6$ mice). **e**, Mice velocity (mean \pm s.e.m.) centred around LED stimulations during place field induction, separated by whether induction was successful (magenta, $n = 15$ sessions) or failed (grey, $n = 13$ sessions) from 10 mice (VGAT-Cre and VIP-Cre). Notice that mice slightly slow down during light presentation (1-1.5s stimulations) but continue running at relatively constant and high speeds. **f**, Difference in speed before and after LED stimulations from **e** for each condition. (+), $P=0.53$; (-), $P=0.85$ (one-sample t-tests). (+) vs (-), $P=0.75$ (t-test). **g**, Three-dimensional representation of all recorded interneurons ($n=1208$ from 6 mice) for successful inductions (+) plotted as a function of their distance in situ to the seed neuron (centred at $x, y, z = 0, 0, 0$). Both colour code and circle size indicate the change in IFS between PRE and POST sessions. **h**, Projection of **g** onto the Z-axis (depth) shows no distance-dependent relationship ($n=1208$ from 6 mice, $P=0.29$, Pearson's R). **i**, Projection of **g** onto the X-Y axes. **j**, Euclidean distance (X-Y) to the seed neuron as a function of change in IFS shows significant relationship ($n=1208$ from 6 mice, $P=0.012$, Pearson's R). Red bins represent the

running IFS average value along the XY distance. All box plots represent median (central line) and interquartile range (25th and 75th percentile); whiskers extend to the most-extreme data points (excluding outliers).

[Source data](#)

Extended Data Fig. 5 No immediate spatial reconfiguration of interneurons after place field induction.

a, Average spatial tuning curve for all interneurons ($n = 6$ mice) for the laps before place field induction (pre-stim laps), directly following induction (post-stim laps) and in POST following successful (magenta) or failed (grey) inductions. Interneurons are ordered by their IFS, and centred around the induced location for each condition. **b**, IFS values on a cell-by-cell basis, showing that interneurons do not become immediately negatively selective at the induced location following successful induction. Top, comparison of IFS in pre-stim laps vs. post-stim laps for successful (+) and failed (−) inductions. (−), $P = 0.81$; (+), $P = 0.06$ (Wilcoxon signed rank-tests). (−) vs (+), $P = 0.07$ (Wilcoxon rank-sum test). Bottom, comparison between post-stim laps and POST session (1 h after rest). (−), $P = 0.24$; (+), $P < 10^{-10}$ (Wilcoxon signed rank-tests). (−) vs (+): $P < 10^{-10}$ (Wilcoxon rank-sum test). For top and bottom, interneurons recorded in all three sessions: $n = 1190$ for (+) and $n = 1208$ for (−) from 6 mice. **c**, 2D histogram of interneurons' IFS in pre-stim laps and POST session (same n as **b**). (+), $P < 10^{-10}$; (−), $P < 10^{-10}$ (Pearson's R). **d**, Average IFS values at the session level ($n = 7$ for each condition from 6 mice) before, immediately after and in the POST induction session. (−), all $P > 0.05$ (paired t-tests). (+), prestim vs POST, $P = 0.04$; all others $P > 0.05$ (paired t-tests). **e**, Fraction across 6 mice of negatively selective interneurons ($IFS < 0$) before induction and in the POST session. POST(+) vs prestim(+), $P = 0.0003$; POST(+) vs prestim(−), $P = 0.0003$; POST(+) vs POST(−), $P < 10^{-5}$ (Fisher's exact tests). **f**, Difference in fraction of negatively selective interneurons (mean ± s.e.m.) between prestim and POST for each session ($n = 7$ for each condition from 6 mice). (+) vs (−), $P = 0.028$ (t-test). **g**, Overall fraction of negatively selective interneurons in prestim (top) and POST (bottom) sessions for successful (magenta) and failed (grey) inductions across 6 mice (same n as **e**), calculated as a function of position on the belt

and not only at the location where the seed neuron is induced (corresponding to position 0 here). All box plots represent median (central line) and interquartile range (25th and 75th percentile); whiskers extend to the most-extreme data points (excluding outliers).

[Source data](#)

Extended Data Fig. 6 Photostimulation of a starter neuron entrains activity in other surrounding pyramidal cells.

a, Representative field of view with one starter pyramidal cell (red) electroporated with bReaChES and GCaMP expressed in all PCs. Optogenetic stimulations (arrows) drive activity in the starter neuron and evoke calcium events in other surrounding pyramidal cells. **b**, Quantification of increased fluorescence (post minus pre) for each photostimulation of the seed neuron (left, red, $n = 31$ sessions, $P < 10^{-10}$, t-test) and all other pyramidal cells (right) in 13 mice. The presence of a seed neuron with an excitatory opsin recruits other PCs above chance level. With seed (blue), $n = 31$ sessions, $P < 10^{-5}$; without seed (black), $n = 8$ sessions, $P = 0.59$ (t-tests). With vs without seed, $P = 0.013$ (t-test). **c**, Intersomatic distance between recruited PCs and the starter neuron for successful (magenta, $n = 13$ sessions) and failed inductions (grey, $n = 18$ session), $P = 0.19$ (t-test) from 13 mice. **d**, Number of recruited pyramidal cells for each condition, $P = 0.36$ (Wilcoxon rank-sum test), same n as **c**. **e**, Fraction of recruited pyramidal cells that were place cells in the PRE session before photoinduction, minus the rate of place cells detected in the other non-recruited cells, for each session, $P = 0.28$ (Wilcoxon rank-sum test), same n as **c**. **f**, Fraction of recruited pyramidal cells that are place cells in the POST session after photoinduction, minus the rate of place cells detected in the other non-recruited cells for each session, $P = 0.005$ (Wilcoxon rank-sum test), same n as **c**. **g**, During immobility and before the seed neuron was induced, the recruited neurons are more likely to spontaneously co-fire (see [Methods](#)) than what would be expected by chance – here calculated by selecting an equivalent number of random pairs of neurons ($n = 2205$ pairs from 13 mice with neurons with at least 1 transient, mean \pm s.e.m.). **h**, Similar to **g**, pairwise correlation of activity traces averaged for each session ($n = 28$ containing bouts of immobility before induction, from 13

mice) during immobility before seed induction. Recruited, $P=0.0003$; Shuffled, $P=0.10$ (t-tests). Recruited vs shuffled, $P = 0.027$ (t-test). **i**, This like-to-like relationship among recruited cells is more pronounced for neurons the intersomatic distances of which (mean \pm s.e.m.) are within $150\mu\text{m}$ of one another ($n=2402$ pairs from 13 mice). Same assembly pairs, $P=0.0008$; Shuffled pairs, $P=0.83$ (t-tests). **j**, Pairwise distance (mean \pm s.e.m.) of place field centroids for recruited and shuffled neurons ($n = 494$ pairs from 13 mice) during navigation in laps preceding induction. Chance level is represented by a dashed line: Recruited, $P<10^{-5}$; Shuffled, $P = 0.89$ (t-tests). Recruited vs shuffled, $P<10^{-5}$ (t-test). **k**, Similar to **i**, this effect is more pronounced for closer neurons (mean \pm s.e.m.). Same assembly pairs, $P=0.048$; Shuffled pairs, $P=0.40$ (t-tests), same n as **j**. All box plots represent median (central line) and interquartile range (25th and 75th percentile); whiskers extend to the most-extreme data points (excluding outliers).

[Source data](#)

Extended Data Fig. 7 Place field induction in an individual neuron does not influence the global representation of the environment.

a, Representative examples of five sessions (from 5 distinct mice) showing the location of the place field of recruited neurons that became place cells from PRE to POST, for each condition (POST+: successful induction in the seed neuron, POST-: failed induction). Position 0 represents the location where the seed neuron was induced in PRE. **b**, Left: Heat maps representing the activity for all recruited cells as a function of position on the belt, centred around the induced location. Photoinduction (labelled ‘during stim’) drives a large increase in activity in the recruited cells, which was not present before induction (left, ‘before stim’). Right: distribution of the peaks of the spatial responses before ($n = 243$) and during ($n = 306$) photoinduction from 13 mice ($P < 10^{-10}$, two-sample Kolmogorov-Smirnov test). During, $P < 10^{-10}$; before, $P = 0.19$ (Kolmogorov-Smirnov uniformity tests). **c**, Left: Place field distribution of all the non-recruited place cells in the POST session for each condition. Right: Distribution of place field

peaks from 13 mice ($P = 0.13$, two-sample Kolmogorov-Smirnov test). (+) ($n = 1175$), $P = 0.67$; (-) ($n = 1177$), $P = 0.26$ (Kolmogorov-Smirnov uniformity tests). **d**, Left: Place field distribution of non-recruited cells which formed a field in the POST session (not place cells in PRE but place cells in POST), for each condition from 13 mice. Right: Distribution of place field peaks ($P = 0.12$, two-sample Kolmogorov-Smirnov test). (+) ($n = 856$), $P = 0.34$; (-) ($n = 904$), $P = 0.10$ (Kolmogorov-Smirnov uniformity tests).

Source data

Extended Data Fig. 8 Computational network model with single neurons and preferential connectivity cannot explain inverse selectivity in presynaptic interneurons.

a, Model with a single seed pyramidal cell. For all following analyses, the structure and parameters of the network is similar to Fig. 4 with the same number of seed neurons ($n = 40$). Specifically, the seed neuron has both random and specific connectivity with interneurons, with the same N_s (number of units within the subnetwork). **b**, Right: average activity of interneurons from the subnetwork of the starter cell (subnet.) and from the rest of the network (rand.). Right: in-field selectivity (IFS, mean \pm s.e.m.) for interneurons presynaptic ($n=2322$) to the starter cell (presyn., $n = 2322$) and others (rand., $n = 1696$). **c**, Same as **b** (mean \pm s.e.m.), when there is no depression between the starter cell and interneurons ($d=0$; $n = 2301$ presyn.; $n = 1699$ rand.). **d**, Same as **b** (mean \pm s.e.m.), for stronger depression rate of synapses ($d=50$; $n = 2283$ presyn.; $n = 1717$ rand.). **e–h**, Simulation of the network model with different sizes of the pyramidal cell-interneuron subnetwork (N_s). Other parameters are the same as in Extended Data Fig. 9, which is copied here in **f** for comparison. **e**, IFS values (mean \pm s.e.m.) for 10 pyramidal cells and 10 interneurons ($N_s = 10$; $n = 4436$ presyn.; $n = 3564$ rand.). **f**, IFS values (mean \pm s.e.m.) for $N_s = 15$ ($n = 4611$ presyn.; $n = 3389$ rand.). **g**, IFS values (mean \pm s.e.m.) for $N_s = 20$ ($n = 4843$ presyn.; $n = 3157$ rand.). **h**, IFS values (mean \pm s.e.m.) for $N_s = 25$ ($n = 5064$ presyn.; $n = 2936$ rand.). The results are robust to change of the parameter, especially larger subnetworks lead to more prominent presence of the anti-tuning in presynaptic inhibition. Smaller subnetworks make the detection of

anti-tuning difficult, although the effect is still observable in the average activity.

[Source data](#)

Extended Data Fig. 9 Computational model with subnetwork structure with different connectivity motifs.

a, Spatial tuning of all pyramidal cells (left) and interneurons (right) in the networks from 40 simulations (similar in the following **b–e**), sorted according to their in-field selectivity (IFS). Position is expressed relative to the location of place formation in the starter cells, respectively. **b**, Left: Average activity of interneurons within the subnetwork (subnet.) and from outside (rand.) as a function of position. Right: IFS (mean \pm s.e.m.) for interneurons presynaptic to starter cells (presyn., n = 2335) and others (rand., n = 1789). The results are shown for the full model (Fig. 4) with pyramidal cell-interneuron subnetwork structure (illustrated on the top). **c**, Same as **b** (mean \pm s.e.m.) for network structures with random connectivity and without the specific connectivity structure of the starter-cell-interneuron subnetwork (n = 1964 presyn.; n = 2043 rand.). **d**, Same as **b** (mean \pm s.e.m.) without the specific connectivity of starter-PCs, while starter-interneurons preserve their specific connectivity (n = 2339 presyn.; n = 1669 rand.). **e**, Schematic illustration of the reorganization of activity and network interactions following field formation. The starter cell elevates the activity of pyramidal cells and interneurons within the subnetwork at its selective location (left), which is followed by depression of pyramidal cells-to-interneurons connections, leading to the diminished activity of interneurons within the subnetwork at that location (right).

[Source data](#)

Extended Data Fig. 10 Alternative model with direct disinhibitory circuitry.

a, Top: schematic of the circuit before field formation. A starter pyramidal cell (PC) contacts two interneuron entities (INT1 and INT2) with excitatory connections. INT1 (interneuron-selective interneuron such as VIP) exerts

static inhibition onto INT2, which projects back to PC. Bottom: in this model, formation of a field in the starter PC drives INT1s and INT2s, but a stronger connectivity with INT1 leads to the depression of INT2 responses.

b, Evolution of neuronal activity of the starter PC (left), INT1 (middle) and INT2 (right) following place field formation of the PC on lap 1. **c**, Average tuning curves before field formation (initial), during the formation (middle) and after field has formed (final), showing that INT2 ultimately exhibits negative tuning at that field location. **d**, Evolution of the synaptic weights as a function of time (laps) during the process of field formation. This model has experimentally testable predictions that we performed. **e**, To do so, we performed calcium imaging in VIP-Cre mice, known to genetically label a subset of interneuron-specific interneurons (INT1) and single-cell electroporation in an individual PC (seed) to perform place field induction. Left: schematic of the experiment. Right: In vivo two-photon image of GCaMP-expressing VIP interneurons (green) and a single CA1 PC expressing GCaMP and bReaChES (red). Scale bar is 50 μ m. **f**, PSTH (mean \pm s.e.m.) centred at the onset of the LED photostimulation for all VIP interneurons and a shuffle trace in which LED onset was randomly chosen during the imaging session ($n = 6$ sessions in 4 mice). **g**, Box plots representing the increased activity following LED stimulation. Data, $P=0.18$; Shuff., $P=0.30$ (t-tests). Data vs Shuff, $P=0.8$ (t-test). The lack of increased activity during photostimulation goes against the prediction of our model that field formation should elevate responses in the INT1 population. **h**, Distribution of in-field selectivity (IFS) at the induced location for all VIP interneurons before photoinduction (PRE, $n = 774$), and after successful (POST(+), magenta, $n = 439$) and failed (POST(-), grey, $n = 353$) inductions. Data from $n = 14$ sessions in 4 mice. All $P > 0.05$ (unpaired t-tests). The lack of development of positive selectivity is not consistent with our model (see **c**). **i**, Average spatial tuning curve for all interneurons for the laps before place field induction (PRE), and in the POST session following successful (magenta) or failed (grey) inductions. Interneurons are ordered by their IFS, and centred around the induced location for each condition. **j**, Box plots representing IFS values for all VIP-positive interneurons (same n as **h**). PRE vs POST(-), $P = 0.43$; PRE vs POST(+), $P = 0.37$ (t-tests). All box plots represent median (central line) and interquartile range (25th and 75th percentile); whiskers extend to the most-extreme data points (excluding outliers).

[Source data](#)

Supplementary information

Supplementary Tables

This file contains Supplementary Tables 1–4.

Reporting Summary

Source data

[Source Data Fig. 1](#)

[Source Data Fig. 2](#)

[Source Data Fig. 3](#)

[Source Data Fig. 4](#)

[Source Data Extended Data Fig. 1](#)

[Source Data Extended Data Fig. 2](#)

[Source Data Extended Data Fig. 3](#)

[Source Data Extended Data Fig. 4](#)

[Source Data Extended Data Fig. 5](#)

[Source Data Extended Data Fig. 6](#)

[Source Data Extended Data Fig. 7](#)

[**Source Data Extended Data Fig. 8**](#)

[**Source Data Extended Data Fig. 9**](#)

[**Source Data Extended Data Fig. 10**](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Geiller, T., Sadeh, S., Rolotti, S.V. *et al.* Local circuit amplification of spatial selectivity in the hippocampus. *Nature* **601**, 105–109 (2022). <https://doi.org/10.1038/s41586-021-04169-9>

- Received: 10 May 2021
- Accepted: 15 October 2021
- Published: 01 December 2021
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04169-9>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04169-9>

| [Section menu](#) | [Main menu](#) |

- Article
- Open Access
- [Published: 10 November 2021](#)

Pre-existing polymerase-specific T cells expand in abortive seronegative SARS-CoV-2

- [Leo Swadling](#) [ORCID: orcid.org/0000-0002-0537-6715](#)¹,
- [Mariana O. Diniz](#)¹ [na1](#),
- [Nathalie M. Schmidt](#) [ORCID: orcid.org/0000-0002-9841-8418](#)¹ [na1](#),
- [Oliver E. Amin](#)¹ [na1](#),
- [Aneesh Chandran](#)¹ [na1](#),
- [Emily Shaw](#)¹ [na1](#),
- [Corinna Pade](#)²,
- [Joseph M. Gibbons](#) [ORCID: orcid.org/0000-0002-7238-2381](#)²,
- [Nina Le Bert](#) [ORCID: orcid.org/0000-0003-0502-2527](#)³,
- [Anthony T. Tan](#)³,
- [Anna Jeffery-Smith](#)^{1,2},
- [Cedric C. S. Tan](#)⁴,
- [Christine Y. L. Tham](#) [ORCID: orcid.org/0000-0002-2913-7591](#)³,
- [Stephanie Kucykowicz](#) [ORCID: orcid.org/0000-0002-8849-218X](#)¹,
- [Gloryanne Aidoo-Micah](#)¹,
- [Joshua Rosenheim](#) [ORCID: orcid.org/0000-0003-0171-2053](#)¹,
- [Jessica Davies](#)¹,
- [Marina Johnson](#)⁵,
- [Melanie P. Jensen](#)^{6,7},
- [George Joy](#)^{6,8},
- [Laura E. McCoy](#) [ORCID: orcid.org/0000-0001-9503-7946](#)¹,
- [Ana M. Valdes](#) [ORCID: orcid.org/0000-0003-1141-4471](#)^{9,10},
- [Benjamin M. Chain](#) [ORCID: orcid.org/0000-0002-7417-3970](#)¹,
- [David Goldblatt](#) [ORCID: orcid.org/0000-0002-0769-5242](#)⁵,
- [Daniel M. Altmann](#)¹¹,
- [Rosemary J. Boyton](#) [ORCID: orcid.org/0000-0002-5608-0797](#)^{12,13},

- [Charlotte Manisty](#)^{6,8},
- [Thomas A. Treibel](#)^{6,8},
- [James C. Moon](#)^{6,8},
- [COVIDsortium Investigators](#),
- [Lucy van Dorp](#) ORCID: orcid.org/0000-0002-6211-2310⁴,
- [Francois Balloux](#) ORCID: orcid.org/0000-0003-1978-7715⁴,
- [Áine McKnight](#)²,
- [Mahdad Noursadeghi](#) ORCID: orcid.org/0000-0002-4774-0853¹,
- [Antonio Bertoletti](#) ORCID: orcid.org/0000-0002-2942-0485^{3,14} &
- [Mala K. Maini](#) ORCID: orcid.org/0000-0001-6384-1462¹

Nature volume 601, pages 110–117 (2022)

- 123k Accesses
- 4 Citations
- 4507 Altmetric
- [Metrics details](#)

Subjects

- [Immunological memory](#)
- [Infection](#)
- [SARS-CoV-2](#)
- [T cells](#)

Abstract

Individuals with potential exposure to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) do not necessarily develop PCR or antibody positivity, suggesting that some individuals may clear subclinical infection before seroconversion. T cells can contribute to the rapid clearance of SARS-CoV-2 and other coronavirus infections^{1,2,3}. Here we hypothesize that pre-existing memory T cell responses, with cross-protective potential against SARS-CoV-2 (refs. [4,5,6,7,8,9,10,11](#)), would expand in vivo to support rapid viral control, aborting infection. We measured SARS-CoV-2-reactive T cells, including those against the early transcribed replication-transcription complex (RTC)^{12,13}, in intensively monitored healthcare workers (HCWs) who tested repeatedly negative according to PCR, antibody binding and neutralization assays (seronegative HCWs (SN-HCWs)). SN-HCWs had stronger, more multispecific

memory T cells compared with a cohort of unexposed individuals from before the pandemic (prepandemic cohort), and these cells were more frequently directed against the RTC than the structural-protein-dominated responses observed after detectable infection (matched concurrent cohort). SN-HCWs with the strongest RTC-specific T cells had an increase in *IFI27*, a robust early innate signature of SARS-CoV-2 (ref. [14](#)), suggesting abortive infection. RNA polymerase within RTC was the largest region of high sequence conservation across human seasonal coronaviruses (HCoV) and SARS-CoV-2 clades. RNA polymerase was preferentially targeted (among the regions tested) by T cells from prepandemic cohorts and SN-HCWs. RTC-epitope-specific T cells that cross-recognized HCoV variants were identified in SN-HCWs. Enriched pre-existing RNA-polymerase-specific T cells expanded in vivo to preferentially accumulate in the memory response after putative abortive compared to overt SARS-CoV-2 infection. Our data highlight RTC-specific T cells as targets for vaccines against endemic and emerging *Coronaviridae*.

[Download PDF](#)

Main

There is wide variability in the outcome of exposure to SARS-CoV-2, ranging from severe illness to asymptomatic infection, to those individuals who remain negative according to standard diagnostic tests. Recent studies have identified SARS-CoV-2 T cell reactivity in prepandemic samples [5,6,7,8,9,10,11,15,16,17,18](#) and isolated cases of exposed individuals who have not seroconverted with single-time-point screening [4,16,19,20,21,22](#). We studied an intensively monitored cohort of HCWs with potential exposure during the first UK pandemic wave (23 March 2020), comparing those with or without PCR and/or antibody evidence of SARS-CoV-2 infection. We postulated that, in HCWs for whom PCR and the most sensitive binding and neutralizing antibody tests remained repeatedly negative (SN-HCWs), T cell assays might distinguish a subset of SN-HCWs with a subclinical, rapidly terminated (abortive) infection. We hypothesized that these individuals would exhibit pre-existing memory T cells with cross-reactive potential, obviating the time required for de novo T cell priming and clonal expansion. In SN-HCWs, and in an additionally recruited cohort of medical students and laboratory staff with stored prepandemic samples that remained seronegative after close contact with cases, we had the opportunity to compare SARS-CoV-2-specific memory T cells with those that were already present in the same individual before, or at the time of, potential exposure.

We included an analysis of the understudied T cells directed against the core RTC within open reading frame 1ab (ORF1ab) (RNA polymerase co-factor non-structural protein 7 (NSP7), RNA polymerase NSP12 and helicase NSP13, hereafter the RTC); these are putative targets for pre-existing responses with pan-*Coronaviridae* reactivity,

because they are likely to be highly conserved due to their key early roles in the viral life cycle. Consistent with this, in cases in which immunity against other viruses (including hepatitis B virus (HBV), hepatitis C virus (HCV), HIV and Japanese encephalitis virus (JEV)) has been described in exposed seronegative individuals, T cells were more likely to target non-structural proteins, such as polymerase, compared with in individuals with a seropositive infection^{[23](#),[24](#),[25](#),[26](#),[27](#)}.

SARS-CoV-2 T cells in seronegative HCWs

We compared T cell reactivity in intensively monitored HCWs with a laboratory-confirmed infection or SN-HCWs, matched for exposure risk and demographic factors (COVIDsortium; Fig. [1a](#) and Extended Data Table [1](#)). Additional control cohorts included healthy adults who were sampled in London, UK, or Singapore before SARS-CoV-2 circulation in humans (prepandemic cohort; Fig. [1a](#)). SN-HCWs were defined by negative weekly diagnostic tests (baseline–week 16, SARS-CoV-2 PCR, nasopharyngeal swab; anti-spike-1 IgG and anti-nucleoprotein (NP) IgG/IgM seroassays^{[28](#)}; Fig. [1b–d](#)). Having previously reported a range of neutralizing antibody titres at week 16 in laboratory-confirmed infections, we examined neutralizing antibodies in SN-HCWs. Two HCWs with neutralizing antibody titres that were just above the threshold were excluded from further analyses; the remaining SN-HCWs were negative by pseudotype assay (Fig. [1e](#)), with a subset also confirmed to be negative at three time points for authentic virus neutralization (Extended Data Fig. [1a](#)). SN-HCWs could have become PCR negative by recruitment; however, non-seroconverters after PCR positivity were rare (2.6% of PCR-positive HCWs negative by all three seroassays^{[16](#)}) and antibody responses are unlikely to have waned before recruitment^{[28](#)}. Furthermore, SN-HCWs lacked detectable SARS-CoV-2 spike-specific memory B cells, which we have shown persist after waning of neutralizing antibodies^{[29](#)} (Extended Data Fig. [1b](#); below the detection threshold). Thus, the SN-HCWs represented a cohort of intensely monitored HCWs who resisted classical laboratory-confirmed infection.

Fig. 1: SARS-CoV-2-specific T cells in SN-HCWs.

 **figure1**

a, Design of the HCW and prepandemic cohorts. nAb, neutralizing antibodies. **b**, Cycle threshold values for the *E* gene PCR analysis in SN-HCWs and HCWs with a laboratory (lab)-confirmed infection (undetectable at 40 cycles was assigned 41). **c**, **d**, Anti-spike S1 (**c**) and anti-NP antibody (**d**) titres in SN-HCWs (baseline to week 16; $n = 58$; dotted lines at assay positivity cut-off and at average peak (AvPos) response in laboratory-confirmed infection). **e**, Pseudovirus neutralization at week 16. The crossed circles represent individuals who were excluded from SN-HCW group ($IC_{50} > 50$). **f**, SARS-CoV-2 proteome highlighting RTC and structural regions assayed for T cell responses (peptide subpools are identified by the numbered boxes) and the number of overlapping 15-mer peptides (or mapped epitope peptides (MEP) for spike). **g–j**, IFN γ ELISpot analyses. **g**, **h**, Viral proteins recognized by individuals coloured by specificity (**g**) and the number of viral proteins targeted by group (**h**). **i**, **j**, The magnitude of the T cell response coloured by viral protein (**i**) and the cumulative magnitude of the T cell response by group (**j**). The red bar shows the geometric mean. For **e**, **h**, the red bar shows the median. For **h**, **j**, statistical analysis was performed

using Kruskal–Wallis tests with Dunn’s correction. M, membrane; SFCs, spot-forming cells. For **b–e**, **g–j**, participants were from the COVIDsortium HCW cohort.

Source Data.

We quantified SARS-CoV-2-specific memory T cells by ELISpot using unbiased stimulation with overlapping peptides covering structural proteins and the less-well-studied non-structural proteins of the RTC (Fig. [1f](#)). As previously described, when using sensitive assays^{[5,6,7,9,17,18](#)} (such as 400,000 peripheral blood mononuclear cells (PBMCs) per well IFN γ ELISpot analysis used here^{[8,16](#)}), some SARS-CoV-2-reactive T cells were detectable in prepandemic samples; however, their multispecificity was significantly lower compared to the week 16 group with a laboratory-confirmed infection (Fig. [1g, h](#); structural responses at week 16 previously reported^{[16](#)}). By contrast, SN-HCWs had SARS-CoV-2-specific T cells that were comparable in breadth to infected HCWs at week 16 and significantly more multispecific than prepandemic samples (Fig. [1g, h](#)). The SARS-CoV-2-specific T cells of SN-HCWs targeted more protein pools and had an approximately fivefold higher cumulative magnitude of responses compared with those of the prepandemic cohort, with an overall strength equivalent to the infected cohort at week 16 (Fig. [1i, j](#)).

T cells from prepandemic samples tended to not target both halves of NP (NP1 and NP2 subpools), whereas around 50% of SN-HCWs and HCWs with a laboratory-confirmed infection did, confirming our earlier suggestion^{[8](#)} that this serves as a simple proxy measure of a multispecific response (Extended Data Fig. [1c–e](#)). Taken together, we found a higher magnitude and breadth of SARS-CoV-2-specific T cells in HCWs who repeatedly tested PCR and antibody negative compared with individuals in the prepandemic cohort.

RTC-specific T cells and *IFI27* in SN-HCWs

We next investigated whether T cell memory differs in SN-HCWs versus HCWs with laboratory-confirmed infection. Anti-viral T cells recognizing influenza A, Epstein–Barr virus (EBV) and cytomegalovirus (CMV) (together, FEC) were equivalent between the three cohorts (Extended Data Fig. [2a](#)). However, the relative immunodominance of T cells against SARS-CoV-2 structural versus RTC proteins differed between the groups. The laboratory-confirmed-infection group had more responses to structural proteins (spike, membrane, NP and ORF3a) than to RTC (NSP7, NSP12, NSP13) (Fig. [2a, b](#)). Memory T cells against structural proteins tended to positively correlate with viral load, whereas RTC responses did not show this association (Extended Data Fig. [2b](#)). By contrast, T cells of the SN-HCWs targeted both structural and RTC regions, with significantly more RTC-specific T cells compared with either the infected or prepandemic groups (Fig. [2a](#) and Extended Data

Fig. 2c, d). Prepandemic samples had a ratio of RTC to structural responses that did not differ significantly from that in SN-HCWs (Fig. 2b), pointing to a possible influence of pre-existing responses on the pool of T cells expanding in SN-HCWs. A further small group (10%) of HCWs had PCR-confirmed infection but lacked detectable neutralizing antibodies at week 16, some of the individuals in this group also lacked binding antibodies; this subgroup was similarly enriched for RTC-reactive T cells (Extended Data Fig. 2e, f). Taken together, this suggests that the structural proteins, which are abundantly produced during active infection, are dominant T cell targets after mild infection, whereas T cells in SN-HCWs preferentially focus on the RTC.

Fig. 2: RTC-specific T cell and *IFI27* signature in SN-HCWs.

 figure2

a, b, IFN γ ELISpot analysis at week 16. **a**, The magnitude of T cell response to structural regions and the RTC. **b**, The ratio of the T cell response to the RTC versus structural regions. The percentage of the cohort with a ratio above 1 (RTC > structural) is shown below. For **a, b**, the red bar shows the geometric mean. **c**, *IFI27* transcript signal by reverse transcription PCR (RT-PCR) in unexposed prepandemic samples ($n = 59$), baseline (BL) samples in HCWs who remained PCR negative and seronegative throughout follow-up ($n = 99$), SN-HCWs with weak ($n = 5$, <50 SFCs per 10^6 PBMCs; Extended Data Fig. 4a) or strong ($n = 15$, >50 SFCs per 10^6 PBMCs) RTC-specific T cells (baseline and peak signal (weeks 0–5)), and HCWs at the time of PCR positivity (PCR $^+$). **d**, The longitudinal *IFI27* signal in SN-HCWs with weak or strong RTC-specific T cell responses (n values as in **c**). For **c, d**, the red bar shows the median, with 2 s.d. either side of the prepandemic cohort mean highlighted in grey; the percentage with raised *IFI27* above the mean + 2 s.d. is indicated below. Statistical analysis was performed using Kruskal–Wallis analysis of variance (ANOVA) with Dunn’s correction (**a–d**). Mann–Whitney paired *t*-test for paired BL versus peak (**c**). For **a–d**, participants were from the COVIDsortium HCW cohort.

Source Data

To confirm the T cell identity of ELISpot responses in SN-HCWs at week 16, we expanded them with RTC peptides and detected both CD4 $^+$ and CD8 $^+$ SARS-CoV-2-

specific T cells dividing (CellTrace violet (CTV) dilution) and producing IFN γ (Extended Data Fig. 3a and Extended Data Table 2). Their post-expansion frequencies tended to be lower than control influenza A/EBV/CMV-specific responses in the same donors but proportional to their differing ex vivo frequencies, indicating comparable proliferative potential (Extended Data Fig. 3b). In vitro-expanded RTC-specific T cells in SN-HCWs were also highly functional, producing multiple cytokines in tandem (Extended Data Fig. 3c,d). Most of the SARS-CoV-2-specific T cells expanded from SN-HCWs were CD4 $^{+}$; however, CD8 $^{+}$ T cells were also detectable in the majority of individuals (Extended Data Fig. 3e).

Our data raised the possibility that SARS-CoV-2 infection in HCWs represents a spectrum, with some SN-HCWs expanding T cells as a result of a subclinical infection that was not detectable by PCR or antibody seroconversion. To test this postulate, we measured the interferon-inducible transcript *IFI27* in the blood, which has recently been shown to detect SARS-CoV-2 infection at, or one week before, PCR positivity (specificity of 0.95 and sensitivity of 0.84)¹⁴. Of the 25% of SN-HCWs with the strongest post-exposure RTC-specific T cell responses (Extended Data Fig. 4a), 40% (that is, 10% of SN-HCW group) already had *IFI27* levels at recruitment that were above the threshold set on the basis of a cohort of unexposed prepandemic samples, although their levels tended to be lower than those in individuals with a laboratory-confirmed infection (Fig. 2c). To further estimate the frequency of abortive infections we tested a larger cohort of 99 unselected SN-HCW baseline samples, and found that a comparable proportion (9.1%) had *IFI27* induction above the prepandemic threshold (Fig. 2c). The *IFI27* signal peaked above the prepandemic threshold in 93.3% of those with strong RTC-specific T cells over weeks 0–5, but in none with weak or undetectable RTC-specific T cells (Fig. 2c). *IFI27* levels showed a cumulative increase, peaking at 3–5 weeks after the UK lockdown (23 March 2020) (Fig. 2d; by which time all of the first-wave laboratory-confirmed infections had occurred (Fig. 1b)). By contrast, *IFI27* was unchanged over weeks 0–5 in SN-HCWs with weak or absent RTC-specific responses, resulting in a lower *IFI27* slope and variance (Fig. 2d and Extended Data Fig. 4b,c). The peak *IFI27* level correlated with NSP7 T cells at week 16, with the latter correlating more strongly with NSP12 and other RTC-specific responses compared with structural responses. Neither *IFI27* or T cell specificity correlated with age, sex or other demographic factors, such as exposure type, in this small cohort (Extended Data Fig. 4d and Extended Data Table 1).

In summary, during a period of high transmission at the start of the first UK pandemic wave, a low-level systemic interferon response indicative of virus exposure was detectable selectively in individuals who had the strongest SARS-CoV-2-specific T cells after exposure, despite them lacking PCR or antibody confirmation of infection. Extrapolating from previous data showing that *IFI27* is induced at the time of incident

infection and correlates with viral load¹⁴, this is consistent with a low-level infection among SN-HCWs with stronger RTC-specific T cell responses.

Targeting of conserved RNA polymerase

A transient/abortive infection that is not detectable by PCR or seroconversion could conceivably result from a lower viral inoculum and/or from a more efficient innate and/or adaptive immune response. The latter would be favoured by pre-existing memory T cells with the potential to expand rapidly after cross-recognition of early viral products of SARS-CoV-2 replication. Early T cell proliferation and T-cell-receptor clonal expansion, even before the virus is detectable, has been observed during mild SARS-CoV-2 infection^{17,30} and expansion of virus-specific T cells predates antibody induction after mRNA vaccination^{2,31}. Having found that the SN-HCW group is enriched for SARS-CoV-2-specific T cells, particularly against RTC, we investigated the possibility that some of these represented expansions of pre-existing cross-reactive responses.

Probable candidates for the source of pre-existing T cells that cross-recognize SARS-CoV-2 are previous infections with closely related human endemic common cold coronaviruses (α -HCoV 229E, NL63 and β -HCoV HKU1, OC43). We bioinformatically determined the sequence homology of all possible SARS-CoV-2-derived 15-mer peptides to a curated set of HCoV sequences (Supplementary Table 1). RTC proteins, which are expressed at the first stage of the SARS-CoV-2 life cycle¹³, had 15-mer sequences of high homology to HCoV^{32,33} (Fig. 3a). In particular, NSP7-, NSP12- and NSP13-derived 15-mers had 6.3%, 29.9% and 31.0% higher average sequence homology to the four HCoV species, respectively, compared with structural-protein-derived 15-mers (all $P < 0.001$; Fig. 3b). NSP12, which was the largest of these proteins, represented the region with the most homology overall among human-infecting *Coronaviridae*. We further assessed the diversity across global circulating SARS-CoV-2 sequences (13,785, representative subsample of 611,893 sequences, GISAID, 27 July 2021; Extended Data Fig. 5a) using Nei's genetic diversity index and an estimate of the minimal number of independent mutational events (homoplasies) at any nucleotide. By both metrics, the RTC proteins NSP12 and NSP13 were among the most conserved across SARS-CoV-2 clades (Fig. 3c and Extended Data Fig. 5b, d) and were significantly more conserved than many structural proteins (Extended Data Table 3).

Fig. 3: Cross-reactive T cells targeting conserved RNA polymerase.

 **figure3**

a, Sequence homology of SARS-CoV-2-derived peptide sequences to HCoV sequences. The columns show 15-mer SARS-CoV-2-derived peptides. The rows show HCoV genome records. Cells are coloured by the level of homology of the 15-mer to a particular HCoV proteome. Cells with no fill indicate that a sequence homology of <40% was observed. **b**, The average sequence homology of 15-mers covering SARS-CoV-2 proteins, or regions (pink, structural (S, M, NP and ORF3a); black, RTC (NSP7, NSP12 and NSP13)), to HCoV sequences. Viral proteins that were not assayed for T cell responses are shown in grey. **c**, The nucleotide diversity along the SARS-CoV-2 genome estimated with Nei's genetic diversity index across each viral protein for all SARS-CoV-2 clades (subsampling; Extended Data Fig. 5a). **d, e**, IFN γ ELISpot analysis of the magnitude of T cell responses to individual SARS-CoV-2 proteins in unexposed prepandemic samples (**d**) and SN-HCWs at week 16 (**e**). The frequency of responders is shown as doughnut charts above. The bar shows the geometric mean. ND, not done. Statistical analysis was performed using Kruskal–Wallis tests with Dunn's correction. Participants were from the COVIDsortium HCW cohort.

[Source Data](#).

Importantly, the highly conserved RNA polymerase (NSP12) was also the region among those tested in prepandemic samples that was most commonly targeted by T cells, with the highest average magnitude and frequency of responders (Fig. 3d). Notably, the same preferential targeting of NSP12 was observed in a geographically distinct cohort of prepandemic samples from Singapore (Fig. 3d). Pre-existing T cells had the potential to recognize all of the viral antigens tested, including those with less conservation across HCoV, as previously described^{5,7,17,34}. Responses against these regions were further enriched in SN-HCWs (Fig. 3d, e; Mann–Whitney U-test, $P < 0.0001$ for all except for ORF3a ($P = 0.0006$) and NSP13 ($P = 0.0003$)), suggesting many sources of pre-existing and de novo responses contribute to T cell memory in exposed seronegative individuals. Despite potential demographic confounding factors between cohorts (Extended Data Table 1), as with prepandemic samples, T cells of SN-HCWs preferentially targeted NSP12 (Fig. 3e). Thus, the viral protein that is most commonly targeted by pre-existing T cells is also the largest conserved region between *Coronaviridae*, suggesting exposure to HCoV is one probable source of cross-reactive T cells.

To further examine the potential for cross-reactivity due to previous infection with seasonal HCoV, we mapped new and previously described^{6,8,18,35} RTC-specific CD4⁺ and CD8⁺ T cell epitopes in SN-HCWs, revealing high sequence conservation with HCoV (Extended Data Table 4 and Extended Data Fig. 6a, b). We identified cross-reactivity against the HLA-A*02:01 restricted epitope in NSP7. A subset of T cells co-stained with MHC class I pentamers loaded with SARS-CoV-2 and HKU1 sequence peptide ex vivo, and bound to SARS-CoV-2 peptide-loaded pentamer after expansion for 10 days with either peptide (Extended Data Fig. 6c). T cells from 3 out of 5 HLA-A*02:01⁺ SN-HCWs tested had stronger responses to the HKU1 sequence than to other seasonal HCoV or SARS-CoV-2 (Extended Data Fig. 6d, e). This suggested that previous HKU1 infection primed these NSP7 responses that are able to cross-recognize the SARS-CoV-2 sequence, albeit with reduced efficiency. HLA-B*35⁺ SN-HCWs also showed variable cross-recognition of seasonal HCoV variant sequences of an NSP12 epitope (Extended Data Fig. 6f).

An alternative explanation for expanded T cells with cross-reactive potential in SN-HCWs is an infection with a seasonal coronavirus during the first wave of SARS-CoV-2 infections in London. As expected, all HCWs had detectable anti-spike IgG against the four endemic HCoV and, as previously described³⁶, spike-specific antibodies against betacoronavirus OC43 were increased in those with PCR-detectable infection and SARS-CoV-2-specific seroconversion (Extended Data Fig. 7). However, there was no difference in endemic HCoV titres in HCWs who had strong RTC-specific T cells and raised IFI27 compared with those with weak or absent RTC-specific responses (Extended Data Fig. 7), making it improbable that HCoV infection itself accounted for the SARS-CoV-2-reactive T cells that we detected in SN-HCWs.

In summary, RTC regions such as polymerase that are expressed in the first stage of the viral life cycle are highly conserved among HCoV and are preferentially targeted by T cells in prepandemic and SN-HCW samples. A subset of T cells from donors who were able to abort infection could cross-recognize SARS-CoV-2 and HCoV sequences at individual RTC epitopes, pointing to previous infection with HCoV as one source of pre-existing cross-reactive T cells.

Polymerase-specific T cells in abortive infection

To examine whether pre-existing cross-reactive and/or rapidly generated de novo RTC-specific T cells expand in vivo, we obtained paired PBMC samples before and after SARS-CoV-2 exposure. Medical students and laboratory staff (contact cohort, $n = 23$) who were sampled before the coronavirus disease 2019 (COVID-19) pandemic (winter 2018–2019), were resampled after close contact with individuals with SARS-CoV-2 infection, with or without IgG seroconversion and with or without PCR positivity (contact cohort; Extended Data Table 5). Parallel analysis of pre- and post-exposure/infection PBMCs demonstrated expansion of the RTC over structural responses in the close-contact seronegative group (Fig. 4a). By contrast, the group with serological confirmation of infection showed the expected in vivo expansion of pre-existing structural SARS-CoV-2-reactive T cells, with no significant increase in RTC-specific T cells (Fig. 4a and Extended Data Fig. 8a). We observed in vivo expansion of pre-existing NSP12 responses in 4 out of 5 individuals who remained seronegative after exposure to SARS-CoV-2, resulting in a significant increase in NSP12 but not control FEC responses (Extended Data Fig. 8b,c). Four out of five remaining seronegative close contacts had newly detected, presumed de novo, low-level responses after exposure (Extended Data Fig. 8c).

Fig. 4: In vivo expansion of polymerase-specific T cells in abortive infection.

 **figure4**

a–e, IFN γ ELISpot analysis. **a**, The magnitude of the T cell response in seronegative individuals who had close contact with cases (green) or in seropositive individuals with infection (orange) to the RTC, structural proteins (Str), a summed total, and an influenza A, EBV and CMV (FEC) peptide pool (grey seronegative/seropositive combined), before and after exposure/infection. Data are mean \pm s.e.m. *P* values are shown at the top. **b**, The change in magnitude of NSP12 T cell response between

recruitment and post-exposure in SN-HCWs (subgroup with top 19 RTC responses at week 16; Extended Data Fig. 4a). Expanded, greater than twofold change. **c**, The magnitude of paired pre- and post-exposure T cell responses to individual 9–15-mer peptides (individual responses; Extended Data Fig. 8g) from RTC or the control FEC peptide pool in SN-HCWs (weeks 16–26, 11 responses from 9 SN-HCWs). CI, confidence intervals. **d**, The magnitude of the T cell response to individual SARS-CoV-2 proteins (top) and to subpools (~40 overlapping peptides; bottom) within the RTC at week 16 in HCWs with a laboratory-confirmed infection or SN-HCWs. **e**, Pre-existing NSP12-specific T cell responses in baseline samples from SN-HCWs and the laboratory-confirmed infection group (PCR positive after baseline or seroconversion at least 4 weeks after recruitment). The doughnut plot above shows frequency. For c–e, the red lines (c, e) and bars (d) show the geometric mean. Statistical analysis was performed using Wilcoxon tests (a, c), Mann–Whitney *U*-test and Fisher’s exact test (d, e). For a, participants were from the contact cohort (Extended Data Table 5). For b–e, participants were from the COVIDsortium HCW cohort (Extended Data Table 1).

Source Data.

We next reverted to the SN-HCW group, in which small volume PBMC collections were available from the time of recruitment, enabling the targeted analysis of baseline T cells in those with the strongest RTC responses at week 16. NSP12-specific T cells were already detectable at the baseline in 79% of those SN-HCWs with the strongest NSP12 responses after exposure (Fig. 4b). NSP12 responses expanded in vivo on average 8.4-fold between recruitment and week 16, with no corresponding change in FEC responses (Fig. 4c and Extended Data Fig. 8d). We confirmed the expansion at week 16 of pre-existing RTC-specific T cells at the subpool (Extended Data Fig. 8e,f) and individual peptide (Fig. 4c and Extended Data Fig. 8g) levels. Moreover, many T cells were newly detected after exposure (Extended Data Fig. 8g), reflecting either de novo priming or expansion of responses that were previously below the limit of assay detection (example of expanded response undetectable by ex vivo ELISpot; Extended Data Fig. 8h). All of the HCWs with newly detected or expanding/contracting NSP12-specific T cells had both NP1- and NP2-reactive T cells after exposure (Extended Data Fig. 8i), whereas only 2 out of 5 individuals with no change in NSP12 had these specificities, suggesting that they may not have had the same level of SARS-CoV-2 exposure. The fold change in NSP12 between recruitment and the week 16 follow-up correlated with the total SARS-CoV-2 response, supporting its use to identify those seronegative individuals with expanded T cell immunity after exposure (Extended Data Fig. 8j).

Finally, we examined whether there was a preferential enrichment of RTC-specific responses in SN-HCWs compared with HCWs with a laboratory-confirmed infection at week 16. Notably, the RNA polymerase NSP12 and its cofactor NSP7 were the only proteins that induced higher-magnitude T cell responses in seronegative individuals in

whom detectable infection was not established compared with those with overt infection (Fig. 4d). T cells in SN-HCWs targeted a larger number of regions of NSP12 (subpools of about 40 overlapping 15-mers; Fig. 1g) compared with T cells in the prepandemic or seropositive cohorts (Extended Data Fig. 8k). T cells targeting several regions of NSP12 and other RTC pools were enriched in SN-HCWs compared to HCWs with a laboratory-confirmed infection (Fig. 4d (bottom)). To examine whether the reduced frequency of NSP12-specific T cells in the 16 week memory response of those with laboratory-confirmed infection was reflective of their repertoire at the time of encountering SARS-CoV-2, we obtained baseline PBMCs from a subset of individuals who were sampled before PCR positivity or more than 4 weeks before seroconversion. NSP12-specific T cells were already significantly lower at the baseline in those who went on to develop laboratory-confirmed infection compared with in SN-HCWs (Fig. 4e and Extended Data Fig. 8l,m), supporting a potential role in protection from PCR-detectable infection and seroconversion.

Conclusions

We provide T cell and innate transcript evidence for abortive, seronegative SARS-CoV-2 infection. Longitudinal samples from SN-HCWs and an additional cohort showed that RTC-specific (particularly polymerase) T cells were enriched before exposure, expanded in vivo and preferentially accumulated in those in whom SARS-CoV-2 failed to establish infection compared with those with overt infection.

The differential biasing of T cells towards early-expressed non-structural SARS-CoV-2 proteins in HCWs without seroconversion may reflect repetitive occupational exposure to very low viral inocula, reported to drive the induction of non-structural T cells in HIV, simian immunodeficiency virus (SIV) and HBV^{26,37,38}. Such repetitive exposure would be congruent with the observed protracted induction of the innate signal *IFI27* and the development of de novo T cells in some SN-HCWs.

However, we also documented the expansion of pre-existing T cells, with responses that are capable of cross-recognizing epitope variants between seasonal HCoV and SARS-CoV-2. Cross-reactive SARS-CoV-2-specific CD8⁺ T cells directed against epitopes that are highly conserved among HCoV are now well described, with pre-existing T cells frequently targeting essential viral proteins with low scope for tolerating mutational variation, such as those in ORF1ab^{6,18,32}. The abundant SARS-CoV-2-specific CD4⁺ T cells may also contribute to protection in SN-HCWs by antibody-independent mechanisms, such as antiviral cytokine and chemokine production. HCWs have higher frequencies of HCoV-reactive T cells compared with the general public¹⁹, and recent HCoV infection is associated with a reduced risk of severe COVID-19 infection³⁹, probably in part attributable to cross-reactive

neutralizing antibodies^{40,41}, however, pre-existing T cells have also been implicated^{15,42}. The early induction of T cells, before detectable antibodies in mild infection³⁰ and concurrent with mRNA vaccination efficacy, support a role for pre-existing cross-reactive memory T cells^{2,31}.

Pre-existing RTC-specific T cells, at a higher frequency than naive T cells and poised for immediate reactivation on antigen cross-recognition, would be expected to favour early control, explaining their enrichment after abortive compared to classical infection. However, the relative contribution of viral inoculum and cross-reactive T cells needs to be further dissected in human challenge experiments and animal models. A caveat of this work is that we analysed only peripheral immunity; it is plausible that mucosal-sequestered antibodies⁴³ had a role in our seronegative cohort. It also remains possible that innate immunity mediates control in abortive infections, with RTC-biased T cell responses being generated as a biomarker of low-grade infection. Interferon-independent induction of RIG-I has been proposed to abort SARS-CoV-2 infection by restraining the viral lifecycle before sgRNA production¹³; this would favour the presentation of epitopes from ORF1ab, released into the cytoplasm in the first stage of the viral life cycle¹², while blocking the production of structural proteins from pregenomic RNA. This raises the possibility that some SARS-CoV-2-infected cells could be recognized and removed by ORF1ab-reactive T cells without widespread production of structural proteins and mature virion formation.

We have described the induction of innate and cellular immunity without seroconversion, highlighting a subset of individuals in whom the risk of SARS-CoV-2 reinfection and immunogenicity of vaccines should be specifically assessed. The HCWs who we studied were exposed to Wuhan Hu-1 and had partial protection from personal protective equipment; it remains to be seen whether abortive infections can occur after exposure to more infectious variants of concern, or in the presence of vaccine-induced immunity. However, clearance without seroconversion points to T cells that may be particularly effective vaccine targets. Cross-protection between coronaviruses is proportional to their sequence homology in mice⁴⁴, making the highly conserved NSP12 region studied here, as well as less studied NSP3/14/16, top candidates for heterologous immunity. Our data highlight the presence of pre-existing T cells in a proportion of donors that are able to expand *in vivo* and target a highly conserved region of SARS-CoV-2 and other *Coronaviridae*. The boosting of such T cells may offer durable pan-*Coronaviridae* reactivity against endemic and emerging viruses, arguing for their inclusion and assessment as an adjunct to spike-specific antibodies in next-generation vaccines.

Methods

COVIDsortium healthcare worker participants

The COVIDsortium bioresource was approved by the ethical committee of UK National Research Ethics Service (20/SC/0149) and registered at <https://ClinicalTrials.gov> ([NCT04318314](#)). Full study details of the bioresource (participant screening, study design, sample collection and sample processing) have previously been described^{[16,45](#)}.

In this cohort and London as a whole, infections peaked for the first pandemic wave of infections during the first week of lockdown (23 March 2020)^{[46](#)}, and we observed approximately synchronous exposure coincident with recruitment; we therefore used this as the benchmark for assessing exposure-generated immunity. Across the main study cohort, 48 participants had positive RT–PCR results with 157 (21.5%) seropositive participants. Furthermore, 79% of positive PCR tests were within the first 2 weeks of follow-up and no HCWs tested PCR positive after week 5 of follow-up^{[14,46](#)} (Fig. [1b](#)), with seroconversion within the first 3 weeks of follow-up for most^{[28](#)}. Infections were asymptomatic or mild with only two hospital admissions (none requiring intensive care admission). The cross-sectional case controlled substudy ($n = 129$) collected samples at 16–18 weeks after the first UK lockdown (Fig. [1a](#)). Power calculations were performed before week 16 substudy sampling to determine the sample size needed to test the hypothesis that HCWs with pre-existing T cell responses are enriched in exposed seronegative group at a range of incidence of infection, assuming 50% of the total cohort had pre-existing T cell responses. Sample sizes of 18–64 per group were estimated. An age-, sex- and ethnicity-matched nested substudy was designed within the larger ($n = 731$) parent study and 129 attended for 16 week sampling, including high-volume PBMC isolation.

Laboratory-confirmed infection was determined by weekly nasopharyngeal RNA stabilizing swabs and RT–PCR (Roche cobas SARS-CoV-2 test, envelope (*E*) gene) and antibody assay positivity (spike protein 1 IgG Ab assay, EUROIMMUN) and anti-nucleocapsid total antibody assay (Roche) described in detail below. The seronegative HCW group was matched for demographics and exposure to the laboratory-confirmed infected group and was defined by negativity in these 3 tests at all 16 time points as well as negative for neutralizing antibodies at week 16 and at selected prior time points as indicated.

The cohort of medical students and laboratory staff was approved by UCL Ethics (project ID:13545/001) and prepandemic samples from healthy donors were collected and cryopreserved before August 2019 under ethics number 11/LO/0421. All participants provided written informed consent and the study conformed to the principles of the Helsinki Declaration.

Isolation of PBMCs and serum

PBMCs were isolated from heparinized blood samples using Pancoll (Pan Biotech) or Histopaque-1077 Hybri-Max (Sigma-Aldrich) density-gradient centrifugation in SepMate tubes (StemCell) according to the manufacturer's specifications. Isolated PBMCs were cryopreserved in fetal calf serum (FCS) containing 10% DMSO and stored in liquid nitrogen.

Whole-blood samples were collected in SST vacutainers (Vacutte) with inert polymer gel for serum separation and clot activator coating. After centrifugation at 1,000g for 10 min at room temperature, the serum layer was aliquoted and stored at -80 °C. All T cell assays reported here were performed on cryopreserved PBMCs.

Weekly SARS-CoV-2 S1 and NP serology

Weekly Euroimmun anti-SARS-CoV-2 enzyme-linked immunosorbent assay (ELISA; anti-SARS-CoV-2 S1 antigen IgG and the Roche Elecsys anti-SARS-CoV-2 electrochemiluminescence immunoassay (ECLIA; anti-SARS-CoV-2 nucleoprotein IgG/IgM) commercial assays were performed by Public Health England as previously described¹⁶. S1 ELISA: A ratio of ≥ 1.1 was deemed to be positive. A ratio of 11 was taken to be the upper threshold as the assay saturates beyond this point. NP ECLIA: anti-NP results are expressed as a cut-off index (COI) value based on the electrochemiluminescence signal of a two-point calibration, with results COI ≥ 1.0 classified as positive.

Neutralization assays for the pseudotype and authentic virus

SARS-CoV-2 pseudotype neutralization assays were conducted using pseudotyped lentiviral particles as previously described¹⁶. In brief, serum was heat-inactivated at 56 °C for 30 min. Serum dilutions in DMEM were performed in duplicate with a starting dilution of 1 in 20 and 7 consecutive twofold dilutions to a final dilution of 1/2,560 in a total volume of 100 µl. SARS-CoV-2 pseudotyped lentiviral particles (1×10^5 RLU) were added to each well (serum dilutions and controls) and incubated at 37 °C for 1 h. Then, 4×10^4 Huh7 cells suspended in 100 µl complete medium were added per well and incubated for 72 h at 37 °C and 5% CO₂. Firefly luciferase activity (luminescence) was measured using the Steady-Glo Luciferase Assay System (Promega) and a CLARIOStar Plate Reader (BMG Labtech). The curves of relative infection rates (as a percentage) versus the serum dilutions (\log_{10} -transformed values) against a negative control of pooled sera collected before 2016 (Sigma-Aldrich) and a positive neutralizer were plotted using Prism 9 (GraphPad). A nonlinear regression method was used to determine the dilution fold that neutralized 50% (IC₅₀).

Authentic SARS-CoV-2 microneutralization assays were performed as previously described⁴⁷. In brief, a mixture of serum dilutions in DMEM (1 in 20 and 11 consecutive twofold dilutions to a final dilution of 1/40,960) and 3×10^4 FFU of SARS-CoV-2 virus (Wuhan Hu-1) were incubated at 37 °C for 1 h. After initial incubation, preseeded Vero E6 cells were infected with the serum–virus samples and incubated (37 °C and 5% CO₂) for 72 h. Cells were then fixed with 100 µl 3.7% (v/v) formaldehyde for 1 h. Cells were washed with PBS and stained with 0.1% (w/v) crystal violet solution for 10 min. After removal of excess crystal violet and air drying, the crystal violet stain was resolubilized with 100 µl 1% (w/v) sodium dodecyl sulfate solution. Absorbance readings were taken at 570 nm using a CLARIOStar Plate Reader (BMG Labtech). Absorbance readings for each well were standardized against technical positive (virus control) and negative (cells only) controls on each plate to determine a percentage neutralization value. A nonlinear regression (curve fit) method was used to determine the dilution fold that neutralized 50% (IC₅₀) using Prism 9 (GraphPad). SARS-CoV-2 is classified as a hazard group 3 pathogen and therefore all authentic SARS-CoV-2 propagation and microneutralization assays were performed in a containment level 3 facility.

Spike ELISA

Seropositivity against SARS-CoV-2 spike was determined for medical student and laboratory staff cohort between July 2020 and Jan 2021 (Extended Data Table 5) by ELISA, as validated and described previously^{40,48,49}. In brief, 9 columns of 96-half-well MaxiSorp plates (Thermo Fisher Scientific) were coated overnight at 4 °C with purified S1 protein in PBS (3 µg ml⁻¹ per well in 25 µl), the remaining 3 columns were coated with goat anti-human F(ab)'2 (1:1,000) to generate an internal standard curve. The next day, plates were washed with PBS-T (0.05% Tween-20 in PBS) and blocked for 1 h at room temperature with assay buffer (5% milk powder PBS-T). Sera were diluted in blocking buffer (1:50). Serum (25 µl) was then added to S1 coated wells in duplicate and incubated for 2 h at room temperature. Serial dilutions of known concentrations of IgG were added to the F(ab)'2 IgG-coated wells in triplicate (Sigma-Aldrich). After incubation for 2 h at room temperature, the plates were washed with PBS-T and 25 µl alkaline phosphatase-conjugated goat anti-human IgG (Jackson ImmunoResearch) at a 1:1,000 dilution in assay buffer added to each well and incubated for 1 h room temperature. Plates were then washed with PBS-T, and 25 µl of alkaline phosphatase substrate (Sigma-Aldrich) added. Optical density values were measured using a MultiskanFC (Thermo Fisher Scientific) plate reader at 405 nm and S1-specific IgG titres interpolated from the IgG standard curve using 4PL regression curve-fitting on GraphPad Prism 8.

HCoV spike meso-scale discovery immunoassay

A multiplexed meso-scale discovery immunoassay was performed as previously described⁵⁰. Plates were coated with 200–400 µg ml⁻¹ spike protein (trimers in prefusion form) from the endemic human coronaviruses HKU1, OC43, 229E and NL63. Antibody concentration is presented in arbitrary units (AU) interpolated from the ECL signal of the internal standard sample using a four-parameter logistic curve fit. Serum samples taken at week 8—the peak time point for spike S1 IgG after PCR-positive SARS-CoV-2 infection—were assayed for HCoV antibodies.

SARS-CoV-2 spike-specific memory B cell staining

Multiparameter flow cytometry was used for ex vivo identification of spike-specific memory B cells staining as previously described²⁹. Biotinylated tetrameric spike (1 µg) was fluorochrome linked by incubating with streptavidin-conjugated APC (Prozyme) and PE (Prozyme) for 30 min in the dark on ice. PBMCs were thawed and incubated with Live/Dead fixable dead cell stain (UV, Thermo Fisher Scientific) and saturating concentrations of phenotyping monoclonal antibodies were diluted in 50% 1× PBS 50% Brilliant Violet Buffer (BD Biosciences): anti-CD3 Bv510 (BioLegend, OKT3, 1:200), anti-CD11c FITC (BD Biosciences, B-ly6, 1:100), anti-CD14 Bv510 (BioLegend, M5E2, 1:200), anti-CD19 Bv786 (BD Bioscience, HIB19, 1:50), anti-CD20 AlexFluor700 (BD Biosciences, 2H7, 1:100), anti-CD21 Bv711 (BD Biosciences, B-ly4, 1:100), anti-CD27 BUV395 (BD Biosciences, L128, 1:100), anti-CD38 Pe-CF594 (BD Biosciences, HIT2, 1:200), IgD Pe-Cy7 (BD Biosciences, IA6-2, 1:100). For identification of SARS-CoV-2-antigen-specific B cells, 1 µg per 500 µl of stain each of tetrameric spike-APC and spike-PE were added to cells. Cells were incubated in the staining solution for 30 min at room temperature, washed with PBS and subsequently fixed with the FoxP3 Buffer Set (BD Biosciences) according to the manufacturer's instructions. All of the samples were acquired on the BD Fortessa-X20 flow cytometer. Data were analysed using FlowJo v.10.7 (TreeStar). Example gating and positivity cut-off have previously been reported²⁹. The magnitude of the SARS-CoV-2 spike-specific memory B cell population is expressed as a percentage of memory B cells (gated as: lymphocytes, singlets, Live, CD3⁻CD14⁻CD19⁺, CD20⁺, excluding: CD38^{hi}, IgD⁺ and CD27⁺CD21⁻) binding both PE- and APC-labelled spike.

SARS-CoV-2 peptides

Full lists of the peptides contained in pools of overlapping peptides covering structural¹⁶ and RTC proteins⁸ have previously been described (15-mer peptides overlapping by 10 amino acids, GL Biochem Shanghai, >80% purity). A list of peptides that overlap NSP12 is provided in Supplementary Table 3. For IFNγ ELISpot

assays, SARS-CoV sequence peptides were used (96.5% sequence homology with Wuhan SARS-CoV-2 consensus sequence, 34/931 amino acids differ; Supplementary Table 3). For epitope mapping, SARS-CoV-2 sequence peptides were used for NSP12-2 and NSP12-5 (GL Biochem Shanghai, >80% purity).

To limit competition for in vitro peptide presentation, we limited stimulations to a maximum of 55 peptides and have, therefore, divided large proteins such as NP into subpools: NP (NP1, NP2, 41 peptides each), M (43 peptides), ORF3a (53 peptides), NSP7 (15), NSP12 (36–37 per pool NSP12-1 to NSP12-5) and NSP13 (39–40 peptides per pool NSP13-1 to NSP13-3). Fifteen-mer peptides covering the predicted SARS-CoV-2 spike epitopes⁸ were used to give a total of 55 peptides in this pool (spike). Optimal 9-mer peptides for CD8⁺ epitopes were custom-synthesized by ThinkPeptides (>70% purity; Supplementary Table 3).

IFN γ ELISpot assay

The IFN γ ELISpot assay was performed as previously described on cryopreserved PBMCs^{8,16,51}. Unless otherwise stated, culture medium for human PBMCs (R10) was sterile 0.22- μ m-filtered RPMI medium (Thermo Fisher Scientific) supplemented with 10% by volume heat-inactivated (1 h, 64 °C) FCS (Hyclone) and 1% by volume 100× penicillin and streptomycin solution (Gibco-BRL).

ELISpot plates (Merck-Millipore, MSIP4510) were coated with human anti-IFN γ antibodies (1-D1K, Mabtech; 10 μ g ml⁻¹) in PBS overnight at 4 °C. The plates were washed six times with sterile PBS and blocked with R10 for 2 h at 37 °C with 5% CO₂. PBMCs were thawed and rested in R10 for 3 h at 37 °C with 5% CO₂ before being counted to ensure that only viable cells were included. PBMCs (400,000 per well) were seeded in R10 and were stimulated for 16–20 h with SARS-CoV-2 peptide pools (2 μ g ml⁻¹ per peptide) at 37 °C in a humidified atmosphere with 5% CO₂. In cases in which insufficient T cells were available, NSP12 pools 1, 2 and 3, and NSP13 pools 1, 2 and 3 were combined into a single well. For baseline measurements, NSP12 pools 1–5 were stimulated in a single well and, in cases in which insufficient T cells were available, a single DMSO well was included. HCWs who did not have a full complement of stimulations were excluded from analysis of total magnitude of breadth of response, resulting in slightly lower *n* values. Internal plate controls were R10 alone (without T cells) and two DMSO wells (negative controls), concanavalin A (ConA, positive control; Sigma-Aldrich) and FEC (HLA I-restricted peptides from influenza, Epstein–Barr virus and CMV; 1 μ g ml⁻¹ per peptide). ELISpot plates were developed with human biotinylated IFN γ detection antibodies (7-B6-1, Mabtech; 1 μ g ml⁻¹) for 3 h at room temperature, followed by incubation with goat anti-biotin alkaline phosphatase (Vector Laboratories; 1:1,000) for 2 h at room temperature, both diluted in

PBS with 0.5% BSA by volume (Sigma-Aldrich), and finally with 50 µl per well of sterile filtered BCIP/NBT Phosphatase Substrate (Thermo Fisher Scientific) for 7 min room temperature. Plates were washed in double-distilled H₂O and left to dry overnight before being read on the AID classic ELISpot plate reader (Autoimmun Diagnostika).

The average of two DMSO wells was subtracted from all peptide-stimulated wells for a given PBMC sample and any response that was lower in magnitude than 2 s.d. of these sample specific DMSO control wells was not considered to be a peptide-specific response (given value 0). Results were expressed as IFN γ SFCs per 10⁶ PBMCs after background subtraction. The geometric mean of all DMSO wells was 9.571 SFCs per 10⁶ PBMCs (3.8 spots). We excluded the results if the negative control wells had >95 SFC per 10⁶ PBMCs or positive control wells (ConA) were negative. T cell responses to SARS-CoV-2 did not correlate with background spots in DMSO wells (for example, the SN-HCW group, Spearman $r = -0.068$, $P = 0.6141$).

Antigen-specific T cell proliferation assay and epitope mapping

Frozen PBMCs were thawed and washed twice with sterile PBS. PBMC were resuspended in 1 ml R10 culture medium (2–10 × 10⁶ PBMCs) and 0.5 µl of 5 mM stock CTV (Thermo Fisher Scientific) was added per sample with mixing. PBMCs were stained in the dark for 10 min at 37 °C in a humidified atmosphere with 5% CO₂. Ten-times volume of cold R10 was added to stop the staining reaction, and cells were incubated for 5 min on ice. Cells were washed in PBS and incubated for 5 min at 37 °C before being transferred to a new tube and were washed again in R10. CTV stained PBMC were plated in 96-well plates (2–4 × 10⁵ PBMCs in 200 µl R10) and stimulated with peptide pools (2 µg ml⁻¹ per peptide) for 10 days in R10 supplemented with 0.5 µg ml⁻¹ soluble anti-CD28 antibodies (Thermo Fisher scientific) and 20 U ml⁻¹ recombinant human IL-2 (Peprotech). CTV-stained and unstained PBMCs were run to confirm efficiency of staining. Then, 100 µl medium was added on day 1, and 100 µl medium was removed and replaced with R10 supplemented with anti-CD28 and IL-2 as above on days 3 and 6. On day 9, PBMCs were restimulated with peptide pools (2 µg ml⁻¹ per peptide) and brefeldin A (10 µg ml⁻¹; Sigma-Aldrich). After 16–18 h restimulation, PBMCs were collected, washed in PBS and stained for fixable live/dead (Near infrared, Thermo Fisher Scientific, 1:1,000), washed in PBS, before being fixed in fix/perm buffer (TF staining buffer kit, eBioscience) for 20 min room temperature. Cells were washed in PBS and incubated in perm buffer (TF staining buffer kit, diluted 1:10 in double-distilled H₂O) for 20 min room temperature, washed in PBS and resuspended in perm buffer with saturating concentrations of anti-human antibodies for intracellular staining: anti-IL-2 PerCp-eFluor710 (Invitrogen, MQ1-17H12, 1:50), anti-TNF α FITC (BD Bioscience, MAb11, 1:100), anti-CD8 α BV785 (BioLegend,

RPA-T8, 1:200), anti-IFN γ BV605 (BD Biosciences, B27, 1:100), anti-IFN γ APC (BioLegend, 4S.B3, 1:50), anti-CD3 BUV805 (BD Biosciences, UCHT1, 1:200), anti-CD4 BUV395 (BD Biosciences, SK3, 1:200), anti-CD154 (CD40L) Pe-Cy7 (BioLegend, 24-31, 1:50) and anti-MIP-1 β PE (BD Biosciences, D21-1351, 1:100). Cells were washed twice in PBS and analysed using the BD LSRII flow cytometer. Cytometer voltages were consistent across batches. Fluorescence minus one (FMOs) and unstimulated samples were used to determine gates applied across samples. Data were analysed using FlowJo v.10.7 (TreeStar).

Optimization experiments showed that the use of recombinant human IL-2 increases non-peptide specific proliferation of T cells but is essential for optimal expansion of proliferating cytokine producing peptide-specific T cells. CTV dilution and staining with anti-human-IFN γ antibodies was used to identify antigen-specific T cells. An unstimulated control well (equivalent DMSO to peptide wells added) was included for each PBMC sample and the percentage of CTV lo IFN γ^+ CD4 $^+$ or CD8 $^+$ cells proliferating in unstimulated wells was subtracted as background cytokine release from all peptide stimulated wells. The T cell proliferation assay above was used to expand SARS-CoV-2-specific T cells and a two-dimensional matrix (Supplementary Table 2) was used such that each 15-mer peptide was represented in 2 pools, aiding the identification of individuals immunogenic 15-mer peptides. T cell responses were then confirmed by repeated expansion with individual 15-mers.

Polyfunctionality, defined as the number of cytokines co-produced by T cells after expansion for 10 days, was assessed using SPICE (v.6.0) and pestle (v.2.0), available at GitHub (<https://naiid.github.io/spice/>)⁵². Responses <0.1% of CD4 $^+$ or CD8 $^+$ T cells were excluded. Boolean gating was used to identify the percentage of T cells making the 31 possible combinations of the following cytokines: IFN γ , TNF, IL-2, CD154, MIP-1 β . Pestle was used to background-subtract the percentage of cytokine-producing cells from unstimulated wells that were run in parallel and to format data for visualization in SPICE. The proportion of T cells making a specific number of cytokines in combination is presented as pie graphs (base mean) and pie arcs represent the proportion making a given cytokine. The RTC-specific T cell polyfunctionality was calculated as an average over T cell responses to NSP7, NSP12 and NSP13 and the structural-specific T cell polyfunctionality is an average of responses to spike, ORF3a, M and NP (Extended Data Fig. 3d).

MHC class I pentamer staining

HLA-A*02-restricted pentamers (Proimmune) of the following specificities were used: SARS-CoV-2 NSP7_{27–35} (KLWAQCVQL) or HCoV HKU1 NSP7_{27–35} (KLWQYCSVVL; ex vivo stains only). For post-expansion staining, antigen-specific T cells were expanded with a cognate 9-mer peptide of SARS-CoV-2 or HCoV HKUW

sequence for 8–10 days as above ($2 \mu\text{g ml}^{-1}$ per peptide) in R10 supplemented with $0.5 \mu\text{g ml}^{-1}$ soluble anti-CD28 antibodies and 20 U ml^{-1} recombinant human IL-2; medium was added on days 1, 3 and 6 before pentamer staining. For ex vivo staining, PBMCs were thawed, washed twice in PBS. Pentamers were centrifuged at 13,000 rpm. for 10 min before use. PBMCs ($0.5–2 \times 10^6$) were stained with 1 μl pentamers at room temperature for 20 min in 50 μl PBS in a 96-well plate. PBMCs were further stained with Blue fixable Live/dead (Invitrogen, 1:1,000) for 20 min at 4°C , and surface-stained with a mixture of saturating concentrations of monoclonal antibodies for 30 min at 4°C : anti-CD3 BUV805 (BD Biosciences, UCHT1, 1:200), anti-CD4 BUV395 (BD Biosciences, SK3, 1:200), anti-CD56 Pe-Cy7 (BD Biosciences, NCAM16.2, 1:100), anti-CD8 α Alexa700 (BioLegend, RPA-78, 1:200), post-expansion CD19 Bv786 (BD Biosciences, HIB19, 1:100). PBMCs were fixed with 1% paraformaldehyde and flow cytometry was performed as above using a BD LSRII flow cytometer. Data were analysed using FlowJo v.10.7 (TreeStar). During analysis, stringent gating criteria were applied (the gating strategy is shown in Extended Data Fig. 6c) with doublet, dead cell, CD19 $^+$ B cell (post-expansion) and CD56 $^+$ NK/NKT exclusion to minimize non-specific binding contamination. HLA-mismatched PBMC (non-HLA-A*02) and fluorescence minus one controls for pentamers were stained in parallel to assess non-specific binding (Extended Data Fig. 6c).

Coronaviridae family sequence homology analyses

The sequence homology of SARS-CoV-2-derived peptides to HCoV sequences was computed as previously described³². In brief, the SARS-CoV-2 proteome ([NC_045512.2](#)) was decomposed into 15-mer peptide sequences overlapping by 14 amino acids. A protein BLAST search of each 15-mer peptide was then performed against a custom sequence database comprising 2,531 *Coronaviridae* sequences³². Homology values of each SARS-CoV-2-derived peptide to viral accessions with ‘229E’, ‘OC43’, ‘NL63’ or ‘HKU1’ included in the species name and that were isolated from human hosts were retained (Supplementary Table 1). Moreover, to determine whether the conservation of 15-mer peptides differed between the SARS-CoV-2 proteins, the average homology of peptides within each protein was computed. A permutation test was conducted to test whether the difference in average homology between the two proteins, Δh , was statistically significant. In brief, the protein membership of each 15-mer peptide was permuted (1,000 iterations). The Δh of two proteins was then calculated at each iteration, resulting in a final null distribution of Δh values. P values were computed as the number of permutations that yielded a Δh at least as extreme as the observed Δh of the two proteins. Custom scripts used to perform the homology searches, heatmap visualization and permutation testing are available at GitHub (https://github.com/cednotsed/tcell_cross_reactivity_covid.git).

For sequence alignments of immunogenic 15-mers or at described MHC class I-restricted epitopes, reference protein sequences for ORF1ab (accession numbers: [QHD43415.1](#), [NP_828849.2](#), [YP_009047202.1](#), [YP_009555238.1](#), [YP_173236.1](#), [YP_003766.2](#) and [NP_073549.1](#)) were downloaded from the NCBI database (<https://www.ncbi.nlm.nih.gov/protein/>) as previously described⁸. Sequences were aligned using the MUSCLE algorithm with the default parameters and percentage identity was calculated in Geneious Prime 2020.1.2 (www.geneious.com). Alignment figures were generated using Snapgene 5.1 (GSL Biotech).

SARS-CoV-2 species genome diversity analyses

For genome diversity analysis, a complete masked alignment was downloaded from the GISAID^{53,54} EpiCoV database on 26 July 2021 together with a GISAID Audacity phylogeny comprising 611,893 accessions (a full list and metadata are available at Figshare (<https://figshare.com/s/049d53f789a8b111b87e>)). The alignment was subsampled to include 800 of each defined NextStrain phylogenetic clade, as provided by GISAID metadata. For clades containing less than 800 accessions, all representatives of that clade were included, resulting in a comprehensive sampling over the global phylogeny of 13,785 accessions encompassing the genomic diversity of SARS-CoV-2 to date (Extended Data Fig. [5a](#)). Diversity along the genome was assessed using two metrics of diversity: the number of recurrent mutational emergences (homoplasies) at any position and Nei's genetic diversity index⁵⁵. Homoplasy counts per locus were computed through application of the HomoplasyFinder screening pipeline⁵⁶ against a maximum likelihood phylogeny constructed over the 13,785-genome alignment. Nei's genetic diversity index was computed as $\frac{1}{I} \sum_{i=1}^I p_i^2$, where I is the count of distinct alleles at a position, and $p_i = (i=1, \dots, I)$ is the frequency of allele i in the studied alignment. The average homoplasy count per locus per gene region and average Nei's genetic diversity per gene region were computed by normalizing the per-locus values to gene length for all ORFs and NSPs according to the reference annotations of GISAID reference genome EPI_ISL_402124. Significant differences between all pairwise combinations of ORF/NSP were assessed using the Wilcoxon rank-sum test implemented in `compare_means()` in the R package `ggpubr` v.0.4.0 (Extended Data Table [3](#)).

IFI27 qPCR

Total RNA from Tempus blood was extracted using the Tempus Spin RNA isolation kit (Applied Biosystems, 4380204). cDNA was obtained using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Quantitative PCR (qPCR) was performed using the TaqMan Fast Advanced Master Mix (Applied Biosystems)

on the ABI StepOnePlus Real-Time PCR machine (Applied Biosystems). The following cycling conditions were used: 95 °C for 2 min; followed by 40 cycles of 95 °C for 3 s and 60 °C for 30 s. *IFI27* and *GAPDH* were amplified using the TaqMan Gene Expression Assay probes Hs01086373_g1 (*IFI27*) and Hs02786624_g1 (*GAPDH*). *GAPDH* was used as a housekeeping gene control. The unexposed prepandemic control HCW cohort for qPCR analysis was described previously⁵⁷

Correlogram plot

A pairwise correlation matrix between variables was calculated and visualized as a correlogram using corrplot (<https://github.com/taiyun/corrplot>) in R v.3.5.3 with R studio v.1.0.153. The Spearman's rank correlation coefficient r is indicated by the size and colour of the circles. Only correlations with $P < 0.05$ are shown. Variables are ordered by hierarchical clustering.

Statistics and reproducibility

Data were assumed to have a non-Gaussian distribution and nonparametric tests were used throughout. For single-paired and unpaired comparisons, Wilcoxon matched-pairs signed-rank tests and a Mann-Whitney U -tests were used. For multiple unpaired comparisons, Kruskal-Wallis one-way ANOVA with Dunn's correction was used. For correlations, Spearman's r test was used. $P < 0.05$ was considered to be significant. Prism v.7.0e and v.8.0 for Mac was used for analysis. Details are provided in the figure legends.

Data reporting

Power calculations were used to estimate the sample size needed for the week 16 substudy (see above). No statistical methods were used to predetermine the sample size. For all of the assays, samples from each cohort were run in parallel to reduce the impact of interbatch technical variation. IFN γ ELISpot assays were performed on HCW cohorts before unblinding of the group (laboratory-confirmed infection or SN-HCW). Other experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

All data analysed during this study are included in this published article and its Supplementary Information. Genomic data analysed was obtained from the publicly available NCBI Virus database and, after registration, from the GISAID EpiCoV repository. The datasets generated during and/or analysed during the current study are available from the corresponding authors on reasonable request. Correspondence and requests for materials should be addressed to M.K.M. or L.S. [Source data](#) are provided with this paper.

Code availability

Custom scripts that were used to perform the homology searches, heatmap visualization and permutation testing are available at GitHub (https://github.com/cednotsed/tcell_cross_reactivity_covid.git).

References

1. 1.

Tan, A. T. et al. Early induction of functional SARS-CoV-2-specific T cells associates with rapid viral clearance and mild disease in COVID-19 patients. *Cell Rep.* **34**, 108728 (2021).

2. 2.

Oberhardt, V. et al. Rapid and stable mobilization of CD8⁺ T cells by SARS-CoV-2 mRNA vaccine. *Nature* **597**, 268–273 (2021).

3. 3.

Zhao, J. J. et al. Airway memory CD4⁺ T cells mediate protective immunity against emerging respiratory coronaviruses. *Immunity* **44**, 1379–1391 (2016).

4. 4.

Sekine, T. et al. Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. *Cell* **183**, 158–168 (2020).

5. 5.

Mateus, J. et al. Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science* **370**, 89–94 (2020).

6. 6.

Nelde, A. et al. SARS-CoV-2-derived peptides define heterologous and COVID-19-induced T cell recognition. *Nat. Immunol.* **22**, 74–85 (2021).

7. 7.

Braun, J. et al. SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. *Nature* **587**, 270–274 (2020).

8. 8.

Le Bert, N. et al. SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* **584**, 457–462 (2020).

9. 9.

Grifoni, A. et al. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* **181**, 1489–1501 (2020).

10. 10.

Schulien, I. et al. Characterization of pre-existing and induced SARS-CoV-2-specific CD8⁺ T cells. *Nat. Med.* **27**, 78–85 (2021).

11. 11.

Loyal, L. et al. Cross-reactive CD4⁺ T cells enhance SARS-CoV-2 immune responses upon infection and vaccination. *Science* **374**, eabh1823 (2021).

12. 12.

Perlman, S. & Netland, J. Coronaviruses post-SARS: update on replication and pathogenesis. *Nat. Rev. Microbiol.* **7**, 439–450 (2009).

13. 13.

Yamada, T. et al. RIG-I triggers a signaling-abortive anti-SARS-CoV-2 defense in human lung cells. *Nat. Immunol.* **22**, 820–828 (2021).

14. 14.

Gupta, R. K. et al. Blood transcriptional biomarkers of acute viral infection for detection of pre-symptomatic SARS-CoV-2 infection: a nested, case-control diagnostic accuracy study. *Lancet Microbe* **2**, E508–E517 (2021).

15. 15.

Mallajosyula, V. et al. CD8⁺ T cells specific for conserved coronavirus epitopes correlate with milder disease in COVID-19 patients. *Sci. Immunol.* **6**, eabg5669 (2021).

16. 16.

Reynolds, C. J. et al. Discordant neutralizing antibody and T cell responses in asymptomatic and mild SARS-CoV-2 infection. *Sci. Immunol.* **5**, eabf3698 (2020).

17. 17.

Low, J. S. et al. Clonal analysis of immunodominance and cross-reactivity of the CD4 T cell response to SARS-CoV-2. *Science* **372**, eabg8985 (2021).

18. 18.

Ferretti, A. P. et al. Unbiased screens show CD8⁺ T cells of COVID-19 patients recognize shared epitopes in SARS-CoV-2 that largely reside outside the spike protein. *Immunity* **53**, 1095–1107 (2020).

19. 19.

da Silva Antunes, R. et al. Differential T-cell reactivity to endemic coronaviruses and SARS-CoV-2 in community and health care workers. *J. Infect. Dis.* **224**, 70–80 (2021).

20. 20.

Wang, Z. et al. Exposure to SARS-CoV-2 generates T-cell memory in the absence of a detectable viral infection. *Nat. Commun.* **12**, 1724 (2021).

21. 21.

Ogbe, A. et al. T cell assays differentiate clinical and subclinical SARS-CoV-2 infections from cross-reactive antiviral responses. *Nat. Commun.* **12**, 2055 (2021).

22. 22.

Gallais, F. et al. Intrafamilial exposure to SARS-CoV-2 associated with cellular immune response without Seroconversion, France. *Emerg. Infect. Dis.* **27**, 113–121 (2021).

23. 23.

Wiegand, J. et al. HBV-specific T-cell responses in healthy seronegative sexual partners of patients with chronic HBV infection. *J. Viral Hepat.* **17**, 631–639 (2010).

24. 24.

Promadej, N. et al. Broad human immunodeficiency virus (HIV)-specific T cell responses to conserved HIV proteins in HIV-seronegative women highly exposed to a single HIV-infected partner. *J. Infect. Dis.* **187**, 1053–1063 (2003).

25. 25.

Heller, T. et al. Occupational exposure to hepatitis C virus: early T-cell responses in the absence of seroconversion in a longitudinal cohort study. *J. Infect. Dis.* **208**, 1020–1025 (2013).

26. 26.

Rowland-Jones, S. et al. HIV-specific cytotoxic T-cells in HIV-exposed but uninfected Gambian women. *Nat. Med.* **1**, 59–64 (1995).

27. 27.

Turtle, L. et al. Human T cell responses to Japanese encephalitis virus in health and disease. *J. Exp. Med.* **213**, 1331–1352 (2016).

28. 28.

Manisty, C. et al. Time series analysis and mechanistic modelling of heterogeneity and sero-reversion in antibody responses to mild SARS-CoV-2 infection. *EBioMedicine* **65**, 103259 (2021).

29. 29.

Jeffery-Smith, A. et al. SARS-CoV-2-specific memory B cells can persist in the elderly who have lost detectable neutralising antibodies. *J. Clin. Invest.*,

<https://doi.org/10.1172/JCI152042> (2021).

30. 30.

Chandran, A. et al. Non-severe SARS-CoV-2 infection is characterised by very early T cell proliferation independent of type 1 interferon responses and distinct from other acute respiratory viruses. Preprint at <https://doi.org/10.1101/2021.03.30.21254540> (2021).

31. 31.

Kalimuddin, S. et al. Early T cell and binding antibody responses are associated with COVID-19 RNA vaccine efficacy onset. *Med* **2**, 682–688 (2021).

32. 32.

Tan, C. C. S. et al. Pre-existing T cell-mediated cross-reactivity to SARS-CoV-2 cannot solely be explained by prior exposure to endemic human coronaviruses. *Infect. Genet. Evol.* **95**, 105075 (2021).

33. 33.

Buckley, P. R., Lee, C. H., Pinho, M. P. & Babu, R. O. HLA-dependent variation in SARS-CoV-2 CD8⁺ T cell cross-reactivity with human coronaviruses. Preprint at <https://doi.org/10.1101/2021.07.17.452778> (2021).

34. 34.

Nelde, A. et al. SARS-CoV-2 T-cell epitopes define heterologous and COVID-19-induced T-cell recognition. *Nat. Immunol.* **22**, 74–85 (2020).

35. 35.

Grifoni, A. et al. SARS-CoV-2 human T cell epitopes: adaptive immune response against COVID-19. *Cell Host Microbe* **29**, 1076–1092 (2021).

36. 36.

Anderson, E. M. et al. Seasonal human coronavirus antibodies are boosted upon SARS-CoV-2 infection but not associated with protection. *Cell* **184**, 1858–1864 (2021).

37. 37.

Clerici, M. et al. T-cell proliferation to subinfectious SIV correlates with lack of infection after challenge of macaques. *AIDS* **8**, 1391–1395 (1994).

38. 38.

Werner, J. M., Abdalla, A., Gara, N., Ghany, M. G. & Rehermann, B. The hepatitis B vaccine protects re-exposed health care workers, but does not provide sterilizing immunity. *Gastroenterology* **145**, 1026–1034 (2013).

39. 39.

Sagar, M. et al. Recent endemic coronavirus infection is associated with less-severe COVID-19. *J. Clin. Invest.* **131**, e143380 (2021).

40. 40.

Ng, K. W. et al. Preexisting and de novo humoral immunity to SARS-CoV-2 in humans. *Science* **370**, 1339–1343 (2020).

41. 41.

Yang, F. et al. Shared B cell memory to coronaviruses and other pathogens varies in human age groups and tissues. *Science* **372**, 738–741 (2021).

42. 42.

Wyllie, D. et al. SARS-CoV-2 responsive T cell numbers and anti-Spike IgG levels are both associated with protection from COVID-19: a prospective cohort study in keyworkers. Preprint at <https://doi.org/10.1101/2020.11.02.20222778> (2021).

43. 43.

Cervia, C. 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 (2021).

44. 44.

Dangi, T. et al. SARS coronavirus vaccines protect against different coronaviruses. Preprint at <https://doi.org/10.1101/2021.06.01.446491> (2021).

45. 45.

Augusto, J. B. et al. Healthcare Workers Bioresource: study outline and baseline characteristics of a prospective healthcare worker cohort to study immune protection and pathogenesis in COVID-19. *Wellcome Open Res.* **5**, 179 (2020).

46. 46.

Treibel, T. A. et al. COVID-19: PCR screening of asymptomatic health-care workers at London hospital. *Lancet* **395**, 1608–1610 (2020).

47. 47.

Reynolds, C. J. et al. Prior SARS-CoV-2 infection rescues B and T cell responses to variants after first vaccine dose. *Science* **5**, eabh1282 (2021).

48. 48.

O’Nions, J. et al. SARS-CoV-2 antibody responses in patients with acute leukaemia. *Leukemia* **35**, 289–292 (2021).

49. 49.

Muir, L. et al. Neutralizing antibody responses after SARS-CoV-2 infection in end-stage kidney disease and protection against reinfection. *Kidney Int. Rep.* **6**, 1799–1809 (2021).

50. 50.

Johnson, M. et al. Evaluation of a novel multiplexed assay for determining IgG levels and functional activity to SARS-CoV-2. *J. Clin. Virol.* **130**, 104572 (2020).

51. 51.

Capone, S. et al. Optimising T cell (re)boosting strategies for adenoviral and modified vaccinia Ankara vaccine regimens in humans. *npj Vaccines* **5**, 94 (2020).

52. 52.

Roederer, M., Nozzi, J. L. & Nason, M. C. SPICE: exploration and analysis of post-cytometric complex multivariate datasets. *Cytom. A* **79**, 167–174 (2011).

53. 53.

Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data—from vision to reality. *Eurosurveillance* **22**, 30494 (2017).

54. 54.

Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID’s innovative contribution to global health. *Glob. Chall.* **1**, 33–46 (2017).

55. 55.

Nei, M. Analysis of gene diversity in subdivided populations. *Proc. Natl Acad. Sci. USA* **70**, 3321–3323 (1973).

56. 56.

Crispell, J., Balaz, D. & Gordon, S. V. HomoplasyFinder: a simple tool to identify homoplasies on a phylogeny. *Microb. Genom.* **5**, e000245 (2019).

57. 57.

Pollara, G. et al. Exaggerated IL-17A activity in human *in vivo* recall responses discriminates active tuberculosis from latent infection and cured disease. *Sci. Transl. Med.* **13**, eabg7673 (2021).

Acknowledgements

We thank all of the patients and control volunteers who participated in this study and all of the clinical staff who helped with recruitment and sample collection; J. Evans at the Rayne Building FACS facility for assistance with flow cytometry assays; and members of all of the contributing and submitting laboratories around the globe who have openly shared large numbers of UK SARS-CoV-2 assemblies. A full list of acknowledgements providing submitting and originating laboratories is provided at Figshare (<https://figshare.com/s/049d53f789a8b111b87e>). The COVIDsortium is supported by funding donated by individuals, charitable Trusts and corporations, including Goldman Sachs, K. C. Griffin, The Guy Foundation, GW Pharmaceuticals, Kusuma Trust and Jagclif Charitable Trust, and enabled by Barts Charity with support from UCLH Charity. Wider support is acknowledged on the COVIDsortium website. Institutional support from Barts Health NHS Trust and Royal Free NHS Foundation Trust facilitated study processes, in partnership with University College London and Queen Mary University London. This study was funded by UKRI/NIHR UK-CIC (supporting L.S. and M.K.M.). M.K.M. is also supported by Wellcome Trust Investigator Award (214191/Z/18/Z) and CRUK Immunology grant (26603), and L.S. by a Medical Research Foundation fellowship (044-0001). M.N. is supported by the

Wellcome Trust (207511/Z/17/Z) and by NIHR Biomedical Research Funding to UCL and UCLH; A.B. by a Special NUHS COVID-19 Seed Grant Call, Project NUHSRO/2020/052/RO5+5/NUHS-COVID/6 (WBS R-571-000-077-733). J.C.M., C.M. and T.A.T. are directly and indirectly supported by the University College London Hospitals (UCLH) and Barts NIHR Biomedical Research Centres and through the British Heart Foundation (BHF) Accelerator Award (AA/18/6/34223). T.A.T. is funded by a BHF Intermediate Research Fellowship (FS/19/35/34374). A.M.V., Á.M., C.M. and J.C.M. were supported by the UKRI/MRC Covid-19 Rapid response grant COV0331. Á.M. and C.P. are supported by Rosetrees Trust, The John Black Charitable Foundation, and Medical College of St Bartholomew's Hospital Trust and NIHR-MRC grant MR/V027883/1. R.J.B. and D.M.A. are supported by the MRC (MR/S019553/1, MR/R02622X/1, MR/V036939/1, MR/W020610/1), NIHR Imperial Biomedical Research Centre (BRC): ITMAT, Cystic Fibrosis Trust SRC (2019SRC015), and Horizon 2020 Marie Skłodowska-Curie Innovative Training Network (ITN) European Training Network (no. 860325). Funding for the HLA imputed data was provided by UKRI/MRC COVID-19 rapid response grant (Cov-0331, MR/V027883/1). L.E.M. is supported by a Medical Research Council Career Development Award (MR/R008698/1). L.v.D. is supported by a UCL Excellence Fellowship. The funders had no role in study design data collection, data analysis, data interpretation or writing of the report.

Author information

Author notes

1. These authors contributed equally: Mariana O. Diniz, Nathalie M. Schmidt, Oliver E. Amin, Aneesh Chandran, Emily Shaw, Mahdad Noursadeghi, Antonio Bertolletti

Affiliations

1. Division of Infection and Immunity, University College London, London, UK

Leo Swadling, Mariana O. Diniz, Nathalie M. Schmidt, Oliver E. Amin, Aneesh Chandran, Emily Shaw, Anna Jeffery-Smith, Stephanie Kucykowicz, Gloryanne Aidoo-Micah, Joshua Rosenheim, Jessica Davies, Laura E. McCoy, Benjamin M. Chain, Aroon D. Hingorani, Mahdad Noursadeghi, Mahdad Noursadeghi & Mala K. Maini

2. Blizard Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Corinna Pade, Joseph M. Gibbons, Anna Jeffery-Smith, Natalie Bullock, Wing-Yiu Jason Lee, Meleri Jones, Áine McKnight & Áine McKnight

3. Emerging Infectious Diseases Program, Duke-NUS Medical School, Singapore, Singapore

Nina Le Bert, Anthony T. Tan, Christine Y. L. Tham & Antonio Bertozetti

4. UCL Genetics Institute, University College London, London, UK

Cedric C. S. Tan, Lucy van Dorp & Francois Balloux

5. Great Ormond Street Institute of Child Health NIHR Biomedical Research Centre, University College London, London, UK

Marina Johnson & David Goldblatt

6. Barts Heart Centre, St Bartholomew's Hospital, Barts Health NHS Trust, London, UK

Melanie P. Jensen, George Joy, Charlotte Manisty, Thomas A. Treibel, James C. Moon, Hakam Abbass, Aderonke Abiodun, Masha Alfarah, Zoe Alldis, Mervyn Andiapen, Jessica Artico, João B. Augusto, Georgina L. Baca, Anish N. Bhuva, Alex Boulter, Ruth Bowles, Ben O'Brien, Nicola Champion, Carmen Chan, Jorge Couto de Sousa, Xose Couto-Parada, Teresa Cutino-Mogue, Rhodri H. Davies, Keenan Dieobi-Anene, Malcolm Finlay, Nasim Forooghi, Matt Hamblin, Gemma Hughes, Rebecca Hughes, Ivie Itua, Victor Jardim, Melanie Petra Jensen, George Joy, Vikas Kapil, Hibba Kurdi, Jonathan Lambourne, Vineela Mandadapu, Katia Menacho, Celina Mfuko, Oliver Mitchelmore, Susana Palma, Kush Patel, Steffen E. Petersen, Brian Piniera, Daniel Pope, Mary Prossora, Amy Richards, Genine Sambile, Andreas Seraphim, George D. Thornton, Art Tucker, Jessry Veerapen, Mohit Vijayakumar, Sophie Welch, Theresa Wodehouse & Lucinda Wynne

7. Department of Cellular Pathology, Northwest London Pathology, Imperial College London NHS Trust, London, UK

Melanie P. Jensen & Melanie Petra Jensen

8. Institute of Cardiovascular Science, University College London, London, UK

George Joy, Charlotte Manisty, Thomas A. Treibel, James C. Moon, Aderonke Abiodun, Gabriella Captur, Lee Howes, Alun Hughes, Rebecca Hughes, George Joy, Hibba Kurdi, Alicja Rapala & George D. Thornton

9. Academic Rheumatology, Clinical Sciences, Nottingham City Hospital, Nottingham, UK

Ana M. Valdes

10. NIHR Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust and University of Nottingham, Nottingham, UK

Ana M. Valdes

11. Department of Immunology and Inflammation, Imperial College London, London, UK

Daniel M. Altmann

12. Department of Infectious Disease, Faculty of Medicine, Imperial College London, London, UK

Rosemary J. Boyton, Sasha N. L. Bailey, Rosemary J. Boyton, David K. Butler, Kai-Min Lin, Diana Munoz-Sandoval, Sam M. Murray, Franziska P. Pieper & Catherine J. Reynolds

13. Lung Division, Royal Brompton & Harefield Hospitals, Guy's and St Thomas' NHS Foundation Trust, London, UK

Rosemary J. Boyton & Rosemary J. Boyton

14. Singapore Immunology Network, A*STAR, Singapore, Singapore

Karen Feehan, Celia Gaier, Michelle Sugimoto & Antonio Bertoletti

15. Division of Medicine, University College London, London, UK

Olivia V. Bracken

16. German Heart Centre and Charité University, Berlin, Germany

Ben O'Brien

17. National Infection Service, Public Health England, Porton Down, UK

Tim Brooks, Anaya Ellis, Gabrielle Harker, Jacqueline Hewson, Jessica Jones, Christopher Moon, Ashley Otter, Amanda Semper, Stephen Taylor & Stephen Thomas

18. Royal Free London NHS Foundation Trust, London, UK

Gabriella Captur, Brooke Douglas, Marianna Fontana, Sarah Louth, Babita Pawarova, Lisa Rannigan, Mihaela Simion & Angelique Smit

19. William Harvey Research Institute, Queen Mary University of London, London, UK

David Collier, Vikas Kapil, Daniel Pope & Matthew Robathan

20. Viral Pseudotype Unit, , Medway School of Pharmacy, Chatham, UK

Cecilia Di Genova & Nigel Temperton

21. Centre for Clinical Pharmacology, University College London, London, UK

Derek Gilroy

22. East London NHS Foundation Trust Unit for Social and Community Psychiatry, Newham Centre for Mental Health, London, UK

Lauren M. Hickling

23. School of Clinical Medicine, University of Cambridge, Cambridge, UK

Ruth Parker & Dan Zahedi

Consortia

COVIDsortium Investigators

- Hakam Abbass
- , Aderonke Abiodun
- , Mashael Alfarih
- , Zoe Alldis
- , Daniel M. Altmann
- , Oliver E. Amin
- , Mervyn Andiapen
- , Jessica Artico
- , João B. Augusto
- , Georgina L. Baca
- , Sasha N. L. Bailey
- , Anish N. Bhuva

- , Alex Boulter
- , Ruth Bowles
- , Rosemary J. Boyton
- , Olivia V. Bracken
- , Ben O'Brien
- , Tim Brooks
- , Natalie Bullock
- , David K. Butler
- , Gabriella Captur
- , Nicola Champion
- , Carmen Chan
- , Aneesh Chandran
- , David Collier
- , Jorge Couto de Sousa
- , Xose Couto-Parada
- , Teresa Cutino-Mogue
- , Rhodri H. Davies
- , Brooke Douglas
- , Cecilia Di Genova
- , Keenan Dieobi-Anene
- , Mariana O. Diniz
- , Anaya Ellis
- , Karen Feehan
- , Malcolm Finlay
- , Marianna Fontana
- , Nasim Forooghi
- , Celia Gaier
- , Joseph M. Gibbons
- , Derek Gilroy
- , Matt Hamblin
- , Gabrielle Harker
- , Jacqueline Hewson
- , Lauren M. Hickling
- , Aroon D. Hingorani
- , Lee Howes
- , Alun Hughes
- , Gemma Hughes
- , Rebecca Hughes
- , Ivie Itua
- , Victor Jardim
- , Wing-Yiu Jason Lee
- , Melanie Petra Jensen

- , Jessica Jones
- , Meleri Jones
- , George Joy
- , Vikas Kapil
- , Hibba Kurdi
- , Jonathan Lambourne
- , Kai-Min Lin
- , Sarah Louth
- , Mala K. Maini
- , Vineela Mandadapu
- , Charlotte Manisty
- , Áine McKnight
- , Katia Menacho
- , Celina Mfuko
- , Oliver Mitchelmore
- , Christopher Moon
- , James C. Moon
- , Diana Munoz-Sandoval
- , Sam M. Murray
- , Mahdad Noursadeghi
- , Ashley Otter
- , Corinna Pade
- , Susana Palma
- , Ruth Parker
- , Kush Patel
- , Babita Pawarova
- , Steffen E. Petersen
- , Brian Piniera
- , Franziska P. Pieper
- , Daniel Pope
- , Mary Prossora
- , Lisa Rannigan
- , Alicja Rapala
- , Catherine J. Reynolds
- , Amy Richards
- , Matthew Robathan
- , Joshua Rosenheim
- , Genine Sambile
- , Nathalie M. Schmidt
- , Amanda Semper
- , Andreas Seraphim
- , Mihaela Simion

- , Angelique Smit
- , Michelle Sugimoto
- , Leo Swadling
- , Stephen Taylor
- , Nigel Temperton
- , Stephen Thomas
- , George D. Thornton
- , Thomas A. Treibel
- , Art Tucker
- , Jessry Veerapen
- , Mohit Vijayakumar
- , Sophie Welch
- , Theresa Wodehouse
- , Lucinda Wynne
- & Dan Zahedi

Contributions

M.K.M. conceived the project and obtained funding. L.S., M.N., A.B. and M.K.M. designed experiments. C.M., T.A.T., J.C.M., M.N. and Á.M. established the HCW cohort. L.S., M.O.D., N.M.S., O.E.A., C.P., J.M.G., S.K., G.A.-M., J.R., J.D. and G.J. collected or processed HCW samples with COVIDsortium investigators. M.K.M., L.S., M.N. and E.S. established medical student/laboratory staff and prepandemic cohorts (UK). L.E.M., M.J., D.G. and COVIDsortium investigators performed serology. M.K.M. and L.S. designed T cell experiments. L.S., M.O.D., N.M.S. and O.E.A. developed, performed and analysed the T cell experiments. A.B., N.L.B., A.T.T. and C.Y.L.T. performed T cell assays and analysed data from the prepandemic cohort (Singapore). A.C. performed and analysed blood transcriptomic experiments. A.M.V. supervised HLA analysis. Á.M. supervised neutralizing antibody experiments. J.M.G. and C.P. performed and analysed neutralizing antibody experiments. C.C.S.T., L.v.D. and F.B. performed viral sequence analysis. E.S., M.P.J., G.J., R.J.B., C.M., T.A.T., J.C.M., Á.M. and M.O.D. provided or processed essential clinical data. M.J. and D.G. performed and analysed HCoV serology. L.S., M.O.D., N.M.S., O.E.A., A.C., C.P., J.M.G., N.L.B., A.T.T., A.J.-S., C.C.S.T., C.Y.L.T., A.M.V., B.M.C., D.G., L.v.D., D.M.A., R.J.B., C.M., T.A.T., L.E.M., F.B., Á.M., M.N., A.B. and M.K.M. analysed and interpreted the data. L.S. and M.K.M. prepared the manuscript. All of the authors reviewed the manuscript.

Corresponding authors

Correspondence to [Leo Swadling](#) or [Mala K. Maini](#).

Ethics declarations

Competing interests

A.B. is a cofounder of Lion TCR, a biotechnology company that develops T cell receptors for the treatment of virus-related diseases and cancers. R.J.B. and D.M.A. are members of the Global T-cell Expert Consortium and have consulted for Oxford Immunotec outside the submitted work. The other authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Stanley Perlman and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

[Extended Data Fig. 1 SARS-CoV-2 immunity in seronegative HCW – authentic virus neutralization and T cell response in those with NP1+NP2 responses.](#)

a, authentic virus neutralization (Wuhan Hu-1). **b**, Example plots of SARS-CoV-2 spike memory B cell (MBC) staining (gated on: lymphocytes/singlets/Live, CD3-CD14-CD19+/CD20+, excluding CD38^{hi}, IgD+ and CD21+CD27- fractions) and frequency of SARS-CoV-2 spike-specific MBC in pre-pandemic or SN-HCW (wk16; as a percentage of total

MBC). Bars, median. **c**, Proportion of cohorts with T cell responses to NP1 and/or NP2 subpools. **d**, Magnitude of T cell response coloured by viral protein and **e**, summed response to RTC and structural regions in SN-HCW with T cells reactive against both NP1 and NP2 and against one of or neither NP1 or NP2 pools at wk16. Kruskal-Wallis with Dunn's correction. Bars, geomean. **a-e**, COVIDsortium HCW cohort. NP, nucleoprotein; RTC, replication-transcription complex.

[Source Data](#)

Extended Data Fig. 2 T cell responses to RTC and structural regions of SARS-CoV-2 by cohort.

a, T cell response to Flu, EBV and CMV (FEC) MHC class I restricted peptide pool. **b**, E gene RT-PCR cycle threshold value vs. magnitude of T cell response to RTC or structural proteins in HCW with laboratory-confirmed infection. **c**, Magnitude of T cell response to RTC vs. structural regions. **d**, Magnitude of T cell response to RTC (top) and structural regions (bottom) coloured by specificity. **e**, Magnitude of T cell response in laboratory-confirmed infection group in HCW with or without detectable neutralizing antibodies at wk16. **f**, T cell response to RTC coloured by protein in laboratory-confirmed infection group ordered by magnitude. HCW lacking neutralizing antibodies highlighted by arrows below. **a-f**, IFN γ ELISpot wk16. **a**, Red lines, geomean. **e**, Bars, geomean. **b-c** Spearman r. **a,d** Kruskal-Wallis ANOVA with Dunn's correction. **a-f**, COVIDsortium HCW cohort.

[Source Data](#)

Extended Data Fig. 3 Functional and proliferative SARS-CoV-2 specific T cells in seronegative HCWs.

a, (Upper) Example gating of CTV stained PBMC after 10-day peptide stimulation: Lymphocytes (SSC-A vs. FSC-A), single cells (FSC-H vs. FSC-A), Live cells (fixable live/dead-), CD3+, CD4+ or CD8+. Second row: Gated on CD8+ showing cytokine/intracellular protein combinations. Response to immunodominant MHC class I-restricted peptide pool against

Flu, EBV, CMV (FEC) in SN-HCW. (Lower) example CTV and IFN γ staining in a SN-HCW (gated on CD4+ [black] or CD8+ [blue] T cells, percentage CTV $^{\text{lo}}$ IFN γ^+ shown). **b**, Correlation between the magnitude of T cells responses to SARS-CoV-2 pools or FEC after 10-day in vitro expansion (% dual staining for two anti-human IFN γ mAb clones, responses <0.1% of CD3 post-expansion excluded) and ex vivo IFN γ ELISpot in SN-HCWs. Spearman r. **c**, Example plots of dual cytokine or activation marker staining of SARS-CoV-2-specific T cells in an SN-HCW after 10-day expansion with peptide pools (proliferating T cells become CTV $^{\text{lo}}$ as they divide and dilute out marker). SARS-CoV-2-specific T cells highlighted in red (CD4+ CTV $^{\text{lo}}$ IFN γ^+). Percentage of CD4+ shown. **d**, polyfunctionality of CD4+ and CD8+ T cells targeting the RTC or structural regions of SARS-CoV-2 or FEC peptide pool (proportion of cytokine producing T cells that co-producing a given number of cytokines after 10-day peptide stimulation). Pie base, mean. Pie arcs show proportion of cells producing a given cytokine. **e**, Proportion of SARS-CoV-2-specific T cells (CTV $^{\text{lo}}$ IFN γ^+) that are CD4+ or CD8+ after 10-day expansion (the protein specificity is listed above, donor ID (**a-l**, corresponding to raw data in Extended Data Table [2](#)) and peptide subpools used for stimulation listed below). **a-e**, SN-HCW at wk16 COVIDsortium HCW cohort.

[Extended Data Fig. 4 Slope and variance of IFI27 signal in seronegative HCWs.](#)

a, Subsetting SN-HCW group into those with weak (n=5, <50 SFCs per 10⁶ PBMCs) or strong (n=20, >50 SFCs per 10⁶ PBMCs) RTC-specific T cell responses at wk16. **b**, Slope and **c**, variance of IFI27 signal (wk0-5) in SN-HCW with weak (n=5) or strong (n=15) RTC-specific T cell responses at wk16. **d**, Correlation matrix of variables for SN-HCW (colour and size of dots represent spearman's r, only correlations p<0.05 shown; peak IFI27 signal from wk0-5, T cell responses at wk16 to proteins, regions [RTC or structural], or total SARS-CoV-2 response). **b,c**, Mann-Whitney test, Red lines at median. **a-d**, COVIDsortium HCW cohort.

[Source Data](#)

Extended Data Fig. 5 Diversity along SARS-CoV-2 genome.

a, Radial phylogeny of SARS-CoV-2 sequence diversity (611,893 genomes) with the 13,785 accessions subsampled for diversity analysis shown in red. **b**, Genetic diversity (Nei's genetic diversity index) at individual nucleotides along the SARS-CoV-2 genomes, together with the density of polymorphic nucleotides over an 100-nucleotide sliding window shown in grey shading (right y-axis) and **c**, Homoplasies (recurrent mutational emergences) at individual nucleotides, together with the density of the number of homoplasies recorded over an 100-nucleotide sliding window shown in grey shading (right y-axis). **d**, Mean number of homoplasies across a given protein. Viral proteins not assayed for T cell responses are shown in grey.

Source Data

Extended Data Fig 6 Cross-reactive coronavirus-specific T cells in seronegative HCWs.

a, Example 2D-mapping matrix after 10-day expansion with NSP12-3 peptide pool in an SN-HCW (antigen-specific, CTV^{lo}IFN γ +; percentage of CD4+ or CD8+ shown). **b**, Alignment of *Coronaviridae* consensus sequences at immunogenic 15mers peptides (Extended Data Table 4). Conserved amino acids in yellow. **c**, (left) Example gating (lymphocytes (SSC-A,FSC-A)/singlets(FSC-A,FSC-H)/Live(Live-dead⁻)/CD3+CD19-/CD8+CD4-/CD56-; example of staining in HLA-mismatched donor and fluorescence minus one for pentamer shown) and (right above) pentamer stains of PBMC from SN-HCW at wk16-26 ex vivo (co-staining of pentamers loaded with SARS-CoV-2 peptide KLWAQCVQL and HKU1 peptide KLWQYCSV) and (right below) after 10-day expansion with SARS-CoV-2 peptide or HKU1 peptide (stained with SARS-CoV-2 peptide loaded pentamer). Percentage of CD8+ shown. **d**, Alignment of *Coronaviridae* sequences at HLA-A*02-restricted epitope in NSP7 (left) and magnitude of CD8+ T cell response (CTV^{lo}IFN γ +) after 10-day expansion with HCoV variant sequence peptides as a percentage of response with SARS-CoV-2 sequence peptide (middle) or absolute percentage of total CD8+ (right). **e**, Example plot of CTV vs. IFN γ after 10-

day expansion with SARS-CoV-2 or HCoV sequence 9-mer peptides (gated on lymphocytes/singlets/live cells/CD3+/CD56-CD4-/CD8+). **f**, Alignment of *Coronaviridae* sequences at B*035-restricted epitope in NSP12 (left), magnitude of CD8+ T cell response ($\text{CTV}^{\text{lo}}\text{IFN}\gamma+$) after 10-day expansion with HCoV variant sequence peptides as a percentage of response with SARS-CoV-2 sequence peptide (middle) or absolute percentage of total CD8+ (right). **d-f**, Conserved amino acids in yellow. **d-f**, SN-HCW wk16. **a, c-f**, COVIDsortium HCW cohort. **d-f**, SN-HCW wk16. **d,f**, Lines, median.

[Extended Data Fig. 7 Anti-spike IgG to human endemic coronaviruses.](#)

Anti-spike IgG titres were measured post-infection (wk8, time of peak SARS-CoV-2 S1 IgG seropositivity in COVIDsortium HCW cohort) in HCW with laboratory-confirmed infection (n=20), and post-exposure (wk8) in SN-HCW with weak (<50 SFCs per 10^6 PBMCs, n=19) or strong RTC-specific T cell response at wk16 (n=19, >50 SFCs per 10^6 PBMCs). Red lines, geomean.

[Source Data](#)

[Extended Data Fig. 8 In vivo expansion of pre-existing SARS-CoV-2-reactive T cells post-infection or post-exposure.](#)

a, Change in magnitude of T cell response between pre-pandemic and post-infection (upper panel: all proteins, lower panel: NSP12) in seropositive close contacts of cases. **b**, Summary data for paired pre-pandemic and post-exposure NSP12 and Flu/EBV/CMV (FEC) responses in seronegative close contacts of infections. Below; example ELISpot well images from a seronegative close contact (NSP12-4: pre-pandemic 45 and post-exposure 95 SFCs per 10^6 PBMCs). **c**, Change in magnitude of T cell response between pre-pandemic and post-exposure samples (upper panel: all proteins, lower panel: NSP12) from seronegative close contacts of cases. **d**, Summary data for NSP12 and FEC responses in SN-HCW (sub-group with the top RTC response at wk16, n=19, Extended Data Fig. 4a). **e**, Summary

data and **f**, change in magnitude of T cell responses for individual HCW to NSP7 (15 peptide pool) or a single subpool from NSP12 and NSP13 between baseline and post-exposure in SN-HCW (wk16-26, 29 responses from 13 SN-HCW). **g**, Change in magnitude of T cell response to individual 9-15mer peptides pre- and post-exposure in SN-HCW (wk16-26, 11 responses from 9 SN-HCW). **h**, Example plots of CTV^{lo} IFN γ ⁺ SARS-CoV-2-specific T cells after 10day expansion of PBMC from baseline and wk16 with peptide #166 (YVYLPYPDPSRILGA) or unstimulated in an HLA-B*51+ SN-HCW (gated on CD8+, percentage of CD8+ shown, gating strategy Extended Data Fig. [3a](#)). **i**, Proportion of SN-HCW with NP1 + NP2-reactive T cells grouped by those with and without newly detected or expanded NSP12 responses at wk16, Fig. [4b](#). **j**, Correlation between the fold-change in NSP12 between recruitment and wk16 and total response to RTC or structural proteins at wk16 in SN-HCWs. Dotted line at 2-fold increase. **k**, The breadth of the NSP12-specific T cell response (number of subpools recognized, pre-pandemic or wk16). **l** Change in magnitude of the T cell response to NSP12 between baseline (open circles) and wk16 (closed circles) in SN-HCW and HCW with laboratory-confirmed infection. Percentage of responders shown below. **m**, Change in magnitude of NSP12-specific T cell response between pre-pandemic and post-infection in HCW with laboratory-confirmed infection. **a,c,f-g,m**, Expanded, >2-fold increase or >35 SFCs per 10⁶ PBMCs increase. **a**, Red line mean, **d-e,I**, Red line/bars, geometric mean. **k**, red line, median. **b,d,e** Wilcoxon test. **I**, Mann-Whitney (unpaired) and Wilcoxon (paired) tests. **k**, Kruskal-Wallis with Dunn's correction. **j**, Spearman r. **a-c**, Contact cohort, Extended Data Table [5](#). **d-m**, COVIDsortium cohort Extended Data Table [1](#).

[Source Data](#)

Extended Data Table 1 Cohort Demographics

Extended Data Table 2 T cell proliferation assay in seronegative HCW

Extended Data Table 3 SARS-CoV-2 Neis genetic diversity and number of homoplasies per site per gene region

Extended Data Table 4 Immunogenic peptides recognised by CD4+ or CD8+ T cells in seronegative HCW

Extended Data Table 5 Demographics and sampling of Close Contact medical student/laboratory staff cohort

Supplementary information

Reporting Summary

Supplementary Table 1

HCoV sequence accession numbers.

Supplementary Table 2

2D epitope mapping matrices.

Supplementary Table 3

NSP12 Overlapping peptide sequences and optimal 9-mers.

Peer Review File

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Source Data Extended Data Fig. 1

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 4

[Source Data Extended Data Fig. 5](#)

[Source Data Extended Data Fig. 7](#)

[Source Data Extended Data Fig. 8](#)

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Reprints and Permissions

About this article

Cite this article

Swadling, L., Diniz, M.O., Schmidt, N.M. *et al.* Pre-existing polymerase-specific T cells expand in abortive seronegative SARS-CoV-2. *Nature* **601**, 110–117 (2022). <https://doi.org/10.1038/s41586-021-04186-8>

- Received: 17 June 2021
- Accepted: 27 October 2021
- Published: 10 November 2021

- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04186-8>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

Further reading

- [**Cross-reactive memory T cells abort SARS-CoV-2 infection**](#)
 - Alexandra Flemming

Nature Reviews Immunology (2022)
- [**How do people resist COVID infections? Hospital workers offer a hint**](#)
 - Max Kozlov

Nature (2021)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04186-8>

- Article
- [Published: 15 December 2021](#)

Anatomically distinct fibroblast subsets determine skin autoimmune patterns

- [Zijian Xu^{1 na1}](#),
- [Daoming Chen^{1,2 na1}](#),
- [Yucheng Hu³](#),
- [Kaiju Jiang¹](#),
- [Huanwei Huang¹](#),
- [Yingxue Du¹](#),
- [Wenbo Wu¹](#),
- [Jiawen Wang¹](#),
- [Jianhua Sui](#) [ORCID: orcid.org/0000-0002-1272-9662¹](#),
- [Wenhui Wang⁴](#),
- [Long Zhang⁴](#),
- [Shuli Li⁵](#),
- [Chunying Li](#) [ORCID: orcid.org/0000-0003-3004-5376⁵](#),
- [Yong Yang](#) [ORCID: orcid.org/0000-0002-7668-3309⁶](#),
- [Jianmin Chang](#) [⁷](#) &
- [Ting Chen](#) [ORCID: orcid.org/0000-0002-7404-4538^{1,8}](#)

[Nature](#) volume **601**, pages 118–124 (2022)

- 9119 Accesses
- 40 Altmetric
- [Metrics details](#)

Subjects

- [Autoimmunity](#)
- [Pattern formation](#)

Abstract

The skin serves as a physical barrier and an immunological interface that protects the body from the external environment^{1,2,3}. Aberrant activation of immune cells can induce common skin autoimmune diseases such as vitiligo, which are often characterized by bilateral symmetric lesions in certain anatomic regions of the body^{4,5,6}. Understanding what orchestrates the activities of cutaneous immune cells at an organ level is necessary for the treatment of autoimmune diseases. Here we identify subsets of dermal fibroblasts that are responsible for driving patterned autoimmune activity, by using a robust mouse model of vitiligo that is based on the activation of endogenous auto-reactive CD8⁺ T cells that target epidermal melanocytes. Using a combination of single-cell analysis of skin samples from patients with vitiligo, cell-type-specific genetic knockouts and engraftment experiments, we find that among multiple interferon- γ (IFN γ)-responsive cell types in vitiligo-affected skin, dermal fibroblasts are uniquely required to recruit and activate CD8⁺ cytotoxic T cells through secreted chemokines. Anatomically distinct human dermal fibroblasts exhibit intrinsic differences in the expression of chemokines in response to IFN γ . In mouse models of vitiligo, regional IFN γ -resistant fibroblasts determine the autoimmune pattern of depigmentation in the skin. Our study identifies anatomically distinct fibroblasts with permissive or repressive IFN γ responses as the key determinant of body-level patterns of lesions in vitiligo, and highlights mesenchymal subpopulations as therapeutic targets for treating autoimmune diseases.

This is a preview of subscription content

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Fig. 1: scRNA-seq analysis reveals distinct IFN γ -responsive cell types in the skin of patients with vitiligo.

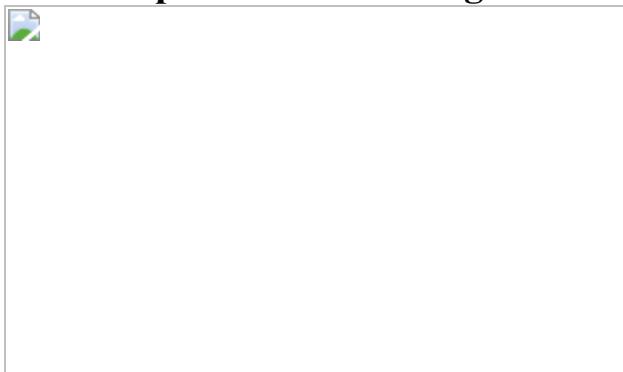


Fig. 2: Paracrine IFN γ signalling mediated by skin stromal cells drives CD8 $^+$ T cell cytotoxicity.



Fig. 3: IFN γ -responsive dermal fibroblasts are uniquely required to orchestrate autoimmune CD8 $^+$ T cells through secreted chemokines.

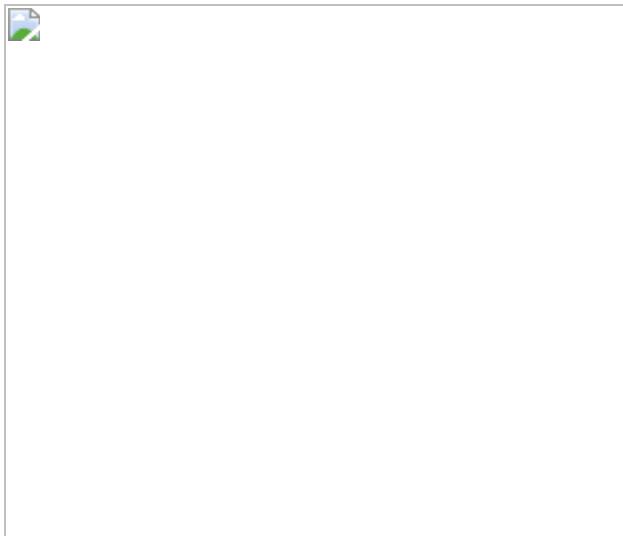


Fig. 4: Anatomically distinct fibroblast subsets with differential IFN γ responses determine organ-level skin autoimmune patterns.



Data availability

The scRNA-seq data, mouse fibroblast RNA-seq data and human fibroblast RNA-seq data have been deposited in the Genome Sequence Archive (GSA) with accession number [PRJCA006797](#). The code for the vitiligo progression model is available at <https://github.com/hydrays/vitiligo>. [Source data](#) are provided with this paper.

References

1. 1.
Larsen, S. B., Cowley, C. J. & Fuchs, E. Epithelial cells: liaisons of immunity. *Curr. Opin. Immunol.* **62**, 45–53 (2019).
2. 2.
Rognoni, E. & Watt, F. M. Skin cell heterogeneity in development, wound healing, and cancer. *Trends Cell Biol.* **28**, 709–722 (2018).
3. 3.
Afshar, M. & Gallo, R. L. Innate immune defense system of the skin. *Vet. Dermatol.* **24**, 32–38 (2013).
4. 4.
Taieb, A. & Picardo, M. Vitiligo. *N. Engl. J. Med.* **360**, 160–169 (2009).
5. 5.
Alikhan, A., Felsten, L. M., Daly, M. & Petronic-Rosic, V. Vitiligo: a comprehensive overview Part I. Introduction, epidemiology, quality of life, diagnosis, differential diagnosis, associations, histopathology, etiology, and work-up. *J Am. Acad. Dermatol.* **65**, 473–491 (2011).

6. 6.

Picardo, M. et al. Vitiligo. *Nat. Rev. Dis. Primers* **1**, 15011 (2015).

7. 7.

Ezzedine, K., Eleftheriadou, V., Whitton, M. & van Geel, N. Vitiligo. *Lancet* **386**, 74–84 (2015).

8. 8.

Taieb, A. Intrinsic and extrinsic pathomechanisms in vitiligo. *Pigment Cell Res.* **13**, 41–47 (2000).

9. 9.

Spritz, R. A. & Santorico, S. A. The genetic basis of vitiligo. *J. Invest. Dermatol.* **141**, 265–273 (2021).

10. 10.

Frisoli, M. L., Essien, K. & Harris, J. E. Vitiligo: mechanisms of pathogenesis and treatment. *Annu. Rev. Immunol.* **38**, 621–648 (2020).

11. 11.

Le Poole, I. C., van den Wijngaard, R. M., Westerhof, W. & Das, P. K. Presence of T cells and macrophages in inflammatory vitiligo skin parallels melanocyte disappearance. *Am. J. Pathol.* **148**, 1219–1228 (1996).

12. 12.

Ogg, G. S., Rod Dunbar, P., Romero, P., Chen, J. L. & Cerundolo, V. High frequency of skin-homing melanocyte-specific cytotoxic T lymphocytes in autoimmune vitiligo. *J. Exp. Med.* **188**, 1203–1208 (1998).

13. 13.

Van den Boorn, J. G. et al. Autoimmune destruction of skin melanocytes by perilesional T cells from vitiligo patients. *J. Invest. Dermatol.* **129**, 2220–2232 (2009).

14. 14.

Maeda, Y. et al. Detection of self-reactive CD8⁺ T cells with an anergic phenotype in healthy individuals. *Science* **346**, 1536–1540 (2014).

15. 15.

Cheuk, S. et al. CD49a expression defines tissue-resident CD8⁺ T cells poised for cytotoxic function in human skin. *Immunity* **46**, 287–300 (2017).

16. 16.

Iannella, G. et al. Vitiligo: pathogenesis, clinical variants and treatment approaches. *Autoimmun. Rev.* **15**, 335–343 (2016).

17. 17.

Rodrigues, M. et al. Current and emerging treatments for vitiligo. *J. Am. Acad. Dermatol.* **77**, 17–29 (2017).

18. 18.

Kovacs, S. O. Vitiligo. *J. Am. Acad. Dermatol.* **38**, 647–666 (1998).

19. 19.

Gregg, R. K., Nichols, L., Chen, Y., Lu, B. & Engelhard, V. H. Mechanisms of spatial and temporal development of autoimmune vitiligo in tyrosinase-specific TCR transgenic mice. *J. Immunol.* **184**, 1909–1917 (2010).

20. 20.

Harris, J. E. et al. A mouse model of vitiligo with focused epidermal depigmentation requires IFN-gamma for autoreactive CD8⁺ T-cell accumulation in the skin. *J. Invest. Dermatol.* **132**, 1869–1876 (2012).

21. 21.

Chatterjee, S. et al. A quantitative increase in regulatory T cells controls development of vitiligo. *J. Invest. Dermatol.* **134**, 1285–1294, (2014).

22. 22.

Yang, L. et al. Interferon-gamma inhibits melanogenesis and induces apoptosis in melanocytes: a pivotal role of CD8⁺ cytotoxic T lymphocytes in vitiligo. *Acta Derm. Venereol.* **95**, 664–670 (2015).

23. 23.

Wang, S. Q. et al. Interferon-γ induces senescence in normal human melanocytes. *PLoS One* **9**, e93232 (2014).

24. 24.

Tulic, M. K. et al. Innate lymphocyte-induced CXCR3B-mediated melanocyte apoptosis is a potential initiator of T-cell autoreactivity in vitiligo. *Nat. Commun.* **10**, 2178 (2019).

25. 25.

Richmond, J. M. et al. Keratinocyte-derived chemokines orchestrate T-cell positioning in the epidermis during vitiligo and may serve as biomarkers of disease. *J. Invest. Dermatol.* **137**, 350–358 (2017).

26. 26.

Biswas, K. B. et al. GPNMB is expressed in human epidermal keratinocytes but disappears in the vitiligo lesional skin. *Sci. Rep.* **10**, 4930 (2020).

27. 27.

Boukhedouni, N. et al. Type-1 cytokines regulate matrix metalloprotease-9 production and E-cadherin disruption to promote melanocyte loss in vitiligo. *JCI Insight* **4**, e133772 (2020).

28. 28.

Ganju, P. et al. Microbial community profiling shows dysbiosis in the lesional skin of vitiligo subjects. *Sci Rep.* **6**, 18761 (2016).

29. 29.

Yuan, X. et al. The impact of NBUVB on microbial community profiling in the lesional skin of vitiligo subjects. *Microb. Pathog.* **140**, 103943 (2020).

30. 30.

Lang, K. S. et al. HLA-A2 restricted, melanocyte-specific CD8⁺ T lymphocytes detected in vitiligo patients are related to disease activity and are predominantly directed against MelanA/MART1. *J. Invest. Dermatol.* **116**, 891–897 (2001).

31. 31.

Namazi, M. R. Neurogenic dysregulation, oxidative stress, autoimmunity, and melanocytorrhagy in vitiligo: can they be interconnected? *Pigment Cell Res.* **20**, 360–363 (2007).

32. 32.

Lotti, T., D'Erme, A. M. & Hercogova, J. The role of neuropeptides in the control of regional immunity. *Clin. Dermatol.* **32**, 633–645 (2014).

33. 33.

Palermo, B. et al. Specific cytotoxic T lymphocyte responses against Melan-A/MART1, tyrosinase and gp100 in vitiligo by the use of major

histocompatibility complex/peptide tetramers: the role of cellular immunity in the etiopathogenesis of vitiligo. *J. Invest. Dermatol.* **117**, 326–332 (2001).

34. 34.

Mantovani, S. et al. Dominant TCR- α requirements for a self antigen recognition in humans. *J. Immunol.* **169**, 6253–6260 (2002).

35. 35.

Byrne, K. T., Zhang, P., Steinberg, S. M. & Turk, M. Autoimmune vitiligo does not require the ongoing priming of naive CD8 T cells for disease progression or associated protection against melanoma. *J. Immunol.* **192**, 1433–1439 (2014).

36. 36.

Mosenson, J. A. et al. Mutant HSP70 reverses autoimmune depigmentation in vitiligo. *Sci. Transl. Med.* **5**, 174ra128 (2013).

37. 37.

Miao, X. et al. PD-L1 reverses depigmentation in Pmel-1 vitiligo mice by increasing the abundance of Tregs in the skin. *Sci Rep.* **8**, 1605 (2018).

38. 38.

Riding, R. L., Richmond, J. M. & Harris, J. E. Mouse model for human vitiligo. *Curr. Protoc. Immunol.* **124**, e63 (2019).

39. 39.

Richmond, J. M. et al. Resident memory and recirculating memory T cells cooperate to maintain disease in a mouse model of vitiligo. *J. Invest. Dermatol.* **139**, 769–778 (2019).

40. 40.

Rashighi, M. et al. CXCL10 is critical for the progression and maintenance of depigmentation in a mouse model of vitiligo. *Sci. Transl. Med.* **6**, 223ra223 (2014).

41. 41.

Yu, Z. et al. Hoxc-dependent mesenchymal niche heterogeneity drives regional hair follicle regeneration. *Cell Stem Cell* **23**, 487–500 (2018).

42. 42.

Groom, J. R. & Luster, A. D. CXCR3 in T cell function. *Exp. Cell. Res.* **317**, 620–631 (2011).

43. 43.

Chang, H. Y. et al. Diversity, topographic differentiation, and positional memory in human fibroblasts. *Proc. Natl Acad. Sci. USA.* **99**, 12877–12882 (2002).

44. 44.

Paus, R., Ito, N., Takigawa, M. & Ito, T. The hair follicle and immune privilege. *J. Investig. Dermatol. Symp. Proc.* **8**, 188–194 (2003).

45. 45.

Farber, E. M., Nickoloff, B. J., Recht, B. & Fraki, J. E. Stress, symmetry, and psoriasis: possible role of neuropeptides. *J. Am. Acad. Dermatol.* **14**, 305–311 (1986).

46. 46.

Boehncke, W.-H. & Schön, M. P. Psoriasis. *Lancet* **386**, 983–994 (2015).

47. 47.

Greb, J. E. et al. Psoriasis. *Nat. Rev. Dis. Primers* **2**, 16082 (2016).

48. 48.

Gabrielli, A., Avvedimento, E. V. & Krieg, T. Scleroderma. *N. Engl. J. Med.* **360**, 1989–2003, (2009).

49. 49.

Kaul, A. et al. Systemic lupus erythematosus. *Nat. Rev. Dis. Primers* **2**, 16039 (2016).

50. 50.

Boyd, A. S. & Neldner, K. H. Lichen planus. *J. Am. Acad. Dermatol.* **25**, 593–619 (1991).

51. 51.

Tang, Y., Harrington, A., Yang, X., Friesel, R. E. & Liaw, L. The contribution of the Tie2⁺ lineage to primitive and definitive hematopoietic cells. *Genesis* **48**, 563–567 (2010).

52. 52.

Deng, L. et al. A novel mouse model of inflammatory bowel disease links mammalian target of rapamycin-dependent hyperproliferation of colonic epithelium to inflammation-associated tumorigenesis. *Am. J. Pathol.* **176**, 952–967 (2010).

53. 53.

McCubbrey, A. L., Allison, K. C., Lee-Sherick, A. B., Jakubzick, C. V. & Janssen, W. J. Promoter specificity and efficacy in conditional and inducible transgenic targeting of lung macrophages. *Front. Immunol.* **8**, 1618 (2017).

54. 54.

Le Gal, F. A. et al. Direct evidence to support the role of antigen-specific CD8⁺ T cells in melanoma-associated vitiligo. *J. Invest.*

Dermatol. **117**, 1464–1470 (2001).

55. 55.

Lo, J. A. et al. Epitope spreading toward wild-type melanocyte-lineage antigens rescues suboptimal immune checkpoint blockade responses. *Sci. Transl. Med.* **13**, eabd8636 (2021).

56. 56.

Hua, C. et al. Association of vitiligo with tumor response in patients with metastatic melanoma treated with pembrolizumab. *JAMA Dermatol.* **152**, 45–51 (2016).

57. 57.

Herzberg, B. & Fisher, D. E. Metastatic melanoma and immunotherapy. *Clin. Immunol.* **172**, 105–110 (2016).

58. 58.

Nardin, C. et al. Vitiligo under anti-PD1 therapy is associated with increased survival in melanoma patients. *J. Am. Acad. Dermatol.* **82**, 770–772 (2020).

59. 59.

Halprin, K. M. Epidermal "turnover time"—a re-examination. *Br. J. Dermatol.* **86**, 14–19 (1972).

60. 60.

Gellatly, K. J. et al. scRNA-seq of human vitiligo reveals complex networks of subclinical immune activation and a role for CCR5 in T_{reg} function. *Sci. Transl. Med.* **13**, eabd8995 (2021).

61. 61.

Chen, T. et al. An RNA interference screen uncovers a new molecule in stem cell self-renewal and long-term regeneration. *Nature* **485**, 104–108 (2012).

62. 62.

Lander, A. D. Morpheus unbound: reimagining the morphogen gradient. *Cell* **128**, 245–256 (2007).

63. 63.

Chen, C. C. et al. Organ-level quorum sensing directs regeneration in hair stem cell populations. *Cell* **161**, 277–290 (2015).

Acknowledgements

This work was supported by grants from the National Key R&D Program of China (2017YFA0103500), the National Basic Research Program of China 973 Programs (2014CB849602), National Natural Science Foundation of China (11671415) and the Beijing Municipal Natural Science Foundation (7172192). We thank the NIBS Animal Facility for their expert handling and care of mice; the NIBS Biological Resource Centre for FACS; the NIBS imaging facility for assistance with the confocal microscope experiment; the NIBS sequencing centre for sample processing of 10x genomics RNA-seq; V. Hearing for providing the DCT and TYR peptide sequences for in-house-made antibodies; F. Shao for help with importing *Ifngr1* KO mice; and all members of the laboratory of T.C. for discussions and technical support.

Author information

Author notes

1. These authors contributed equally: Zijian Xu, Daoming Chen

Affiliations

1. National Institute of Biological Sciences, Beijing, China

Zijian Xu, Daoming Chen, Kaiju Jiang, Huanwei Huang, Yingxue Du, Wenbo Wu, Jiawen Wang, Jianhua Sui & Ting Chen

2. Peking University–Tsinghua University–National Institute of Biological Sciences Joint Graduate Program, School of Life Sciences, Peking University, Beijing, China

Daoming Chen

3. Academy for Multidisciplinary Studies, Beijing National Center for Applied Mathematics, Beijing Advanced Innovation Center for Imaging Theory and Technology, Capital Normal University, Beijing, China

Yucheng Hu

4. Peking University Third Hospital, Beijing, China

Wenhui Wang & Long Zhang

5. Department of Dermatology, Xijing Hospital, Xi'an, China

Shuli Li & Chunying Li

6. Institute of Dermatology, Chinese Academy of Medical Sciences and Peking Union Medical College, Nanjing, China

Yong Yang

7. Department of Dermatology, Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China

Jianmin Chang

8. Tsinghua Institute of Multidisciplinary Biomedical Research, Tsinghua University, Beijing, China

Ting Chen

Contributions

T.C. and Z.X. conceived the project, designed the experiments and wrote the manuscript. Z.X. and D.C. performed most of the *in vivo* experiments. Y.D. participated in the lentivirus injection experiment. K.J. performed some of the immunofluorescent staining experiment. Z.X., H.H. and J.W. conducted bioinformatics analysis of single-cell sequencing data. W. Wu collected fibroblasts from different body positions and participated in the cell culture experiment. D.C., H.H., J.W. and Z.X. conducted the bioinformatics analysis of the RNA-seq data. Y.H. developed the mathematical model. L.Z., W. Wang, Y.Y. and J.C. collected samples from patients with vitiligo and healthy donors. S.L., C.L. and J.C. collected clinical data from patients with vitiligo and carried out statistical analysis. J.S. provided the CD8 depletion antibody in the CD8⁺ T cell depletion experiment.

Corresponding authors

Correspondence to [Jianmin Chang](#) or [Ting Chen](#).

Ethics declarations

Competing interests

All authors declare no competing interests related to this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Comprehensive analysis of different cell populations from patients with vitiligo and healthy donors.

a, Representative images of the skin of patients with vitiligo showing characteristic bilateral symmetric lesion patterns in different body regions. **b**, Experimental workflow and representative FACS profiles to obtain single cells from human skin biopsies for single cell RNA-seq. **c**, Clinical information for all donors with the number of cells sequenced in single-cell analysis. **d**, t-SNE visualization of all collected cells showing eight main cell types with distinct clusters. **e**, Heat map analysis of differentially expressed genes in each cell type. **f–h**, violin plot (**f**), feature plot (**g**), and dot plot (**h**) analysis of signature genes for each cell type. **i**, Volcano plot and GO analysis of genes enriched in melanocytes from the M1 cluster compared to M2 cluster. **j**, T-SNE projection of T cells from patients with vitiligo and healthy donors. **k**, Volcano plot showing genes differentially expressed (>2-fold, t test $p < 0.01$) in each T cell sub-cluster. **l**, **m**, Representative immunofluorescent images (**l**) and quantification (**m**) of T cell subtype markers GZMA, ZNF683, CD40LG, and FOXP3 in T cells (CD3+) in the skin of patients with vitiligo. **n**, GO categories of genes enriched in each cell type from the skin of patients in the progressive state compared to healthy donor skin. **o**, Heat map analysis of 20 IFN γ response genes in eight cell types from the skin of patients in the progressive state and healthy donor skin. **p**, Violin plot showing the expression of *IFNG* in T cell subtypes. **q**, Violin plots showing *IFNGR1* and *IFNGR2* expression in different skin cell types. **r**, Representative immunofluorescent staining images, quantification (left and middle), and regression analysis (right) between density of CD8+ T cell and pSTAT1+ cell in the skin of patients with vitiligo. **s**, Representative immunofluorescent staining images of pSTAT1 signal in each cell type in the skin of patients with vitiligo. Scale bars, 50 μ m (**l**, **r**, **s**). For exact p values, see Source Data. For statistics, p summary and sample sizes, see [Methods](#).

[Source data](#)

Extended Data Fig. 2 Characterization of the melanoma-T_{reg}-induced vitiligo mouse model.

a, Schematic diagram of the melanoma/Treg-induced vitiligo mouse model. First the dorsal skin in the right flank of 8–9-week old C57 mice was inoculated with B16F10 melanoma cells (Day 0), then CD4 depletion antibodies were injected on Days 4 and 10. The tumour was surgically removed on Day 12 to prevent interference of subsequent analysis. **b**, Representative hair coat images of control and vitiligo induced mice at Day 300 after the vitiligo induction procedure. At 4 weeks after induction, dorsal skin hair follicles close to the B16F10 tumour cell injection and surgical removal sites started to show depigmentation as a result of wounding-induced new hair growth. Then the depigmented hair follicles expanded and eventually rendered the whole dorsal hair coat depigmented at ~Day 300. **c**, Representative tail skin images of mice at Day 0, 45, 60, 90, 120, 240, 360, 480, 720 post melanoma/Treg-induced vitiligo induction procedure. Melanocytes in mouse dorsal skin are located in hair follicles but not in the epidermis; only in mouse tail skin are melanocytes located in both the hair follicle and epidermis similar to human skin. Skin epidermis depigmentation is the defining feature of vitiligo pathology. Therefore, we only used mouse tail skin for vitiligo analysis. In tail skin, the depigmentation was initially patchy and then progressed to eventually cover the entire epidermal surface, which did not recover even more than a year later. **d**, Representative whole-mount immunofluorescent staining and density plot images of DCT+ melanocytes and CD8+ T cells in mouse tail skin epidermis at Day 0, 19, 26, 33 after the vitiligo induction procedure. Prior to vitiligo induction, very few if any CD8+ T cells could be detected in epidermis. Starting from Day 19 after the vitiligo induction procedure, CD8+ T cells infiltration and small regions of melanocyte loss could be observed. Note, melanocytes loss only occurred in regions where CD8+ T cells locally aggregated into clusters; as the CD8+ T cells clusters continuously expanded so did the regions of melanocyte loss correspondingly. **e**, FACS quantifications of CD45+CD3+CD8+ T cells and CD117+ melanocyte in mice with or without vitiligo induction confirmed the loss of melanocytes and enrichment of CD8+ T cells after vitiligo induction. **f**, Representative whole-mount immunofluorescent staining images and quantifications of DCT+ melanocytes and CD8+ T cells in mouse tail skin epidermis with combined or individual B16F10 inoculation and CD4 depletion antibody treatment, showing the melanoma/Treg-induced vitiligo model requires both B16F10 inoculation and CD4

depletion antibody injection. **g**, Schematic diagram, representative whole-mount immunofluorescent staining images and quantification of DCT+ melanocytes and CD8+ T cells in the skin epidermis of melanoma/Treg-induced vitiligo model with or without CD8 depletion antibody treatment, showing CD8+ T cells are responsible for melanocyte loss in the melanoma/Treg-induced vitiligo mouse model. **h, i**, Representative whole-mount images (**h**) and FACS profiles (**i**) with corresponding quantification of the percentage of CD8+ T cells that express CD3 or CD11c in mouse epidermis, showing majority of the CD8+ cells are CD3+ T cells, but not CD11c+ dendritic cells. **j**, Representative whole-mount immunofluorescent staining images and density plot images of CD8+ T cells and DCT+ melanocytes in tail skin epidermis at Day 19, 26, 33 after vitiligo induction procedure. Note the continuously expanding CD8+ T cell clusters in skin epidermis. **k**, Representative whole-mount immunofluorescent staining images and quantifications of Ki67+, CD3+ T cells, and DCT+ melanocytes in tail skin epidermis at Day 33 after vitiligo induction procedure. Enlarged image on the left represents border of lesion skin, the right one represents lesion region. Quantification showed the proliferation rates of CD8+ T cells at the border versus inside the clusters are equivalent; and percentage of Ki67+ cells in melanocytes, T cells and keratinocytes showed majority of the proliferating cells in skin epidermis are keratinocytes. These data indicated that the continuous expansion of CD8+ T cell clusters mainly results from skin-infiltrated CD8+ T cells being actively recruited into the border regions. Scale bars, 500 μm (**c, d, f, g, j**), 100 μm (**h, k**). For exact p values, see Source Data. For statistics, p summary and sample sizes, see [Methods](#).

Source data

[Extended Data Fig. 3 The melanoma–T_{reg}-induced vitiligo mouse model recapitulates hallmarks of human vitiligo.](#)

a, UMAP projection of skin immune cells from patients with vitiligo and healthy donors. We identified 5 major clusters in immune cells, including two adaptive immune cell clusters (CD4+ T cell and CD8+ T cell) and three innate immune cell clusters (Langerhans cell, macrophage, dendritic cell 1/2). Other immune cells of very low abundance in human skin, such as

mast cells and $\gamma\delta$ T cells, could also be observed in our single-cell RNA-seq data, but owing to their limited number these cell types were not analysed individually. **b–e**, Heat map (**b**), dot plot (**c**), violin plot (**d**) and feature plot (**e**) analysis of signature genes in each immune cell subtype. **f**, Quantification and comparison of five immune cell types between progressive-state patients with vitiligo and healthy donors by scRNA-seq analysis. **g**, Representative FACS profiles analysing immune cell types in mice vitiligo skin, including epidermal TCR $\gamma\delta$ ^{high} and TCR $\gamma\delta$ ^{low} $\gamma\delta$ T cell present in mouse. **h, i**, Quantification and comparison of immune cell types between vitiligo (at Day 33) and control mice. Data in f-g showed the predominantly enriched immune cell type is CD8+ T cell in both patients with vitiligo and the melanoma/Treg-induced vitiligo mouse model, whereas the other immune cell types did not show significant difference. **j**, Representative whole-mount immunofluorescent staining images of CD4+ T cells in WT skin, and Tregs in *Foxp3*^{Cre}; *Ai14* skin, indicating at Day 33 of vitiligo analysis, the CD4+ T cell population have recovered from systematical CD4+ T cell depletion. **k**, Heat map analysis of cytotoxicity-related genes expressed in skin CD8+ T cells of patients with vitiligo and vitiligo-induced mice, indicating that in molecular level, this melanoma/Treg-induced vitiligo mouse model elicits endogenous auto-reactive CD8+ cytotoxic T cells similar to human vitiligo. Ctl: control, Vit: vitiligo. Scale bars, 200 μ m (**k**). For exact p values, see Source Data. For statistics, p summary and sample sizes, see [Methods](#).

[Source data](#)

[**Extended Data Fig. 4 Analysis of wild-type and *Ifngr1* KO mice with three different vitiligo induction methods.**](#)

a–j, Analysis of WT and *Ifngr1* KO mice with melanoma/Treg-induced vitiligo model. **a**, Schematic diagram. **b**, Representative tail skin images of WT and *Ifngr1* KO mice at Day 60 after vitiligo induction. **c, d**, Representative whole-mount images (**c**) and quantification (**d**) of tail skin epidermal melanocytes and CD8+ T cells at Day 33 after vitiligo induction. **e, f**, Representative FACS profile of CD117+ melanocytes and CD3+CD8+ T cells (**e**), and quantification of CD117+ melanocytes (**f**) in the tail skin epidermis at Day 33 after vitiligo induction. **g**, Whole-mount views of

scales in tail skin epidermis (dotted line) in control and vitiligo mice. This is the unit area we used to quantify melanocyte or CD8+ T cell density in wholemount analysis throughout this paper. **h**, Scatter plots and correlation analysis of melanocyte number versus CD8+ T cell number in each scale of WT and *Ifngr1* KO tail skin at Day 33 after vitiligo induction. **i**, Representative whole-mount images of DCT+ melanocytes in tail skin epidermis of WT and *Ifngr1* KO mice at 4 months after vitiligo induction. **j**, QPCR analysis of *GZMA*, *GZMB*, *PRF1*, and *CCL5* expression in skin CD8+ T cells isolated from WT and *Ifngr1* KO mice at Day 33 after vitiligo induction. **k–p**, Analysis of WT and *Ifngr1* KO mice with *Pmel* transgenic spontaneous vitiligo mouse model. **k**, Schematic diagram. **l–m**, Representative tail skin images (**l**, P70) and whole-mount immunofluorescent staining images (**m**, P42) of *Pmel*;WT and *Pmel*; *Ifngr1* KO mice. **n**, **o**, Representative FACS profiles (**n**) and quantification (**o**) of CD117+ melanocytes and CD45+CD3+CD8+ T cells in tail skin epidermis of WT, *Ifngr1* KO, *Pmel*;WT and *Pmel*; *Ifngr1* KO mice at P42. **p**, Scatter plots and correlation analysis of melanocyte number versus CD8+ T cell number in each scale in *Pmel*;WT and *Pmel*; *Ifngr1* KO tail skin at P42. **q–x**, Analysis of WT and *Ifngr1* KO mice with adoptive T cell transfer-based vitiligo mouse model. **q**, Schematic diagram. **r**, Quantification of CD117+ melanocytes and CD45+CD3+CD8+ T cells in tail skin epidermis of WT mice after γ irradiation, with or without WT *Pmel* T cell transfer or AVV-hPMEL intraperitoneal injection. These data indicate sub-lethal irradiation alone, irradiation with hPMEL-AAV alone, or irradiation with *Pmel* CD8+ T cell transfer alone did not result in melanocyte loss or CD8+ T cell infiltration in WT tail skin epidermis. **s**, Representative immunofluorescent staining images and FACS quantification of epidermal melanocyte of mice post γ -irradiation 26 days with controls, indicating after 5 Gy γ irradiation leads to ~2-fold increase of melanocyte number in tail skin epidermis compared to untreated mice. **t**, Representative FACS profiles and quantification of CD3+CD8+VB13+ *Pmel* T cells in spleen or skin of WT and *Ifngr1* KO mice at Day 26 with or without adoptive transfer-based vitiligo model induction. These data showed although both WT and *Ifngr1* KO host mice contained the same number of transferred WT TCR VB13+ *Pmel* T cells in the spleen, only WT host mice exhibit robust VB13+ *Pmel* T cell infiltration in skin, whereas the *Ifngr1* KO host mice did not. **u–v**, Representative tail skin images (**u**, Day 60) and whole-mount

immunofluorescent images (**v**, Day 26) of WT and *Ifngr1* KO mice post vitiligo model-induction. Scale bars, 500 μm. **w**, Representative FACS profiles and quantification of epidermal CD117+ melanocytes in WT and *Ifngr1* KO mice at Day 26 post vitiligo model-induction. **x**, Scatter plots and correlation analysis of melanocyte number versus CD8+ T cell number in each scale in WT and *Ifngr1* KO mice at Day 26 post vitiligo model-induction. Scale bars, 500 μm (**b**, **c**, **g**, **i**, **l**, **m**, **s**, **u**, **v**). For exact p values, see Source Data. For statistics, p summary and sample sizes, see [Methods](#).

[Source data](#)

[Extended Data Fig. 5 IFN \$\gamma\$ -responsive skin stromal cells are required for vitiligo progression.](#)

a–c, Analysis of six main immune cell types in tail skin of WT and *Ifngr1* KO mice with different vitiligo induction methods. **a**, Day 33 after melanoma/Treg-induced vitiligo model induction. **b**, P42 of *Pmel*;WT and *Pmel;Ifngr1* KO mice. **c**, Day 26 after adoptive transfer-based vitiligo model-induction. **d**, Timeline of the graft and vitiligo induction assay. Representative whole-mount immunofluorescent staining images of grafted WT tail skin epidermis on WT host at Day 0 and Day 21 with or without vitiligo induction were showed. These data indicate tail skin did not spontaneously develop vitiligo after the graft procedure alone; only after vitiligo induction in the host mice did the grafted skin develop vitiligo as indicated by melanocyte loss and large amounts of CD8+ T cells infiltration. **e**, **f**, Lineage tracing experiment of CD8+ T cells: **e**, Schematic diagram WT tail skin graft on *CD4*^{Cre};*mTmG* host and representative immunofluorescent staining images of CD8+ T cells in the grafted skin epidermis after vitiligo induction; **f**, FACS analysis and quantification of spleen cells from *CD4*^{Cre};*Ai6* mouse to determine GFP labelling efficiency in CD8+ T cells. Cells were pre-gated on CD45+ live singlets, indicating that the infiltrated CD8+ T cells were derived from the host mice rather than the grafted skin. These data indicate the graft infiltrated CD8+ T cells were derived from the host mice rather than the grafted skin. **g**, Quantification of melanocytes and CD8+ T cells number in grafted skin without vitiligo model induction in host mice. Donor skin pairs grafted onto the same host mouse are linked by lines. **h**, Representative

immunofluorescent staining of the junction region between grafted C57 tail skin and host dorsal skin of a *membrane-Tomato* (*mT*) transgenic mouse. After full-thickness C57 tail skin was grafted onto membrane-*tdTomato* (*mT*)-expressing host mice (*Rosa-mT*), in which all host cells are genetically labelled as *mT*+, host derived dermal cells indeed invaded into the grafted skin dermis. But neither K14+ keratinocytes nor DCT+ melanocytes migrated from the host to grafted skin. **i**, Representative immunofluorescent staining image of pSTAT1 signal in immune cell (CD45+), smooth muscle cell (a-SMA+), Langerhans cells (Langerin+) of grafted *Ifngr1* KO tail skin on WT host after vitiligo induction on a host mouse. Enlarged image on the right represents Langerin+ cells in the epidermis. Scale bars, 50 μ m (**d, e, h, i**). For exact p values, see Source Data. For statistics, p summary and sample sizes, see [Methods](#).

[Source data](#)

[**Extended Data Fig. 6 IFN \$\gamma\$ -responsive skin fibroblasts are required for driving vitiligo pathogenesis.**](#)

a, Schematic diagram and representative FACS profiles of skin cells isolation. Detailed gating strategies are described in method section. QPCR analyses of cell-type-specific signature genes in FACS-isolated populations include *KRT14* for keratinocytes, *DCT* for melanocytes, *CD45* for immune cells, *PDGFRA* for fibroblasts, and *CD31* for endothelial cells respectively. QPCR validation of *IFNGR1* knockout used FACS-purified endothelial cells, keratinocytes, immune cells, melanocytes, fibroblasts and CD8+ T cells from *Tek*^{Cre}; *IFNGR1*^{f/f}, *K14*^{Cre}; *IFNGR1*^{f/f}, *Csf1r*^{Cre}; *IFNGR1*^{f/f}, *Tyr*^{CreER}; *IFNGR1*^{f/f}, *CD4*^{Cre}; *IFNGR1*^{f/f}, and *Pdgfra*^{CreER}; *IFNGR1*^{f/f} mice, all with WT mice as controls. **b**, Representative whole-mount immunofluorescent staining images of melanocytes and CD8+ T cells in control and six cell-type-specific conditional knockout lines at Day 33 after vitiligo induction. **c**, Quantification of skin CD8+ T cells in WT and six cell-type-specific knockout lines at Day 33 after vitiligo induction procedure based on wholmount staining. **d, e**, Representative FACS profiles (**d**) and quantification (**e**) of CD117+ epidermal melanocytes and CD3+CD8+ T cells in control and six cell-type-specific conditional knockout lines post vitiligo induction. **f**, Schematic diagram of

melanoma/Treg-induced vitiligo procedure and representative tail skin images (Day 60) of WT and $Pdgfra^{CreER};IFNGR1^{fl/fl}$ cKO mice after vitiligo induction. Vit. induc.: Vitiligo induction. Scale bars, 500 μ m (**b, f**). For exact p values, see Source Data. For statistics, p summary and sample sizes, see [Methods](#).

[Source data](#)

Extended Data Fig. 7 IFN γ -responsive fibroblasts are both necessary and sufficient to mediate local recruitment of CD8 $^+$ T cells through secreted chemokines.

a–c, Analysis of WT and $Pdgfra^{CreER};IFNGR1^{fl/fl}$ cKO mice with *Pmel* transgenic spontaneous vitiligo mouse model. **a**, Schematic diagram and representative tail skin images (P70) and representative wholemount images (P42) of *Pmel*;WT and *Pmel*; $Pdgfra^{CreER};IFNGR1^{fl/fl}$ cKO mice. **b**, Scatter plots and correlation analysis of melanocyte number versus CD8+ T cell number in each scale from *Pmel*;WT and *Pmel*; $Pdgfra^{CreER};IFNGR1^{fl/fl}$ cKO mice at P42. **c**, Representative FACS profiles and quantification of tail skin epidermal CD117+ melanocytes and CD3+CD8+ T cells in WT, $Pdgfra^{CreER};IFNGR1^{fl/fl}$ cKO, *Pmel*;WT and *Pmel*; $Pdgfra^{CreER};IFNGR1^{fl/fl}$ cKO mice at P42. **d–g**, Analysis of WT and $Pdgfra^{CreER};IFNGR1^{fl/fl}$ cKO mice with adoptive transfer-based vitiligo model. **d**, Representative FACS profiles and quantification of CD3+CD8+VB13+ *Pmel* T cells in spleen or tail skin of WT and $Pdgfra^{CreER};IFNGR1^{fl/fl}$ cKO mice at Day 26 with or without adoptive transfer-based vitiligo model induction. **e**, Schematic diagram and representative tail skin images (Day 60) and representative wholemount immunofluorescent staining images (Day 26) of WT and $Pdgfra^{CreER};IFNGR1^{fl/fl}$ cKO mice after vitiligo induction. **f**, Scatter plots and correlation analysis of melanocyte number versus CD8+ T cell number in each scale WT and $Pdgfra^{CreER};IFNGR1^{fl/fl}$ cKO mice at Day 26 after vitiligo induction. **g**, Representative FACS profiles and quantification of tail epidermal CD117+ melanocytes, and quantification of tail epidermal CD117+ melanocytes and CD3+CD8+ T cells of WT and $Pdgfra^{CreER};IFNGR1^{fl/fl}$ cKO mice at Day 26. **h**, Section immunofluorescent analysis of

intradermally injected RFP-labelled fibroblast. Note injected RFP+ fibroblasts located in lower dermis. **i**, Quantification of T cell transwell migration assay showing T cells migration ratio using control medium without fibroblast, corresponding to Fig. 3d. **j**, Analysis of human and mouse vitiligo fibroblasts up regulated genes. Left, heat map showing differential expressed genes in tail skin fibroblasts from WT control, WT vitiligo-induced, and *Ifngr1* KO vitiligo-induced mice. Middle, volcano plot of differentially expressed genes in human fibroblasts from progressive-state patients with vitiligo compared with healthy donors. Red dots denote genes >1.5 fold upregulated (*t* test $p < 0.01$) in fibroblasts from progressive-state patients with vitiligo. Right, heat map analysis of the 28 common upregulated secreted factors in fibroblasts from patients with vitiligo and mice. **k**, Quantification of migration ratio from the T cell transwell migration assay with various chemokines. **l**, Representative immunofluorescent staining images of CXCL9 and CXCL10 signal with T cells (CD3+) in the skin of patients with vitiligo. **m**, **n**, Representative immunofluorescent staining images of CXCL9 (**m**) and CXCL10 (**n**) signals in melanocytes (DCT+), fibroblasts (PDGFRA+), endothelial cells (CD31+), smooth muscle cells (a-SMA+), keratinocytes (K14+), mononuclear phagocytes (CD11c+), Langerhans cells (Langerin+), and T cells (CD3+) in the skin of patients with vitiligo. Scale bars, 500 μm (**a**, **e**, **l**), 50 μm (**h**), 100 μm (**m**, **n**). For exact p values, see Source Data. For statistics, p summary and sample sizes, see [Methods](#).

[Source data](#)

[**Extended Data Fig. 8 Fibroblasts directly affect CD8 \$^+\$ T cells through the CXCL9/CXCL10–CXCR3 axis in vitiligo.**](#)

a, Schematic diagram of the experimental procedure of lentivirus-infected cell type analysis. FACS profiles and quantification showing majority of lentivirus-infected cell population in the skin is fibroblast. **b**, QPCR analysis validated high knockdown efficiency of shRNAs targeting CXCL9/CXCL10 in vitro. **c**, Top, experimental design of in vivo knockdown efficiency detection in different dermal cell types. Bottom, qPCR analysis in FACS-isolated virus-infected (RFP+) fibroblasts, endothelial cells, and immune cells, indicating CXCL9 and CXCL10

expression were specifically silenced in virus-infected fibroblasts, not in other cell types. **d**, Representative whole-mount images showing dermis and overlying epidermis with injection of indicated shRNA-expressing lentivirus. Box-whisker plots of epidermal melanocyte number in each scale of mice tail skin injected with the indicated shRNAs. **e**, FACS quantification of immune cell types in tail skin of WT and *Pdgfra*^{CreER}; *IFNGR1*^{fl/fl} cKO mice with three different vitiligo induction methods. **f**, Feature plot showing CXCR3 expression pattern and quantification of CXCR3 in all cell clusters from patients with vitiligo and healthy donors. **g**, Representative FACS profiles and quantification of CXCR3+ cells composition in vitiligo-induced mice (at Day 33). Scale bars, 500 µm (**d**). For exact p values, see Source Data. For statistics, p summary and sample sizes, see [Methods](#).

[Source data](#)

[**Extended Data Fig. 9 Anatomically distinct dermal fibroblasts have intrinsic differences in the IFN \$\gamma\$ response.**](#)

a, Heat map of upregulated genes in human dermal fibroblasts from different anatomical regions after IFN γ treatment (> 2-fold, *t* test p < 0.01). **b**, *HOX* expression pattern (based on RNA-seq results) in human dermal fibroblasts from eight different anatomic regions with or without IFN γ treatment during in vitro culture. **c**, QPCR analysis of *HOXB8*, *HOXC8*, *HOXB13*, and *HOXD11* in human dermal fibroblasts from eight anatomic regions. **d**, *JAK1*, *JAK2*, *IFNGR1*, and *STAT1* expression pattern (based on RNA-seq results) in human dermal fibroblasts from eight different anatomic regions with or without IFN γ treatment. **e**, QPCR analysis of *TNFSF10*, *IL6*, and *CTSH* in human dermal fibroblasts from eight anatomically distinct regions after IFN γ treatment. **f**, Left, representative whole-mount images and quantification of hair follicle density in the dorsal and paw dorsal skin of mice at P21. Right, representative section image and quantification of skin thickness in the dorsal and paw dorsal skin of mice at P60. **g**, Representative immunofluorescent images, and quantification of hair follicle located melanocytes in ventral and paw ventral skin of control mice and vitiligo mice (at Day 300). **h**, FACS analysis and quantification of skin CD3+CD8+ T cells and CD117+ melanocytes from paw dorsal, paw

ventral, dorsal, and ventral regions at Day 300 after vitiligo induction. **i**, Schematic diagram of fibroblast isolation and qPCR analysis of *CXCL9* and *CXCL10* in fibroblasts from four distinct regions of long-term vitiligo mice. **j**, Quantification of migration ratio using medium without fibroblast, corresponding to Fig. [4g](#). **k**, Relationship between melanocyte number and infiltrated CD8+ T cell number in each scale after vitiligo induction based on quantification data in Extended Data Fig. [4h](#). Scale bars, 100 μ m (**f**), 50 μ m (**g**). For exact p values, see Source Data. For statistics, p summary and sample sizes, see [Methods](#).

[Source data](#)

[**Extended Data Fig. 10 Mathematical modelling reveals that fibroblasts direct collective CD8 \$^{\pm}\$ T cell local activity.**](#)

a, The mathematic model developed to predict local CD8+ T cell recruitment and clonal expansion behaviour. The model is a 3D square lattice with two layers. The upper epidermal layer contains T cells and melanocytes. The lower dermal layer contains fibroblasts with different chemotactic abilities. The CD8+ T cell population in skin is considered to be a decentralized system. Each CD8+ T cell is equipped with the means of sensing a change in density. Over time, the behaviour of each cell changes according to its state and the states of its neighbouring cells and the surrounding signals. The collective pattern can be globally modulated by changing the parameters governing local cell-cell interactions. Once the T cell surrounding a melanocyte exceeds a threshold number, melanocyte death and IFN γ secretion occur. The IFN γ signal induce neighbouring fibroblast chemotactic effect to recruit nearby CD8+ T cells, reaching the local CD8+ T cell density threshold for adjacent melanocyte cytotoxicity and IFN γ secretion. This positive feedback loop between CD8+ T cells and fibroblasts could ensure T cell clonal expansion and vitiligo progression. **b**, The mathematic model predicts the expansion process of a single T cell clone. Five representative time points from T0 to T80 show the initial T cell cluster state and the subsequent T cell expansion process over 80 time units. In model 1 with normal chemotactic fibroblasts, full spectrum of T cell cluster formation and expansion patterns observed in WT vitiligo mouse model were reproduced (first row, fibroblasts with normal chemotactic

effect, also shown in Supplementary Videos 1, 2). In model 2, in which the fibroblasts were incapable of chemotaxis, we obtained patterns observed in *Pdgfra*^{CreER}; *IFNGR1*^{fl/fl} cKO mice. In this model, although T cells could still randomly aggregate in the epidermis, they failed to propagate this effect and recruit more T cells to the initial site (second row, fibroblasts with no chemotactic effect, also shown in Supplementary Videos 3, 4). In model 3, in which the chemotactic effect of fibroblasts was turned down to 1/2 of the normal value, random CD8+ T cell aggregates only recruited a limited number of T cells, resulting in slow T cell cluster expansion and melanocyte loss in one or two directions (third row, fibroblasts have 1/2 normal chemotactic effect, also shown in Supplementary Videos 5, 6). **c**, The mathematic model predicts large-scale T cell clone expansion over the long term. Four representative time points from T0 to T450 show the initial state and the subsequent T cell cluster expansion process over 450 time units. CD8+ T cells in the normal chemotaxis model efficiently coordinate so as to achieve clonal expansion and melanocyte clearance. CD8+ T cells in the no chemotaxis model fail to undergo clonal expansion and melanocyte death is detected. CD8+ T cells in the weak chemotaxis model (migration ability decreases to 1/2) generate small T cell clones and expand slowly. **d**, The mathematic model predicts the T cell clone distribution pattern under regional variant fibroblasts with different levels of chemotactic effect to T cell. In this model, the fibroblasts marked in blue have a normal chemotactic effect, and those marked in green have a weak chemotactic effect. In this region, T cell chemotaxis (migration ability) decrease to 1/2 of normal value. Six representative time points from T0 to T600 show the initial state and subsequent T cell cluster expansion (also shown in Supplementary Video 7). The results show that the T cell clones are more likely to expand and generate white patches on the normal chemotactic region. Both white patches and T cell clone patterns are highly correlated with the regional fibroblast variants. **e, f**, QPCR analysis validated high knockdown efficiency (**e**) and the effect blocking IFN γ downstream signal (**f**) in vitro for shRNAs targeting *IFNGR1*, *JAK1*, or *STAT1*. **g**, Representative epidermis whole-mount immunofluorescent staining images and density plot of shRNA-mediated knockdown assay, relative to the corresponding dermis in Fig. 4k. **h**, Box-whisker plots and correlation analysis of T cell number versus percentage of infected fibroblasts (upper panels), and melanocyte number versus percentage of

infected fibroblasts (lower panels) in each scale of *in vivo* mosaic fibroblast knockdown experiment. **i**, Scatter plots of median of melanocyte number versus percentage of infected fibroblasts in each scale of *in vivo* mosaic fibroblast knockdown experiment. Scale bars, 500 μm (g). For exact p values, see Source Data. For statistics, p summary and sample sizes, see [Methods](#).

[Source data](#)

Supplementary information

[Supplementary Information](#)

This file contains Supplementary Tables 1–8 and their accompanying legends.

[Reporting Summary](#)

[Modelling with normal chemotaxis](#)

Supplementary Video 1 . The mathematical model predicts one T cell clone expansion under normal chemotaxis.

[Modelling with normal chemotaxis in large scale](#)

Supplementary Video 2 . The mathematical model predicts large-scale T cell clone expansion under normal chemotaxis.

[Modelling with no chemotaxis](#)

Supplementary Video 3 . The mathematical model predicts one T cell clone expansion under no chemotaxis.

[Modelling with no chemotaxis in larger scale](#)

Supplementary Video 4 . The mathematical model predicts large-scale T cell clone expansion under no chemotaxis.

Modelling with weak chemotaxis

Supplementary Video 5 . The mathematical model predicts one T cell clone expansion when chemotaxis decreases to half of the normal value.

Modelling with weak chemotaxis in larger scale

Supplementary Video 6 . The mathematical model predicts large-scale T cell clone expansion when chemotaxis decreases to half of the normal value.

Modelling with regional variant chemotaxis

Supplementary Video 7 . The mathematical model predicts the T cell clone distribution pattern under regional variant fibroblasts with different levels of chemotactic effect to T cell. In this model, the fibroblasts marked in blue have normal chemotactic effect, whereas those marked in green have weak chemotactic effect. In this region, T cell chemotaxis decreases to half of the normal value.

Source data

[**Source Data Fig. 1**](#)

[**Source Data Fig. 2**](#)

[**Source Data Fig. 3**](#)

[**Source Data Fig. 4**](#)

[**Source Data Extended Data Fig. 1**](#)

[**Source Data Extended Data Fig. 2**](#)

[**Source Data Extended Data Fig. 3**](#)

[**Source Data Extended Data Fig. 4**](#)

[**Source Data Extended Data Fig. 5**](#)

[**Source Data Extended Data Fig. 6**](#)

[**Source Data Extended Data Fig. 7**](#)

[**Source Data Extended Data Fig. 8**](#)

[**Source Data Extended Data Fig. 9**](#)

[**Source Data Extended Data Fig. 10**](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Xu, Z., Chen, D., Hu, Y. *et al.* Anatomically distinct fibroblast subsets determine skin autoimmune patterns. *Nature* **601**, 118–124 (2022).
<https://doi.org/10.1038/s41586-021-04221-8>

- Received: 02 March 2020
- Accepted: 05 November 2021
- Published: 15 December 2021

- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04221-8>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04221-8>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 08 December 2021](#)

Non-genetic determinants of malignant clonal fitness at single-cell resolution

- [Katie A. Fennell](#)^{1,2} ORCID: orcid.org/0000-0002-7814-4851,
- [Dane Vassiliadis](#) ORCID: orcid.org/0000-0002-7814-4851^{1,2} ORCID: orcid.org/0000-0002-7814-4851,
- [Enid Y. N. Lam](#) ORCID: orcid.org/0000-0001-5843-7836^{1,2},
- [Luciano G. Martelotto](#)³,
- [Jesse J. Balic](#)¹,
- [Sebastian Hollizeck](#) ORCID: orcid.org/0000-0002-9504-3497^{1,2},
- [Tom S. Weber](#)^{4,5},
- [Timothy Semple](#)^{1,3},
- [Qing Wang](#)⁶,
- [Denise C. Miles](#)^{4,5},
- [Laura MacPherson](#)^{1,2},
- [Yih-Chih Chan](#) ORCID: orcid.org/0000-0003-2177-5406^{1,2},
- [Andrew A. Guirguis](#) ORCID: orcid.org/0000-0003-3134-0596^{1,2},
- [Lev M. Kats](#) ORCID: orcid.org/0000-0001-8742-8138^{1,2},
- [Emily S. Wong](#) ORCID: orcid.org/0000-0003-0315-2942^{6,7},
- [Sarah-Jane Dawson](#) ORCID: orcid.org/0000-0002-8276-0374^{1,2,3},
- [Shalin H. Naik](#) ORCID: orcid.org/0000-0003-0299-3301^{4,5} &
- [Mark A. Dawson](#) ORCID: orcid.org/0000-0002-5464-5029^{1,2,3}

Nature volume **601**, pages 125–131 (2022)

- 15k Accesses

- 2 Citations
- 213 Altmetric
- [Metrics details](#)

Subjects

- [Cancer stem cells](#)
- [Leukaemia](#)
- [Transcriptomics](#)
- [Tumour heterogeneity](#)

Abstract

All cancers emerge after a period of clonal selection and subsequent clonal expansion. Although the evolutionary principles imparted by genetic intratumour heterogeneity are becoming increasingly clear¹, little is known about the non-genetic mechanisms that contribute to intratumour heterogeneity and malignant clonal fitness². Here, using single-cell profiling and lineage tracing (SPLINTR)—an expressed barcoding strategy—we trace isogenic clones in three clinically relevant mouse models of acute myeloid leukaemia. We find that malignant clonal dominance is a cell-intrinsic and heritable property that is facilitated by the repression of antigen presentation and increased expression of the secretory leukocyte peptidase inhibitor gene (*Slpi*), which we genetically validate as a regulator of acute myeloid leukaemia. Increased transcriptional heterogeneity is a feature that enables clonal fitness in diverse tissues and immune microenvironments and in the context of clonal competition between genetically distinct clones. Similar to haematopoietic stem cells³, leukaemia stem cells (LSCs) display heritable clone-intrinsic properties of high, and low clonal output that contribute to the overall tumour mass. We demonstrate that LSC clonal output dictates sensitivity to chemotherapy and, although high- and low-output clones adapt differently to therapeutic pressure, they coordinately emerge from minimal residual disease with increased expression of the LSC program. Together, these data provide

fundamental insights into the non-genetic transcriptional processes that underpin malignant clonal fitness and may inform future therapeutic strategies.

This is a preview of subscription content

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Clonal dominance is cell-intrinsic.

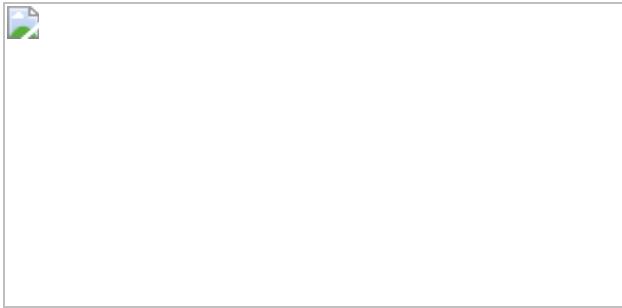


Fig 2: *Slpi* regulates malignant clonal dominance *in vivo*.

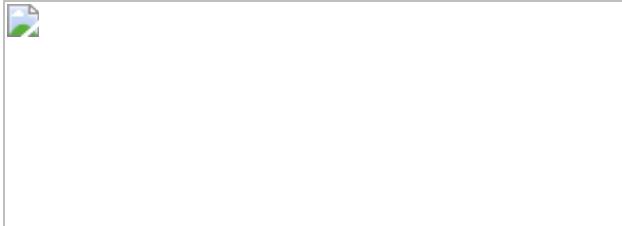
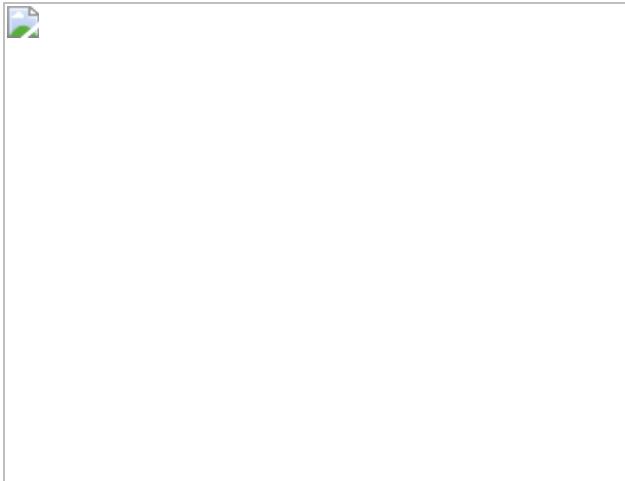


Fig. 3: Non-genetic inheritance of the transcriptional program defining clonal dominance.



Fig 4: Clonal dynamics under therapeutic pressure.



Data availability

All high-throughput sequencing data supporting the findings of this study have been deposited in the NCBI Gene Expression Omnibus under primary accession code [GSE161676](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE161676).

Code availability

The code that support the findings of this study are available online (https://atlassian.petermac.org.au/bitbucket/scm/daw/splintr_paper_code and https://atlassian.petermac.org.au/bitbucket/scm/daw/splintr_barcode_seq_pipeline).

References

1. 1.

Turajlic, S., Sottoriva, A., Graham, T. & Swanton, C. Resolving genetic heterogeneity in cancer. *Nat. Rev. Genet.* **20**, 404–416 (2019).

2. 2.

Marine, J. C., Dawson, S. J. & Dawson, M. A. Non-genetic mechanisms of therapeutic resistance in cancer. *Nat. Rev. Cancer* **20**,

743–756 (2020).

3. 3.

Rodriguez-Fraticelli, A. E. et al. Single-cell lineage tracing unveils a role for TCF15 in haematopoiesis. *Nature* **583**, 585–589 (2020).

4. 4.

Lin, D. S. et al. DiSNE movie visualization and assessment of clonal kinetics reveal multiple trajectories of dendritic cell development. *Cell Rep.* **22**, 2557–2566 (2018).

5. 5.

Greaves, M. When one mutation is all it takes. *Cancer Cell* **27**, 433–434 (2015).

6. 6.

Visvader, J. E. Cells of origin in cancer. *Nature* **469**, 314–322 (2011).

7. 7.

Krivtsov, A. V. et al. Cell of origin determines clinically relevant subtypes of MLL-rearranged AML. *Leukemia* **27**, 852–860 (2013).

8. 8.

Pendleton, M., Lindsey, R. H. Jr, Felix, C. A., Grimwade, D. & Osheroff, N. Topoisomerase II and leukemia. *Ann. N. Y. Acad. Sci.* **1310**, 98–110 (2014).

9. 9.

Somervaille, T. C. & Cleary, M. L. Identification and characterization of leukemia stem cells in murine MLL-AF9 acute myeloid leukemia. *Cancer Cell* **10**, 257–268 (2006).

10. 10.

Krivtsov, A. V. et al. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature* **442**, 818–822 (2006).

11. 11.

Dahlin, J. S. et al. A single-cell hematopoietic landscape resolves 8 lineage trajectories and defects in Kit mutant mice. *Blood* **131**, e1–e11 (2018).

12. 12.

Matsuo, H. et al. Fusion partner-specific mutation profiles and KRAS mutations as adverse prognostic factors in MLL-rearranged AML. *Blood Adv.* **4**, 4623–4631 (2020).

13. 13.

The Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* **368**, 2059–2074 (2013).

14. 14.

Shaffer, S. M. et al. Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature* **546**, 431–435 (2017).

15. 15.

Raj, A. & van Oudenaarden, A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216–226 (2008).

16. 16.

Larsson, A. J. M. et al. Genomic encoding of transcriptional burst kinetics. *Nature* **565**, 251–254 (2019).

17. 17.

MacPherson, L. et al. HBO1 is required for the maintenance of leukaemia stem cells. *Nature* **577**, 266–270 (2020).

18. 18.

Coombs, C. C., Tavakkoli, M. & Tallman, M. S. Acute promyelocytic leukemia: where did we start, where are we now, and the future. *Blood Cancer J.* **5**, e304 (2015).

19. 19.

Jones, C. L., Inguva, A. & Jordan, C. T. Targeting energy metabolism in cancer stem cells: progress and challenges in leukemia and solid tumors. *Cell Stem Cell* **28**, 378–393 (2021).

20. 20.

Zuber, J. et al. Mouse models of human AML accurately predict chemotherapy response. *Genes Dev.* **23**, 877–889 (2009).

21. 21.

Duy, C. et al. Chemotherapy induces senescence-like resilient cells capable of initiating AML recurrence. *Cancer Discov.* **11**, 1542–1561 (2021).

22. 22.

Bell, C. C. et al. Targeting enhancer switching overcomes non-genetic drug resistance in acute myeloid leukaemia. *Nat. Commun.* **10**, 2723 (2019).

23. 23.

Rambow, F. et al. Toward minimal residual disease-directed therapy in melanoma. *Cell* **174**, 843–855 (2018).

24. 24.

Fong, C. Y. et al. BET inhibitor resistance emerges from leukaemia stem cells. *Nature* **525**, 538–542 (2015).

25. 25.

Shluss, L. I. et al. Tracing the origins of relapse in acute myeloid leukaemia to stem cells. *Nature* **547**, 104–108 (2017).

26. 26.

Zorita, E., Cusco, P. & Filion, G. J. Starcode: sequence clustering based on all-pairs search. *Bioinformatics* **31**, 1913–1919 (2015).

27. 27.

Silver, D. P. & Livingston, D. M. Self-excising retroviral vectors encoding the Cre recombinase overcome Cre-mediated cellular toxicity. *Mol. Cell* **8**, 233–243 (2001).

28. 28.

Kebschull, J. M. & Zador, A. M. Cellular barcoding: lineage tracing, screening and beyond. *Nat. Methods* **15**, 871–879 (2018).

29. 29.

Naik, S. H., Schumacher, T. N. & Perie, L. Cellular barcoding: a technical appraisal. *Exp. Hematol.* **42**, 598–608 (2014).

30. 30.

Ye, M. et al. Hematopoietic differentiation is required for initiation of acute myeloid leukemia. *Cell Stem Cell* **17**, 611–623 (2015).

31. 31.

Naik, S. H. et al. Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature* **496**, 229–232 (2013).

32. 32.

Langmead, B. Aligning short sequencing reads with Bowtie. *Curr. Protoc. Bioinform.* **32**, 11.7.1–11.7.14 (2010).

33. 33.

Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

34. 34.

Scheinin, I. et al. DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. *Genome Res.* **24**, 2022–2032 (2014).

35. 35.

McLaren, W. et al. The Ensembl variant effect predictor. *Genome Biol.* **17**, 122 (2016).

36. 36.

Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019).

37. 37.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).

38. 38.

Stoeckius, M. et al. Cell hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 224 (2018).

39. 39.

Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019).

40. 40.

Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38–44 (2018).

41. 41.

Zappia, L. & Oshlack, A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *Gigascience* **7**, giy083 (2018).

42. 42.

Finak, G. et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, 278, (2015).

43. 43.

Michailidou, K. et al. Association analysis identifies 65 new breast cancer risk loci. *Nature* **551**, 92–94, (2017).

44. 44.

Marjanovic, N. D. et al. emergence of a high-plasticity cell state during lung cancer evolution. *Cancer Cell* **38**, 229–246 (2020).

45. 45.

Azizi, E. et al. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell* **174**, 1293–1308 (2018).

Acknowledgements

We thank members of the Dawson laboratory, particularly S. Jackson, O. Gilan and K. Knezevic for technical assistance with experiments; F. Rossello for advice and discussions related to data analysis and D. Zalcenstein and J. Gao for advice regarding lentiviral barcode design and cloning. We thank the following funders for fellowship and grant support: Peter MacCallum Foundation Grant (to D.V.), VCA mid-career fellowship (to E.Y.N.L.), CSIRO Future Science Fellowship in Synthetic Biology (to T.S.W.), NHMRC Early Career Fellowship grant (no. 1052195, to D.C.M.) NHMRC Investigator Grant (no. 1196749, to M.A.D.), Cancer Council Victoria Dunlop Fellowship (to M.A.D), Howard Hughes Medical Institute international research scholarship (no. 55008729, to M.A.D), NHMRC Investigator Grant (no. 1196755, to S.J.D.), CSL Centenary fellowship (to S.J.D.) and NHMRC Project Grant (no. 1128984, to S.J.D. and M.A.D.).

Author information

Author notes

1. These authors contributed equally: Katie A. Fennell, Dane Vassiliadis

Affiliations

1. Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia
Katie A. Fennell, Dane Vassiliadis, Enid Y. N. Lam, Jesse J. Balic, Sebastian Hollizeck, Timothy Semple, Laura MacPherson, Yih-Chih Chan, Andrew A. Guirguis, Lev M. Kats, Sarah-Jane Dawson & Mark A. Dawson
2. Sir Peter MacCallum Department of Oncology, The University of Melbourne, Melbourne, Victoria, Australia

Katie A. Fennell, Dane Vassiliadis, Enid Y. N. Lam, Sebastian Hollizeck, Laura MacPherson, Yih-Chih Chan, Andrew A. Guirguis, Lev M. Kats, Sarah-Jane Dawson & Mark A. Dawson

3. The University of Melbourne Centre for Cancer Research, The University of Melbourne, Melbourne, Victoria, Australia

Luciano G. Martelotto, Timothy Semple, Sarah-Jane Dawson & Mark A. Dawson

4. The Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria, Australia

Tom S. Weber, Denise C. Miles & Shalin H. Naik

5. The Department of Medical Biology, The University of Melbourne, Melbourne, Victoria, Australia

Tom S. Weber, Denise C. Miles & Shalin H. Naik

6. Victor Chang Cardiac Research Institute, Sydney, New South Wales, Australia

Qing Wang & Emily S. Wong

7. School of Biotechnology and Biomolecular Sciences, UNSW, Sydney, New South Wales, Australia

Emily S. Wong

Contributions

K.A.F., D.V., S.H.N. and M.A.D. designed/analysed experiments. K.A.F., D.V. and M.A.D. wrote the manuscript with contributions from all of the authors. K.A.F. developed SPLINTR with assistance from D.C.M., E.Y.N.L. and D.V. with joint supervision by S.H.N. and M.A.D. All experimental work was performed by K.A.F. and D.V. with assistance from J.J.B., T.S., L.M., A.G. and Y.-C.C. under the supervision of M.A.D. All

computational analyses were performed by D.V. with assistance from E.Y.N.L., S.H., Q.W. and T.S.W. under the supervision of E.S.W., S.H.N. and M.A.D. Critical reagents and resources were provided by L.M.K., S.H.N., S.-J.D. and M.A.D.

Corresponding author

Correspondence to [Mark A. Dawson](#).

Ethics declarations

Competing interests

M.A.D. has been a member of advisory boards for CTX CRC, Storm Therapeutics, Celgene and Cambridge Epigenetix. S.-J.D. has been a member of advisory boards for AstraZeneca. S.-J.D. has received research funding from Genentech. M.A.D. and S.-J.D. receive research funding from CTx CRC. The other authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Ross Levine and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

[Extended Data Fig. 1 Generation and proof of concept of SPLINTR barcode libraries.](#)

a, Three independent SPLINTR barcode libraries were generated, each containing a distinct fluorescent reporter and barcode structure (Methods).

SPLINTR barcodes are constitutively transcribed upon genome integration, enabling tracking of clonally related cells together with their individual transcriptomes. **b**, Base-position analysis of SPLINTR barcode structures. The designated combinations of strong (G or C) and weak (A or T) bases results in barcode structures that can be computationally deconvoluted from other SPLINTR libraries. **c**, Venn diagrams showing number of unique barcodes in each library identified via overlap of two deep-sequenced PCR technical replicates per library. Boxplots showing distributions of Hamming (**d**) and Levenshtein edit (**e**) distances between individual barcodes calculated for each individual library. 1000 barcodes present in each reference library were randomly sampled without replacement ($n = 100$ resampling events per barcode library, GFP library, $n = 170,885$ barcodes, BFP library, $n = 672,982$ barcodes and mCherry library, $n = 1,324,188$ barcodes) and the average hamming/edit distance between all pairwise combinations were computed. Boxplots span the upper quartile (upper limit), median (centre) and lower quartile (lower limit). Whiskers extend a maximum of 1.5x IQR. **f**, Schematic of a pilot *in vitro* SPLINTR barcoding experiment. **g**, Proportional bubbleplot showing barcode distributions derived from barcode-seq from Pool #1 and #2 and scRNA-seq from Pool #1. Bubble size scales with clone size. **h**, Pearson correlation matrix between normalized barcode repertoires from Pool #1 and #2 barcode-seq and Pool #1 scRNA-seq. Correlation R values are shown. **i**, UMAP projection of Pool #1 scRNA-seq dataset. Louvain clusters are indicated. Cells containing a detected SPLINTR barcode are highlighted red. Cells without a barcode are shown in grey. **j**, The number of distinct SPLINTR barcodes detected per cell in the Pool #1 scRNA-seq dataset. Doublets were classified as cells containing a unique combination of two or more SPLINTR barcodes and are shown in red. Singlet cells are in blue. **k**, Total UMI counts per cell for cells grouped according to their predicted doublet status as described in (**j**) (doublet = 469 cells, singlet = 7494 cells, unknown = 2388 cells). Unknown indicates cells in which no SPLINTR barcode was detected. Boxplots span the upper quartile (upper limit), median (centre) and lower quartile (lower limit). Whiskers extend a maximum of 1.5x IQR.

Extended Data Fig. 2 Clonal dominance is an intrinsic feature of leukaemia initiating clones, regardless of cell-of-origin.

Results are representative of $n = 4$ mice per experimental group. **a**, Schematic of an *in vivo* barcoding experiment investigating the clonal competition and dynamics in *MLL-AF9* AML derived from an HSC ($\text{Lin}^-, \text{c-Kit}^+, \text{Sca-1}^+$) or GMP ($\text{Lin}^-, \text{c-Kit}^+, \text{CD34}^+, \text{Fc}\gamma\text{R}^{\text{HI}}$) cell-of-origin. For the FACS gating strategy, see Supplementary Information Figure 1. **b**, Disease burden based on CD45.2 donor cell expression in the bone marrow and spleen of *Ptprc^a* (top) or NSG (bottom) mice transplanted with either HSC or GMP derived leukaemia. All mice were collected at the same timepoint. Bars represent the median. **c**, Histogram detailing the total number of GMP (red) or HSC- derived leukaemic clones (green) comprising the 95th percentile from bulk barcode-seq at Baseline and Disease in *Ptprc^a* (top) or NSG (bottom) mice transplanted with a 50:50 mix of either HSC or GMP-derived leukaemias. **d**, Violin plots showing the clone size expressed as \log_{10} transformed counts per million (CPM) of HSC or GMP-derived clones collected from NSG (top, HSC: $n = 15$ clones, GMP: $n = 14$ clones) or *Ptprc^a* (bottom, HSC: $n = 13$ clones, GMP: $n = 10$ clones) mice that received 50:50 HSC/GMP competition transplants. Boxplots span the upper quartile (upper limit), median (centre) and lower quartile (lower limit). Whiskers extend a maximum of 1.5x IQR. Two-sided T-test. NSG p = 0.55, *Ptprc^a* p = 0.13. **e**, Heat maps of pairwise Pearson correlation values for HSC derived clones (top) and GMP derived clones (bottom) from Barcode-seq samples from mice transplanted with single *MLL-AF9* HSC or GMP-derived leukaemias or a 50:50 HSC/GMP cell mix. **f**, Bubble plot showing distribution of HSC and GMP derived leukaemic clone sizes in barcode-seq data from Baseline (T_0) and disease timepoints from mice that received transplants of 50:50 HSC/GMP cell mix. Clones ordered based on their abundance at the baseline timepoint (Ranked abundance at baseline). Dominant clones (those above 2% frequency) from GMP or HSC groups are coloured, all other clones are shown in grey. HSC-derived leukaemias were barcoded with the GFP SPLINTR library. GMP-derived leukaemias were barcoded with the mCherry SPLINTR library.

Extended Data Fig. 3 SPLINTR identifies both LSC and non-LSC like transcriptional signatures in disease causing clones.

a, Schematic of an *in vivo* experiment to identify the transcriptional features of dominant murine *MLL-AF9 + Kras^{G12D}* leukaemic clones that drive disease. **b**, Correlation plot between the frequencies of SPLINTR barcoded bone marrow (left) and spleen (right) clones at the Baseline and Disease timepoints ($n = 4$ mice). Inset values indicate adjusted R^2 and p-values using the F-test for linear regression on frequencies of barcodes detected at both timepoints. Clone barcode frequencies were quantified by barcode-seq. Dashed line indicates a threshold frequency (2%) above which a clone is considered dominant. ($n = 4$ mice) **c**, UMAP projection of scRNA-seq data from SPLINTR barcoded *MLL-AF9 + Kras^{G12D}* cells taken at the baseline timepoint prior to transplantation. Top row, left to right: Louvain clusters, cells comprising dominant clones, cells coloured according to module enrichment scores for genes upregulated in leukaemic stem cells (LSC Signature - UP). Bottom row, left to right: Cells coloured according to module enrichment scores for genes upregulated in normal long-term haematopoietic stem cells (LT-HSCs), common myeloid progenitors (CMPs) and granulocyte/macrophage progenitors (preGM/GMPs). $q_{10} = 10^{\text{th}}$ quantile, $q_{90} = 90^{\text{th}}$ quantile. Module enrichment scores for genes upregulated in LSCs **(d)**, normal LT-HSC **(e)** or GMPs **(f)** or CMPs **(g)** for each dominant clone vs. all other cells (cells per group left to right: $n = 24, 34, 29, 147, 115, 32, 13115$). Asterisks indicate the result of a two-sided T-test using Holm multiple testing correction for each clone vs. all other cells. Number of cells shown above each category. Adjusted p-values (left to right) for **(d)**: $6e^{-07}, 3.2e^{-07}, 0.33, 0.25, 3e^{-34}, 0.92$. **(e)**: $1, 1, 1, 0.00012, 2.2e^{-14}, 1$. **(f)**: $2.6e^{-07}, 9.1e^{-07}, 1, 1, 1.2e^{-45}, 1$. **(g)**: $1.5e^{-10}, 9.1e^{-11}, 0.76, 6.9e^{-06}, 3.7e^{-55}, 0.76$. *** = $p\text{-adj} < 0.001$. Boxplots for **(d)** to **(g)** span the upper quartile (upper limit), median (centre) and lower quartile (lower limit). Whiskers extend a maximum of 1.5x IQR. **h**, Integration of scRNA-seq data sets from *MLL-AF9* primary leukaemia cells from **(c)** and healthy progenitor HSPC populations from Dahlin et al. 2018. Top-left panel: UMAP projection showing Louvain clusters. Top-middle panel: Cells labelled according to their sample of origin. Primary *MLL-AF9* baseline timepoint cells (*MLL-AF9*), Dahlin et al. lineage negative, c-Kit positive cells (LK), and Dahlin et al. lineage negative, Sca-1 positive and c-Kit positive cells (LSK). Top-right: UMAP projection with dominant primary *MLL-AF9* clones highlighted. Bottom row: UMAP projections showing

gene set module enrichment scores for genes upregulated in Long Term Haematopoietic Stem Cells (LT-HSC) (left), Common Myeloid Progenitors (CMPs) (middle) and Granulocyte/Macrophage progenitors (preGM/GMPs) (right). $q10 = 10^{\text{th}}$ quantile, $q90 = 90^{\text{th}}$ quantile. **i**, Hypergeometric test for Louvain clusters enriched for dominant *MLL-AF9* clones. Negative \log_{10} transformed P values per cluster are shown. Red clusters are significantly enriched for dominant clones. Dashed line indicates a $p = 0.05$ significance threshold. **j**, The percentage of the 3 different cell types comprising each of the Louvain clusters in the integrated dataset. Clusters enriched for dominant *MLL-AF9* clones are shown in red.

Extended Data Fig. 4 Generation and genetic characterization of secondary and tertiary leukaemic cells.

a, Schematic of a SPLINTR experiment using 2° *MLL-AF9* cells. Timepoints where whole exome sequencing (WES) and low-coverage whole genome sequencing (LC-WGS) analysis was conducted are indicated. **b**, Results of whole exome sequencing (WES) of SPLINTR barcoded tertiary *MLL-AF9* leukaemic cells at Endpoint (T_N), expressed as mutations per Megabase (1 million base pairs). Points indicate raw result from individual mice ($n = 2$). Lines indicate the mean. Full lists of variants identified can be found in Extended Data Table 3. **c**, Copy number variant profiles from LC-WGS data of SPLINTR barcoded tertiary *MLL-AF9* leukaemic cells from bone marrow at Endpoint (top) and a sex and age matched healthy wild type C57BL/6 littermate control (bottom). **d**, UMAP projection of reanalysed *MLL-AF9 + Kras^{G12D}* Baseline scRNA-seq data from Fig. 2. Cells comprising Shared dominant clones are highlighted (top) and normalized expression level of transcripts derived from the *MLL-AF9* retroviral construct (bottom).

Extended Data Fig. 5 Pre-existing transcriptional determinants of clonal dominance in *MLL-AF9* leukaemia.

a, Heat maps of pairwise Pearson correlation values for Barcode-seq samples from *MLL-AF9*, *MLL-AF9 + Kras^{G12D}* and *MLL-AF9 + Flt3^{ITD}* groups ($n = 4\text{-}5$ mice). **b**, UMAP projections of Baseline (T_0) scRNA-seq

datasets from *MLL-AF9* (left column), *MLL-AF9 + Kras^{G12D}* (middle column) and *MLL-AF9 + Flt3^{ITD}* (right column) groups as in Fig. 1. Rows show results of Louvain clustering (top), module enrichment scores for genes upregulated in LSCs (middle) or the MHC Class I signature (bottom). Heat maps showing the top differentially expressed genes in dominant clones vs all other cells at the Baseline timepoint for the *MLL-AF9* (**c**) and *MLL-AF9 + Flt3^{ITD}* (**d**) groups. A random sample of all other cells within the same genotype is shown on each heat map. Biological pathways upregulated (**e**) and downregulated (**f**) in dominant clones for each individual genotype and the Competition 80:15:5 group at Endpoint (T_N). Number of genes in each group are indicated at the top. **g**, UMAP projections of scRNA-seq data from Baseline and Disease stage *MLL-AF9 + Kras^{G12D}* cells showing expression patterns of *B2m*, *Myb* and *Slpi*. $q_{10} = 10^{\text{th}}$ quantile, $q_{90} = 90^{\text{th}}$ quantile. Disease scRNA-seq data integrated from two independent mice. **h**, Boxplots of normalized *Slpi* expression (log1p transformed) in cells comprising dominant clones vs. all other cells for the *MLL-AF9* (cells per group left to right: $n = 17, 21, 41, 3, 7, 2, 2, 9088$), *MLL-AF9 + Kras^{G12D}* (cells per group left to right: $n = 14, 18, 29, 443, 63, 386, 9, 27, 36, 120, 136, 1, 45, 17, 80, 23, 162, 50, 1, 48, 13, 41, 10304$) and *MLL-AF9 + Flt3^{ITD}* (cells per group left to right: $n = 3257, 664, 1066, 501, 110, 26, 306, 34, 24, 11, 32$) groups. One-sided Wilcoxon rank sum test with Benjamini-Hochberg multiple test correction for each clone vs. all other cells (Other) within each group. Adjusted p-values for *MLL-AF9* left to right: $0.022, 2e^{-10}, 1, 0.022, 0.022, 0.022, 1$. Adjusted p-values for *MLL-AF9 + Kras^{G12D}* left to right: $0.85, 0.00023, 6.6e^{-09}, 0.13, 7.5e^{-14}, 1, 0.12, 1.7e^{-10}, 2.9e^{-10}, 1.1e^{-36}, 8.8e^{-11}, 0.66, 4.3e^{-05}, 0.51, 8.8e^{-06}, 1.5e^{-09}, 2.2e^{-31}, 6.2e^{-13}, 0.13, 5e^{-13}, 0.03, 2.8e^{-13}$. Adjusted p-values for *MLL-AF9 + Flt3^{ITD}* left to right: $0.24, 1, 1, 1, 1, 0.47, 7.1e^{-17}, 1, 0.066, 0.66, 1, 1, 0.0089, 1, 1, 1, 1, 0.46, 0.46, 1.6e^{-07}, 0.17$. * = p-adj < 0.05 , ** = p-adj < 0.01 , *** = p-adj < 0.001 , **** = p-adj < 0.0001 . **i**, UMAP projections of Disease stage *MLL-AF9* (left column), *MLL-AF9 + Kras^{G12D}* (middle column) and *MLL-AF9 + Flt3^{ITD}* (right column) scRNA-seq datasets. Rows show *B2m* expression (top) and module enrichment scores for genes upregulated (middle) or downregulated (bottom) in LSCs in each dataset. Boxplots span the upper quartile (upper limit), median (centre) and lower quartile (lower limit). Whiskers extend a maximum of 1.5x IQR.

Extended Data Fig. 6 Results of *in vivo* modelling of clonal competition.

Proportional bubble plot of Barcode-seq data showing dominant clones in the Competition 1:1:1 (**a**) and Competition 80:15:5 (**b**) groups at Baseline and Disease timepoints. Red, green and blue regions show *MLL-AF9 + Flt3^{ITD}*, *MLL-AF9* and *MLL-AF9 + Kras^{G12D}* derived clones respectively ($n = 4$ mice per group) **c**, Correlation plots of clonal repertoires of *MLL-AF9 + Kras^{G12D}* clones from baseline and disease samples in the Competition 1:1:1 (left) or Competition 80:15:5 (right) groups ($n = 4$ mice per group) **d**, Venn diagram indicating the number of shared and distinct *MLL-AF9 + Kras^{G12D}* clones at Disease between *Kras* only, Competition 1:1:1 and Competition 80:15:5 groups. **e**, Clone size at Baseline (T_0) expressed as \log_{10} transformed counts per million (CPM) of Competition 80:15:5 specific clones ($n = 22$ clones) vs *Kras* specific clones ($n = 22$ clones). Boxplots span the upper quartile (upper limit), median (centre) and lower quartile (lower limit). Whiskers extend a maximum of 1.5x IQR. One-sided T-test (*, $p = 0.018$). **f**, Representative flow cytometry analysis of the disease burden in bone marrow and spleen 2 weeks after equal numbers of cells from each genotype were injected into separate mice. Plots are representative of 1 mouse per group. **g**, Proportional bubble plot of normalized Barcode-seq data showing clonal repertoire of *Kras* clones from Baseline samples in the Competition 1:1:1, Competition 80:15:5 or *MLL-AF9 + Kras^{G12D}* only groups. **h**, Correlation plots of the barcode repertoires present in Baseline cells sampled from the *MLL-AF9 + Kras^{G12D}* only transplants and the Competition 1:1:1 and Competition 80:15:5 transplants. Inset values indicate adjusted R^2 and p-values using a one-tailed F-test for linear regression on CPM of barcodes detected at both timepoints. Results in this figure are representative of $n = 4$ mice per experimental group.

Extended Data Fig. 7 Competition-specific clones display altered transcriptional burst kinetics compared to KRAS-specific clones.

a, Normalized *Slpi* (top) and *B2m* (bottom) expression in *MLL-AF9 + Kras^{G12D}* Baseline scRNA-seq cells comprising *Kras*-specific clones (red, $n = 617$ cells) Competition-specific clones (blue, $n = 926$ cells) Baseline-specific clones (green, $n = 965$ cells) or all other clones (grey, $n = 12,049$ cells). Boxplots indicate median and inter-quartile range (IQR). Whiskers extend a maximum of 1.5x IQR. Two-sided T-test using Holm multiple testing correction for each group vs. all other clones (Other clones). *Kras*-specific (*Slpi*: ****, p-adj = $2e^{-134}$; *B2m*: ****, p-adj = $6.2e^{-36}$), Competition-specific (*Slpi*: ****, p-adj = $6.3e^{-32}$; *B2m*: ****, p-adj = $7.9e^{-36}$), Baseline-specific (*Slpi*: ****, p-adj = $5.8e^{-13}$; *B2m*: p-adj = 0.91). **b**, Histogram showing the percentage of cells belonging to *Kras*-specific, Competition-specific or Baseline-specific clones present in each Louvain cluster for the *MLL-AF9 + Kras^{G12D}* Baseline scRNA-seq dataset. **c**, Permutation analysis of Shannon diversity scores derived from the Louvain cluster occupancy of *Kras*-specific ($n = 617$ cells) and Competition-specific ($n = 926$ cells) groups. Centre point indicates raw Shannon diversity values with error bars reflecting the standard error of 100 resampling events with replacement. Two-sided Mann Whitney U test. *, p = 0.006847. **d**, Scatter plots of \log_{10} transformed burst frequency and burst size values for 5188 genes with inferred transcriptional kinetics based on *MLL-AF9 + Kras^{G12D}* Baseline scRNA-seq data from *Kras*-specific ($n = 617$ cells) and Competition-specific cells ($n = 926$ cells). Red points indicate genes with a statistically significant difference (Benjamini-Hochberg adjusted p-value < 0.01) in Competition-specific cells vs *Kras*-specific cells. **e**, Venn diagram showing overlap of genes with significantly increased or decreased burst size or burst frequency in the Competition-specific group versus the *Kras*-specific group. Biological process gene ontologies for genes with a significantly increased transcriptional burst frequency (**f**) or burst size (**g**) in Competition-specific cells vs *Kras*-specific cells. Selected gene sets with Benjamini-Hochberg adjusted p-value < 0.05 are shown.

Extended Data Fig. 8 Dominant clones possess a distinct chromatin accessibility landscape.

a, Number of clones comprising the 95th percentile at Baseline (T_0) and Endpoint (T_N) bone marrow and spleen samples in NSG and Ptprc^a mice

that received transplants from the same pool of barcoded *MLL-AF9* + *Kras*^{G12D} cells. **b**, Proportional bubble plots of clones identified via Barcode-seq at Baseline and Disease timepoints for each mouse in each immune background. For each sample, clones are sorted in descending rank order according to their percentage abundance in the Baseline sample ($n = 3$ mice per group) **c**, Permutation analysis of Shannon diversity scores derived from the Louvain cluster occupancy of Ptprc^a only ($n = 398$ cells) or Ptprc^a and NSG specific (Shared, $n = 150$ cells) dominant clones from Fig. 2. Center point indicates raw Shannon diversity values with error bars reflecting the standard error of 1000 resampling events with replacement. Two-sided Mann Whitney U test. *, $p = 1e^{-11}$. **d**, Gene expression heat map showing the top differentially expressed genes in dominant clones from the NSG-only, Ptprc^a-only and shared groups vs all other cells at the Baseline timepoint. A random sample of all other cells is shown on the heat map. Genes of interest are highlighted. **e**, Biological pathways enriched in the set of genes upregulated in NSG-specific clones (NSG_up) or downregulated in NSG (NSG_down) and Ptprc^a specific clones (BL6_down). Number of genes in each group are indicated above. **f**, UMAP projections of scRNA-seq dataset from Baseline *MLL-AF9* + *Kras*^{G12D} cells from Fig. 3 showing expression of *Cebpa*, *Cebpβ* and *Cebpδ*. **g**, UMAP projection of scATAC-seq dataset from Baseline *MLL-AF9* + *Kras*^{G12D} cells from Fig. 3 showing results of Louvain clustering and label transfer to the corresponding scRNA-seq dataset as in Fig. 3. **h**, Results of differential accessibility analysis of cells in scATAC-seq clusters 1 and 7 (Dominant) vs cells in all other clusters (Other) at three exemplar loci *B2m* (top), *Kit* (middle) *Cebpe* (bottom). Boxes indicate regions of significantly altered accessibility in the Dominant group. **i**, The top six enriched transcription factor motifs in regions significantly differentially accessible in cells comprising scATAC-seq clusters 1 and 7 (clusters enriched for dominant clones) vs cells in all other clusters. **j**, Results of *Cebpe*, *Cebpa* and isotype control (IgG/Neg) Chromatin Immuno-precipitation coupled qPCR using primers against the promoter of *Slpi* and a negative-control region (Gene desert) in *MLL-AF9* + *Kras*^{G12D} cells. Mean ± s.d. of $n = 3$ independent experiments are shown. **k**, Western blot analysis of Baseline *Slpi* protein levels in *MLL-AF9* + *Kras*^{G12D} cells transduced with guides targeting *Slpi* or non-targeting guide control immediately prior to the *in vitro* and *in vivo* competition assays in

Fig. 2. (representative of $n = 2$ independent experiments). For uncropped blot, see Supplementary Information Figure 2. I, Flow cytometry analysis of bone marrow cells collected from 50:50 group mice. Representative results from $n = 4$ mice.

Extended Data Fig. 9 Serial transplantation reveals a heritable transcriptional program of clonal dominance.

a, Bubble plots of clones identified via Barcode-seq for Donor mouse #1 and corresponding recipient NSG or Ptprc^a mice at secondary transplant Baseline and Disease timepoints. Bubbleplots indicate primary transplant clones (left) or secondary transplant clones (middle and right). For each sample, clones are sorted in descending rank order according to their percentage abundance in the Baseline sample. $n = 2$ mice per group. **b**, UMAP projection of scRNA-seq datasets from bone marrow (top) and spleen (bottom) derived clones at Baseline prior to secondary transplantation. Cells comprising the top three most dominant clones from the primary transplantation are highlighted. **c**, Network plots of enriched Gene Ontology Biological Process terms for genes upregulated (left) and downregulated (right) in secondary transplant dominant clones. **d**, Gene expression heat map showing the top differentially expressed genes in cells comprising dominant clones from the secondary transplantation (red) vs all other cells (grey). A random sample of all other cells is shown on the heat map. Genes positively or negatively correlated with clonal dominance are highlighted red or blue respectively. **e**, UMAP projection of individual scRNA-seq datasets from bone marrow (top) and spleen (bottom) derived clones at Baseline prior to secondary transplantation. Louvain clusters, cells comprising individual dominant clones in the secondary transplantation and expression patterns of *Slpi*, *Cebpe*, *B2m* and the LSC signature are shown from left to right. q10 = 10th quantile, q90 = 90th quantile.

Extended Data Fig. 10 Distinct responses of high and low output AML clones to chemotherapy treatment.

a, Abundance (proportion) of each dominant clone from the primary transplantation shown across the Baseline #2 and Disease #2 timepoints of

the secondary transplantation. Clone abundance in the bone marrow (top) and spleen (bottom) are shown. Difference to mean indicates the difference between mean frequencies of each clone in the primary versus secondary transplants. **b**, Proportions of secondary transplant clones for Donor mouse #1 and recipient Ptprc^a and NSG mice. High output clones contribute >5% of the total disease burden whereas Mid output clones contribute 1-5%. Low output clones contribute below 1% of the total disease burden. Points represent individual clones and are coloured according to their primary transplant clone identity. Low output clones are not coloured. **c**, UMAP projections of the integrated scRNA-seq dataset shown in Fig. 4 showing *c-Myc* expression, and module enrichment scores for the Boroviak Diapause, LSC and Ramalho stemness gene signatures. q10 = 10th quantile, q90 = 90th quantile. **d**, UMAP projections of Baseline timepoint scRNA-seq data from *MLL-AF9 + Kras^{G12D}* cells prior to transplantation. Top row left to right: Louvain clusters, embedded locations of cells comprising low to High output clones (red), Mid/High to low output clones (blue) clones that are extinguished with chemotherapy treatment (green) and clones that survive chemotherapy (purple). Bottom row left to right: *Slpi* expression, *Cebpe* expression *B2m* expression, module enrichment scores for the Boroviak Diapause and Fridman Senescence signatures. q10 = 10th quantile, q90 = 90th quantile. **e**, Gene expression heat map showing the top differentially expressed genes in cells comprising Mid/High to Low output and Low to High output clones in the Vehicle sample and at each treatment timepoint. Normalized expression shown on scale. **f**, Selected enriched Biological Process gene ontologies for the set of genes up and downregulated in Low to High output clones vs Mid/High to Low output clones in the Vehicle sample and at each treatment timepoint. **g**, Selected enriched Biological Process gene ontologies for the set of genes up and downregulated in Mid/High to Low output clones vs Low to High output clones in the Vehicle sample and at each treatment timepoint.

Supplementary information

Supplementary Information

Supplementary Figs. 1 and 2.

Reporting Summary

Peer Review File

Supplementary Table 1

A list of oligos.

Supplementary Table 2

Hashtag information.

Supplementary Table 3

Exome data.

Supplementary Table 4

Gene sets.

Supplementary Table 5

Bursting analysis.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Fennell, K.A., Vassiliadis, D., Lam, E.Y.N. *et al.* Non-genetic determinants of malignant clonal fitness at single-cell resolution. *Nature* **601**, 125–131 (2022). <https://doi.org/10.1038/s41586-021-04206-7>

- Received: 09 November 2020
- Accepted: 30 October 2021
- Published: 08 December 2021
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04206-7>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

Further reading

- [A heritable, non-genetic road to cancer evolution](#)
 - Tamara Prieto
 - Dan A. Landau

Nature (2022)

[A heritable, non-genetic road to cancer evolution](#)

- Tamara Prieto
- Dan A. Landau

News & Views 08 Dec 2021

Beyond genetics for cancer evolution

- Darren J. Burgess

Research Highlight 05 Jan 2022

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04206-7>

| [Section menu](#) | [Main menu](#) |

- Article
- Open Access
- [Published: 15 December 2021](#)

ER proteins decipher the tubulin code to regulate organelle distribution

- [Pengli Zheng](#) [ORCID: orcid.org/0000-0002-3797-4023¹](#),
- [Christopher J. Obara²](#),
- [Ewa Szczesna³](#),
- [Jonathon Nixon-Abell](#) [ORCID: orcid.org/0000-0003-4169-0012^{1,2}](#) nAff7,
- [Kishore K. Mahalingan³](#),
- [Antonina Roll-Mecak^{3,4}](#),
- [Jennifer Lippincott-Schwartz](#) [ORCID: orcid.org/0000-0002-8601-3501²](#) &
- [Craig Blackstone](#) [ORCID: orcid.org/0000-0003-1261-9655^{1,5,6}](#)

[Nature](#) volume 601, pages 132–138 (2022)

- 14k Accesses
- 137 Altmetric
- [Metrics details](#)

Subjects

- [Endoplasmic reticulum](#)
- [Microtubules](#)

Abstract

Organelles move along differentially modified microtubules to establish and maintain their proper distributions and functions^{1,2}. However, how cells interpret these post-translational microtubule modification codes to selectively regulate organelle positioning remains largely unknown. The endoplasmic reticulum (ER) is an interconnected network of diverse morphologies that extends promiscuously throughout the cytoplasm³, forming abundant contacts with other organelles⁴.

Dysregulation of endoplasmic reticulum morphology is tightly linked to neurologic disorders and cancer^{5,6}. Here we demonstrate that three membrane-bound endoplasmic reticulum proteins preferentially interact with different microtubule populations, with CLIMP63 binding centrosome microtubules, kinecin (KTN1) binding perinuclear polyglutamylated microtubules, and p180 binding glutamylated microtubules. Knockout of these proteins or manipulation of microtubule populations and glutamylation status results in marked changes in endoplasmic reticulum positioning, leading to similar redistributions of other organelles. During nutrient starvation, cells modulate CLIMP63 protein levels and p180–microtubule binding to bidirectionally move endoplasmic reticulum and lysosomes for proper autophagic responses.

[Download PDF](#)

Main

Eukaryotes compartmentalize cellular functions within distinct organelles, and regulation of organelle position is critical for cell health. Organelles are transported bidirectionally by motor and adaptor proteins along microtubules¹, which are modulated by multiple post-translational modifications that comprise part of the ‘tubulin code’. Although this code has been implicated in cargo selection and directed organelle movement², how it is decoded to mediate transport and control distribution remains largely unknown.

Endoplasmic reticulum (ER) comprises structurally and functionally divergent membrane compartments that include interconnected tubules, perinuclear matrices and sheets, and the nuclear envelope^{3,7}. The ER is a compelling candidate for exploiting the complexity of the tubulin code, since it spreads throughout the cytoplasm in association with microtubules and makes abundant organelle contacts^{8,9,10}. Most studies of ER shaping and organelle contacts have emphasized peripheral tubular ER. How the denser perinuclear ER is shaped and asymmetrically distributed remains largely unknown, although three ER membrane-bound proteins—CLIMP63, p180 and KTN1—localize prominently to perinuclear ER and are considered sheet-forming proteins¹¹. Even so, depletion of CLIMP63 may paradoxically lead to the expansion of ER matrices or sheets in the periphery^{11,12}, and perinuclear ER matrices or sheets remain abundant even upon simultaneous knockdown of all three proteins, prefiguring more complex functional roles¹¹.

CLIMP63, p180 and KTN1 position ER

We used CRISPR–Cas9 to knock out these proteins in human U2OS cells stably expressing the ER marker mEmerald–Sec61β (Extended Data Fig. [1a, b](#)). As

previously reported^{11,12}, peripheral ER in CLIMP63-knockout cells is populated with increased numbers of dense matrices or sheets—a ‘dispersed’ phenotype. KTN1 knockout also disperses ER, whereas p180-knockout cells exhibit a contrasting ‘clustered’ ER phenotype, with the peripheral network remaining tubular and perinuclear ER collapsing asymmetrically into a smaller area at one side of the nucleus (Fig. 1a, Extended Data Fig. 1c,d). These morphologic changes are not secondary to alterations in levels of other ER-shaping proteins or cell cycle disruption (Extended Data Fig. 1b,e,f). Double knockout of CLIMP63 and KTN1 substantially disperses ER. Conversely, ER in CLIMP63 and p180 double-knockout cells resembles the wild type, consistent with their opposing single-knockout phenotypes. Surprisingly, p180 and KTN1 double knockout causes more ER clustering than in p180-knockout cells (Fig. 1a, Extended Data Fig. 1d), suggesting a more complex interplay. In CLIMP63–p180–KTN1 triple-knockout cells, high-density ER matrices or sheets are abundant in the perinuclear region (Fig. 1a), although perinuclear ER appears less evenly distributed compared with wild-type cells, with ‘hot spots’ (Extended Data Fig. 1g) that may reflect ER positioning defects.

Fig. 1: CLIMP63, p180 and KTN1 differentially regulate ER morphology.

 **figure1**

a, Representative images of wild-type (WT), knockout (KO), double knockout (DKO) or triple knockout (TKO) of CLIMP63, p180 and/or KTN1 in U2OS cells stably expressing mEmerald–Sec61 β (green, ER marker) and co-labelled with DAPI (blue, nuclear marker), anti- α -tubulin (red, microtubule marker) and anti-TRAP α (magenta, rough ER marker). Perinuclear and peripheral regions (left, outlined) are enlarged on the right. Scale bar, 10 μ m. **b, c**, ER MDR (**b**) and asymmetry (**c**) (see [Supplementary Text](#)) in cells as in **a**. $n = 30$ cells. **d–f**, Quantifications of ER morphology in wild-type or CLIMP63-knockout (**d**), p180-knockout (**e**) or KTN1-knockout (**f**) cells expressing the indicated CLIMP63, p180 or KTN1 constructs. Since recombinant KTN1 levels are typically very low, only cells with detectable KTN1–mApple signals were quantified. Ability (+) or inability (−) of the mutants to bind microtubules (MT) is indicated. $n = 32, 30, 30, 30, 31, 30$ and 30 cells (left to right) in **d**; $n = 59, 62, 64, 61, 34, 70$ and 70 (left to right) in **e**; and $n = 32, 38, 38, 39, 40$ and 36 (left to right) in **f**.

Data are mean \pm s.d. with individual data points shown. Two-tailed *t*-test; *P* values are shown.

Source data.

To quantitatively assess changes in ER morphology and distribution, we devised complementary algorithms. First, we harnessed a statistical approach based on probability density estimation to analyse spatial distributions of fluorescently labelled ER and other organelles. Next, we used an experimentally derived spatial probability mass function, which quantifies fluorescence changes across an image, to calculate the radial distribution and degree of cellular asymmetry of organelles (Extended Data Fig. 2a–g, [Supplementary Text](#)). Single or double knockout of CLIMP63 and KTN1 increases ER mean distribution radius (MDR) (Fig. 1b), indicating that ER is spread more peripherally. By contrast, p180 knockout or p180 and KTN1 double knockout increases ER asymmetry (Fig. 1c). Quantification assessing the rough ER marker TRAP α instead of mEmerald–Sec61 β shows similar results (Extended Data Fig. 2h,i). Microtubule MDR and asymmetry change only slightly (Extended Data Fig. 2j–m).

ER proteins bind subsets of microtubules

We assessed microtubule binding of numerous ER proteins by co-sedimentation. CLIMP63 and p180, both known microtubule-binding proteins^{13,14}, co-sediment with microtubules as expected. KTN1 also sediments robustly with microtubules (Extended Data Fig. 3a,b). Since full-length p180 (p180L) is degraded during cell lysis (Extended Data Fig. 3a,c), we used a smaller, more stable splice variant (p180s) that lacks the numerous ribosome-binding decapeptide repeats present in p180L (Extended Data Figs. 1a, 3b); p180s was undetectable by immunoblotting in the cell lines studied, facilitating identification of recombinant protein (Extended Data Fig. 3d). For each protein, we mapped microtubule-binding domains; only wild-type proteins or mutants capable of microtubule binding restored ER morphology in corresponding knockout cell lines (Fig. 1d–f, Extended Data Figs. 3,4). For instance, CLIMP63 missense mutants R7A, K10A and R70A did not bind microtubules or suppress ER distribution defects in CLIMP63-knockout cells, whereas CLIMP63(H69A), which binds microtubules, rescues the phenotype. A phosphomimetic CLIMP63 mutant defective in microtubule binding¹³ also did not rescue ER distribution defects (Extended Data Fig. 4a–d). For KTN1, only the deletion mutant that binds microtubules suppressed the abnormal ER phenotype (Fig. 1f, Extended Data Fig. 4g–i). Finally, p180s lacking the kinesin-1 binding domain still suppressed the clustered ER phenotype in p180-knockout cells (Extended Data Fig. 4j–l). Thus, despite distinct phenotypes, ER morphology changes in CLIMP63-, p180- and KTN1-knockout cells are likely to all reflect alterations in microtubule binding.

We hypothesized that these proteins bind different microtubule populations and used a proximity ligation assay (PLA) to visualize their microtubule associations in cells (Extended Data Fig. 5a–c). We depleted centrosomal microtubules using centrinone B treatment¹⁵, and Golgi-derived microtubules by knocking down AKAP450¹⁶ (Extended Data Fig. 5d). We found that microtubule association of CLIMP63 was sensitive to centrosome depletion but not Golgi microtubule depletion, whereas KTN1–microtubule association was sensitive to both; p180–microtubule association was not sensitive to depletion of either centrosomes or Golgi microtubules (Extended Data Fig. 5e–h). Admittedly, these microtubule subsets can be interdependent, and centrosome depletion can boost AKAP450-dependent microtubule nucleation at the Golgi^{16,17}. Even so, disrupting Golgi microtubules did not alter centrosome activity¹⁷ or CLIMP63–microtubule association (Extended Data Fig. 5e,f).

We inferred that CLIMP63 preferentially binds centrosomal microtubules, KTN1 preferentially binds perinuclear microtubules derived from either centrosome or Golgi, and p180 preferentially binds more peripheral microtubules regardless of origin. In this scenario, PLA distributions for microtubules with CLIMP63 should be more asymmetric than with p180 or KTN1, and PLA distributions for p180 and microtubules should be more dispersed. Indeed, PLA signals for CLIMP63 and microtubules were more asymmetric than those for p180 and KTN1 with microtubules. However, PLA MDR for p180 with microtubules resembled that for KTN1 with microtubules in wild-type cells (Extended Data Fig. 5i). We reasoned that because ER is densely packed perinuclearly in wild-type cells, PLA signals were also mostly perinuclear, making differences challenging to identify. As a workaround, we quantified PLA distributions in CLIMP63-knockout cells, in which ER is more dispersed (Fig 1a); MDR for p180–microtubule PLA signals was larger than MDR for KTN1–microtubule PLA signals (Extended Data Fig. 5j), suggesting that p180 binds more peripheral microtubules than KTN1. Consistent with this specificity, centrosome depletion led to highly dispersed ER in wild-type but not p180-knockout cells, whereas depletion of Golgi-derived microtubules clustered ER in wild-type but not CLIMP63 and KTN1 double-knockout cells (Extended Data Fig. 5k–n).

Graded binding to modified microtubules

For regulatory specificity, microtubules undergo reversible post-translational modifications including acetylation, detyrosination and glutamylation, which together constitute key elements of the tubulin code². Although CLIMP63, p180 or KTN1 knockout did not affect overall levels of these modifications, tubulin polyglutamylation was decreased in centrosome or Golgi microtubule-depleted cells (Extended Data Fig. 6a–c). We thus considered whether variations in tubulin glutamylation underlie binding selectivity for different microtubule populations and differential effects of the proteins on ER distribution.

CLIMP63 overexpression caused tight ER–microtubule alignment¹³ that is suppressed in centrosome-depleted cells, whereas p180s or KTN1 overexpression did not trigger ER–microtubule alignment (Extended Data Fig. 6d–f). Co-expression of TTLL4, which monoglutamylates microtubules¹⁸ (Extended Data Fig. 6g, h), slightly enhanced ER–microtubule alignment in p180- but not KTN1-overexpressing cells (Extended Data Fig. 6d, e). By contrast, co-expression of TTLL7, which polyglutamylates microtubules^{19,20} (Extended Data Fig. 6g, h), led to significant microtubule–ER alignment in both p180- and KTN1-overexpressing cells (Extended Data Fig. 6d, e). Although co-expression of TTLL7 slightly enhanced ER-microtubule alignment in CLIMP63-overexpressing cells, co-expression with TTLL4 or microtubule de-glutamylases CCP1 or CCP5 (CCP1 shortens glutamate chains, whereas CCP5 is thought to remove the branch-point glutamate²¹, Extended Data Fig. 6g, i) did not influence ER–microtubule alignment (Extended Data Fig. 6d, e). Since CLIMP63–microtubule associations as assessed using PLA were unaffected by overexpression of TTLL4, TTLL7, CCP1 or CCP5 (Extended Data Fig. 6j), we inferred that CLIMP63–microtubule binding is not altered by changes in microtubule glutamylation. PLA signals for KTN1–tubulin were significantly increased by TTLL7 but not TTLL4 and decreased in cells overexpressing CCP1 or CCP5 (Extended Data Fig. 6k). By contrast, PLA signals of p180–tubulin were slightly increased by TTLL4 overexpression, markedly increased by TTLL7, slightly decreased by CCP1, and significantly decreased by CCP5 (Extended Data Fig. 6l). We conclude that KTN1 and p180 respond differentially to glutamylation levels, with KTN1 preferentially associating with polyglutamylated versus monoglutamylated microtubules, whereas p180 broadly associates with mono- and polyglutamylated microtubules.

We purified fragments of p180, KTN1 and CLIMP63 containing their microtubule-binding domains (Extended Data Fig. 6m, n) and investigated binding to differentially glutamylated microtubules in vitro, using TTLL6 to generate microtubules functionalized with polyglutamate chains of various lengths primarily on α -tubulin¹⁸ (Extended Data Fig. 6o). Both p180 and KTN1 showed substantial increases in binding to microtubules polyglutamylated by TTLL6, with only background binding to unmodified microtubules (Fig. 2a, b). Moreover, as the average glutamate number $\langle n_E \rangle$ on α -tubulin increased from 3.5 to 8.3, binding affinities increased in lockstep, with 2.7- and 5.6-fold increases for p180 and KTN1, respectively. Notably, p180 had a 2.9-fold stronger affinity than KTN1 for microtubules with shorter chains (Fig. 2b). Next, we interrogated how microtubule binding is affected by β -tubulin monoglutamylation induced by TTLL4 and polyglutamylation induced by TTLL7^{18,19} (Extended Data Fig. 6p). Both p180 and KTN1 showed binding preferences toward microtubules functionalized with polyglutamates by TTLL7 (Fig. 2c, d) but weaker binding to TTLL4-modified microtubules (Fig. 2c, e). Of note, p180 exhibited higher in vitro binding (3.9-fold) to microtubules monoglutamylated by TTLL4 compared with KTN1, whereas p180 and KTN1 bound similarly to microtubules

polyglutamylated by TTLL7 (Fig. 2c, e). This difference was evident even though numbers of glutamates added by TTLL4 (mean of 1.5) and TTLL7 (mean of 1.2) were similar (Extended Data Fig. 6p), indicating that KTN1 prefers polyglutamate chains introduced by TTLL7 to multiple monoglutamates introduced by TTLL4. At higher glutamylation levels, both KTN1 and p180 formed patches on the microtubule lattice (Fig. 2c), indicative of cooperative binding that may be physiologically relevant in cells when these molecules are tethered and concentrated on the ER membrane. In contrast to p180 and KTN1, CLIMP63 was less responsive to microtubule glutamylation; it lacked detectable binding to unmodified and polyglutamylated microtubules with $\langle n_E \rangle$ of 2.7, exhibiting microtubule binding only when $\langle n_E \rangle$ reached 3.8 (Extended Data Fig. 6q, r). Thus, hyperglutamylation can enhance CLIMP63–microtubule binding in vitro, but since overexpression of CCP1 or CCP5 did not seem to affect CLIMP63–microtubule binding in cells (Extended Data Fig. 6e, j), and centrosome depletion suppressed CLIMP63 overexpression-mediated ER–microtubule alignment (Extended Data Fig. 6f), a different tubulin modification or interaction probably mediates preferential binding of CLIMP63 with centrosome microtubules.

Fig. 2: p180 and KTN1 bind glutamylated and polyglutamylated microtubules, respectively.

 **figure2**

a, Representative micrographs of p180 and KTN1 fragments (cyan) binding to unmodified microtubules or microtubules glutamylated in vitro by TTLL6 (magenta). Internal reflection microscopy images for the microtubule channel were background subtracted and inverted. Average numbers of glutamate molecules (E) added to microtubules as quantified from mass spectroscopy data in each group are indicated. Scale bar, 5 μ m. **b**, Binding of p180 and KTN1 fragments to unmodified and TTLL6-glutamylated microtubules. $n = 110$ (unmodified), 165 ($\alpha + 3.5$ E, $\beta + 1.3$ E) and 112 ($\alpha + 8.3$ E, $\beta + 2.6$ E) microtubules for p180; $n = 141$ (unmodified), 186 ($\alpha + 3.5$ E, $\beta + 1.3$ E) and 156 ($\alpha + 8.3$ E, $\beta + 2.6$ E) microtubules for KTN1. **c**, Representative micrographs of p180 and KTN1 microtubule interacting fragments (cyan) showing binding to unmodified microtubules or microtubules glutamylated by TTLL4 or

TTLL7 (magenta). KTN1 and p180 are shown with different brightness/contrast settings for TTLL4- and TTLL7-modified microtubules, reflecting large differences in binding between mono- and polyglutamylated microtubules. Scale bar, 5 μ m. **d, e**, Affinities of p180 and KTN1 for unmodified microtubules and microtubules glutamylated by TTLL7 (**d**) or TTLL 4 (**e**). The x-axis shows weighted averages of glutamate residues attached to α - and β -tubulin. $n = 185$ (unmodified) and 117 ($\alpha + 0$ E, $\beta + 1.2$ E) microtubules for p180; $n = 128$ (unmodified) and 225 ($\alpha + 0$ E, $\beta + 1.2$ E) microtubules for KTN1; $n = 179$ (unmodified) and 237 ($\alpha + 0$ E, $\beta + 1.5$ E) microtubules for p180; $n = 163$ (unmodified) and 224 ($\alpha + 0$ E, $\beta + 1.5$ E) microtubules for KTN1. **f**, CLIMP63 binds centrosomal microtubules, KTN1 binds perinuclear polyglutamylated microtubules, and p180 can bind peripheral microtubules with less glutamylation. Together, these proteins maintain proper asymmetric ER distribution, which regulates organelle distributions. MT, microtubules. Data are mean \pm s.d. with individual data points shown. Kruskal–Wallis (**b**) or Mann–Whitney tests (**d, e**); P values are shown.

[Source data](#).

Glutamylation regulates ER distribution

Perinuclear microtubules harbour more polyglutamylation, whereas monoglutamylation is generally more prominent peripherally (Extended Data Fig. [7a–d](#)), consistent with KTN1 binding preferentially to perinuclear microtubules and p180 binding to peripheral microtubules. TTLL overexpression glutamylated microtubules throughout the cell, eliminating the relatively discrete perinuclear distribution of polyglutamylated microtubules and thus drawing ER towards the cell periphery, whereas overexpression of CCP1 or CCP5 decreased binding of p180 and KTN1 to perinuclear microtubules; thus, overexpression of TTLL, CCP1 or CCP5 all lead to dispersed ER (Extended Data Fig. [7e–g](#)). In p180 and KTN1 double-knockout cells, overexpression of TTLL4, CCP1 or CCP5 did not change ER MDR (Extended Data Fig. [7h](#)), yet TTLL7 overexpression still had minor effects, possibly through other pathways. When TTLL7 is overexpressed, KTN1 should bind all microtubules, rather than preferring perinuclear ones. Thus, with TTLL7 overexpression, KTN1 knockout resulted in less dispersed ER (Extended Data Fig. [7i](#)). We also knocked down CCP5, which increases tubulin glutamylation. Similar to TTLL4 overexpression, CCP5 knockdown dispersed ER (Extended Data Fig. [7j–m](#)).

We examined several cell lines widely used in ER morphology studies to assess whether they had different microtubule glutamylation levels. Notably, COS7 cells had particularly high polyglutamylation levels (Extended Data Fig. [7c](#)), and although polyglutamylation in COS7 cells remained relatively more perinuclear compared to monoglutamylation and microtubule distribution, the difference was much less than in

U2OS cells (Extended Data Fig. 7*b, d*). We hypothesized that KTN1 knockout in COS7 cells would show a distinct ER phenotype, possibly mimicking TTL7 overexpressing cells (Extended Data Fig. 7*i*). Indeed, although knockout of CLIMP63 or p180 in COS7 cells showed similar phenotypes as in U2OS cells, KTN1 knockout in COS7 cells led to clustered ER (Extended Data Fig. 8*a–d*), in contrast to dispersed ER in KTN1 knockout U2OS cells. Moreover, overexpression of CCP6 (which has similar activity to CCP1) also led to clustered ER (Extended Data Fig. 8*e–h*). We conclude that CLIMP63, p180 and KTN1 preferentially bind centrosomal, polyglutamylated and glutamylated microtubules, respectively, to cooperatively distribute ER (Fig. 2*f*).

Organelle positioning and glutamylation

Live imaging of six organelles²² simultaneously revealed that most have a distribution similar to ER (Extended Data Fig. 9*a*), suggesting that ER might broadly regulate organelle distribution. Notably, in CLIMP63-, p180- and KTN1-knockout cells, all organelles that we examined exhibited similar distribution changes to those of ER—more dispersed in CLIMP63- or KTN1-knockout cells and more asymmetric in p180-knockout cells (Extended Data Fig. 9*b–d*). Moreover, CCP1 overexpression, which disperses ER, also increased MDR for lysosomes, mitochondria and peroxisomes in wild-type cells but not in p180 and KTN1 double-knockout cells (Extended Data Fig. 9*e–g*). Thus, perinuclear ER morphology specifies the distributions of other organelles downstream of microtubule glutamylation.

ER and lysosome movements in autophagy

Perinuclear lysosome clustering, a signature event in early autophagy, is important for proper autophagic flux^{23,24}. Similar to lysosomes, ER migrates perinuclearly during early autophagy, and subsequently redistributes to the periphery (Fig. 3*a, b*, Supplementary Video 1). CLIMP63 levels increased significantly during early autophagy, and this increase did not appear to require new protein synthesis or inhibition of lysosomal or proteasomal degradation (Fig. 3*c*, Extended Data Fig. 10*a*). CLIMP63 knockout prevented ER movement toward the perinuclear region (Fig. 3*b*) and suppressed autophagosome–lysosome fusion and autophagic degradation, but not lysosomal activity (Extended Data Fig. 10*b–f*). Since p180 and KTN1 protein levels remained unchanged (Extended Data Fig. 10*a*), we examined their binding to microtubules. KTN1–microtubule binding did not change upon nutrient starvation, but p180–microtubule binding increased (Fig. 3*d*). Consistently, ER and lysosomes in p180-knockout cells remained perinuclear (Fig. 3*b*), and thus p180-knockout cells showed defects in recovery of mTOR signalling²⁴ after nutrient re-supplementation, but not in autophagic degradation (Extended Data Fig. 10*g, h*).

Fig. 3: ER distribution changes during autophagy.

 figure3



a, U2OS cells stably expressing mEmerald–Sec61 β (green, ER marker) were starved in EBSS for 0, 0.5, or 2 h and labelled with anti-LAMP1 (red, lysosome marker). Scale bar, 10 μ m. **b**, ER and lysosome distribution in wild-type, CLIMP63- or p180-knockout cells starved in EBSS for the indicated times. $n = 55, 53, 52, 52, 53, 54, 51, 50, 51, 50, 46$ and 44 cells (left to right) for both ER and lysosomes. **c**, Top, U2OS cells were starved in EBSS for 0–2 h and immunoblotted. Bottom, protein levels relative to α -tubulin. $n = 5$ experiments. **d**, Average intensities of PLA for p180 and KTN1 with α -tubulin at 0, 0.5, or 2 h of EBSS starvation or puromycin treatment (2 μ g ml $^{-1}$ for 2 h). $n = 273, 254, 188, 180, 143, 144, 182$ cells (left to right). **e**, Schematic of p180 domain composition. The ribosome-binding domain includes 41 positively charged decapeptide repeats, which can potentially bind microtubules once ribosomes are dissociated. Amino acid sequences for several repeats are shown, with positively charged residues in red and negatively charged residues in green. Repeat sequences

vary slightly but are all positively charged. **f**, Under normal cellular conditions, p180L ribosome-binding repeats are occupied by ribosomes and cannot bind microtubules. When cells are starved, ribosomes dissociate from ER, and the repeats can then bind microtubules. **g**, PLA for p180 and RPL3 (ribosome marker) upon starvation or puromycin treatment ($2 \mu\text{g ml}^{-1}$ for 2 h). $n = 230$ cells. Data are mean \pm s.d. with individual data points shown. One-way ANOVA followed by Dunnett's multiple comparisons test for (b), (c); two-sided *t*-test for (d), (g); *P* values are shown. See [Supplementary Information](#) for uncropped western blots.

[Source data](#).

Microtubule modification levels were unaffected by starvation (Extended Data Fig. [10i](#)). Notably, the ribosome-binding region of p180L (the major cellular isoform) includes 41 positively charged decapeptide repeats (Fig. [3e](#)). We hypothesized that this region is occupied by ribosomes under normal cellular conditions but then ribosomes dissociate during starvation, exposing these positively charged regions that can then bind microtubules (Fig. [3f](#)). Indeed, starvation significantly decreased p180–ribosome binding (Fig. [3g](#), Extended Data Fig. [10j](#)); puromycin treatment, which dissociates ribosomes from ER in fed conditions, markedly enhanced p180–microtubule binding (Fig. [3d](#)). In contrast to p180s, which lacks most ribosome-binding decapeptide repeats, p180L overexpression increased ER–microtubule alignment. This alignment was enhanced by starvation or puromycin treatment (Extended Data Fig. [10k–m](#)), further indicating that ribosome-binding repeats of p180L bind microtubules upon ribosome dissociation.

Discussion

Peripheral ER network morphology is maintained by hydrophobic hairpin domain proteins (reticulons and receptor expression enhancing proteins (REEPs)) that shape the tubules. The polygonal network is generated via atlastin-mediated tethering and fusion of tubules at three-way junctions and distributed via cytoskeletal interactions^{[3,25,26](#)}. Much less is known about the dynamic organization of perinuclear ER. Although microtubules have important roles in establishing ER morphology, most studies have emphasized peripheral tubular ER^{[27](#)} or identification of ER proteins that bind microtubules^{[13,14,28](#)}. Proteins including CLIMP63, p180 and KTN1 are enriched in dense, sheet-like perinuclear ER, and they each bind microtubules. However, phenotypes of cells deficient in these proteins differ considerably, raising the question of how microtubule-binding specificity is maintained. Here we demonstrate that CLIMP63, p180 and KTN1 preferentially bind different subsets of microtubules to maintain perinuclear ER in its characteristic distribution, explaining the differential effects of their absence. Furthermore, depletion of centrosome or Golgi-derived microtubules has distinct effects on the microtubule binding of these three proteins.

Microtubule diversity can be achieved via different tubulin gene products, differential interactions with microtubule-associated proteins and numerous post-translational modifications². Modifications are dynamic and rapidly reversible, but evidence for how they affect microtubule-related functions has been limited². We have shown here that KTN1 preferentially binds perinuclear polyglutamylated microtubules with long glutamate chains, whereas p180 binds glutamylated microtubules with either short or long chains. By contrast, CLIMP63 has a higher threshold for response to microtubule glutamylation. We cannot exclude that increased affinity at higher glutamate numbers for TTLL7-modified microtubules stems from additional chains that TTLL7 initiates on tubulin tails, and not only from introduction of longer chains. Conversely, p180 is more sensitive to any increase in glutamate numbers on the tubulin tail and robustly binds both mono- and polyglutamylated microtubules. This differential effect on microtubule binding according to glutamylation state has previously been observed for the microtubule-severing ATPase spastin²⁹ and may represent a general feature of this modification, enabling fine tuning of molecular interactions. Thus, a small difference in the number of glutamates added to tubulin side chains may exert a substantial qualitative effect on ER distribution. Other ER-localized, microtubule-binding proteins³⁰ are likely to contribute to overall cellular ER positioning. Indeed, even in p180 and KTN1 double-knockout cells, TTLL7 overexpression still disperses ER, suggesting the involvement of other ER proteins. Moreover, tubular ER selectively moves along acetylated microtubules²⁷, further indicating that ER distribution is broadly sensitive to microtubule modifications.

The ability of cells to dynamically control ER distribution through differential microtubule modifications has important functional implications. For instance, p180 regulates microtubule remodelling in axons³¹, and axonal microtubules are highly glutamylated². Thus, p180 may affect microtubule remodelling by differentially recognizing glutamylated axonal microtubules. Of note, although dysregulation of ER shaping and microtubule polyglutamylation lead to different neurodegenerative diseases^{32,33}, these diseases share some similar cellular phenotypes, including mitochondrial distribution defects and axon degeneration, suggesting possible convergence.

When ER positioning is disrupted, distributions of other organelles are affected. Microtubules have key roles in organelle distribution¹, and their ability to selectively distribute organelles relies on a tubulin code. Our results indicate that ER distribution is mediated via specific membrane-bound proteins with differential binding to different levels and types of microtubule glutamylation, broadly affecting distributions of most other organelles. ER thus interprets the tubulin code to regulate movement and positioning of cellular organelles. Rather than imbuing each organelle with its own sensing and response mechanisms, cells achieve organizational efficiency by using ER as a first-line sensor and responder. This role is exemplified during nutrient starvation,

when cells increase CLIMP63 protein levels to move ER towards the perinuclear region, which also clusters lysosomes for efficient autophagic degradation. Then, cells harness enhanced p180–microtubule binding to redistribute ER and lysosomes for a proper reset. There are likely to be other ER proteins that also decipher the tubulin code, with important implications for ER function in health and disease.

Methods

Plasmids and reagents

GFP-mCherry-LC3 was a gift from Juan S. Bonifacino. mEmerald-Sec61 β , pcDNA3.1_CLIMP63-HA, mApple-SiT (Golgi apparatus), mKO-SKL (peroxisome), YFP-KDEL (ER) and CFP-LAMP1 (lysosome) were constructed as described previously^{6,7,18,21}. pCMV6_p180s-myc-Flag (RC218816), pCMV6_KTN1-myc-Flag (RC219832), pCMV6_TTLL4-myc-Flag (RC205206) and pCMV6_CCP1-myc-Flag (RC220826) were obtained from Origene Technologies. pcDNA3.1_TTLL6-Flag (OHu07095), pcDNA3.1_CCP5-Flag (Ohu28493), pcDNA3.1_CCP6-Flag (OHu24335) and pcDNA3.1_p180L (OHu24745) were obtained from GenScript. Mutants of CLIMP63, p180s and KTN1 were generated in a pcDNA3.1(+) vector with a C-terminal HA-epitope tag. Specifically, Ser-to-Glu mutants (S3E, S17E, S19E) of CLIMP63 were synthesized by GenScript. pcNDA3.1-KTN1, pN1-KTN1-mApple, pC1-sp-mScarlet-KTN1, pN1-KTN1-mNeonGreen, 2×Strep-p180s_29-381-mNeonGreen, 2×Strep-p180s_29-381-mNeonGreen and CLIMP63-mNeonGreen-2×Strep were also constructed using standard cloning procedures. 5' and 3' UTR of KTN1 were synthesized by Integrated DNA Technologies and ligated into the pN1-KTN1-mNeonGreen vector. Human TTLL4, TTLL6, TTLL7, CCP1, CCP5 and CCP6 inserts were also cloned into pCMV14-3×Flag and pN1-mApple. DNA oligonucleotides were synthesized by Integrated DNA Technologies.

LD540 was provided by C. Thiele³⁴. Camptothecin (S1288), DC661 (S8808), etoposide (S1225) and MG132 (S2619) were purchased from Selleckchem. LysoSensor Green DND 189 (L7535), G418 (11811098) and puromycin (A1113802) were from Thermo Fisher Scientific. Centrinone B (CNB, CN5690) was from Tocris. GTP (G8877), ATP, Taxol (T7402), tunicamycin (SML1287), Earle's Balanced Salts (EBSS, E2888) and Duolink In Situ PLA kit (DUO92002, DUO92004, DUO92008, DUO92013) were from Sigma-Aldrich. The Cathepsin L Activity Assay Kit (Fluorometric) (ab65306) was obtained from Abcam.

Cell culture and transfections

All cell lines were obtained from the American Type Culture Collection: HEK 293T (CRL-11268), COS7 (CRL-1651), HeLa (CCL-2), RPE1 (CRL-4000) and U2OS (HTB-96) cells. HEK 293T, COS7 and HeLa cells were cultured in Dulbecco's Modified Eagle Medium (DMEM, Thermo Fisher Scientific 11995065), RPE1 cells were cultured in DMEM/F12 (1:1) Medium (Thermo Fisher Scientific, 11330-057), and U2OS cells were cultured in McCoy's 5A medium (Thermo Fisher Scientific 16600108); all were supplemented with 10% fetal bovine serum (FBS, Thermo Fisher Scientific 2614007) and 1×penicillin/streptomycin/amphotericin B (Thermo Fisher Scientific 15240112) at 37 °C with 5% CO₂. HEK 293T cells were transfected with Avalanche-Everyday Transfection Reagent (EZ Biosystems, EZT-EVDY-1). U2OS and COS7 cells were electroporated using Cell Line Nucleofector Kit V (for U2OS; Lonza, VVCA-1003) and Cell Line Nucleofector Kit R (for COS7; Lonza, VVCA-1001) following the manufacturer's instructions. Amounts of key plasmids transfected are (per 1 × 10⁶ U2OS cells or 5 × 10⁵ COS7 cells): 0.3 µg CLIMP63-HA, 0.5 µg p180s-HA, 2 µg KTN1-mApple, 1 µg CLIMP63-mEmerald (for overexpression), 2 µg p180s-mEmerald, 4 µg p180L-mEmerald, 4 µg KTN1-mEmerald, as well as (in various vectors) 1 µg TTLL4, 1 µg TTLL6, 0.3 µg TTLL7, 4 µg CCP1, 0.5 µg CCP5, and 1 µg CCP6.

For RNAi knock down of AKAP450, two siRNAs targeting ATATGAACACAGCTTATGA and AACTTGAAAGTTAACTATCAA were synthesized by Eurofins Genomics. Cells were transfected using Avalanche-Omni Transfection Reagent (EZ Biosystems, EZT-OMNI-1) with 20 pmol siRNA for 3 days. For RNAi knockdown of CCP5, ON-TARGETplus siRNA sets targeting human CCP5 were purchased from Horizon Discovery (LQ-009468-00-0005), and 60 pmol siRNAs were transfected per 1 × 10⁶ U2OS cells using Lonza Cell Line Nucleofector Kit V.

Antibodies

Primary antibodies used: mouse monoclonal anti-AKAP450 (BD Biosciences, 611518, Clone 7/AKAP450, immunoblot 1:250), rabbit polyclonal anti-Atlastin2 (Bethyl Laboratories, A303-333A, immunoblot 1:500), rabbit polyclonal anti-Atlastin3 (Proteintech, 16921-1-AP, immunoblot 1:1,000), rabbit monoclonal anti-Catalase (Cell Signaling Technology, 12980, clone D4P7B, immunofluorescence 1:800), mouse monoclonal anti-Climp63 (Enzo, ALX-804-604, clone G1/296, immunofluorescence 1:500 immunoblot 1:5,000), mouse monoclonal anti-Flag M2 (Sigma-Aldrich, F1804, clone M2, immunoblot 1:1,000), rabbit polyclonal anti-GFP (MBL, 598, immunoblot 1:5,000, immunofluorescence 1:500), mouse monoclonal anti-GM130 (BD Biosciences, 610822, Clone 35/GM130, immunofluorescence 1:200), rabbit polyclonal anti-GM130 (Proteintech, 11308-1-AP, immunofluorescence 1:200), mouse monoclonal anti-HA (Covance, MMS-101P, clone 16B12, immunofluorescence 1:500, immunoblot 1:5,000), rabbit polyclonal anti-kinectin (Proteintech, 19841, immunoblot

1:2,000), rabbit monoclonal anti-kinectin (Cell Signaling Technology, 13243, clone D5F7J, immunofluorescence 1:100), mouse monoclonal anti-Lamp1 (DSHB, clone 1D4B, immunofluorescence 1:2,000), rabbit polyclonal anti-LC3 (Cell Signaling Technology, 4108, immunofluorescence 1:200, immunoblot 1:1,000), rabbit polyclonal anti-Lunapark (Sigma-Aldrich, HPA014205, immunoblot 1:250), mouse monoclonal anti-Myc (Santa Cruz, sc-40, clone 9E10, immunoblot 1:2,000), rabbit polyclonal anti-p180 (Thermo Fisher Scientific, PA5-21392, immunofluorescence 1:500, immunoblot 1:5,000), rabbit polyclonal anti-Pericentrin (Abcam, ab4448, immunofluorescence 1:1,000), rabbit polyclonal anti-polyglutamylation (polyE) (AdipoGen, AG-25B-0030, immunofluorescence 1:200, immunoblot 1:1,000), mouse monoclonal anti-glutamylation clone GT335 (AdipoGen, AG-20B-0020, immunofluorescence 1:200, immunoblot 1:200), rabbit polyclonal anti-REEP2 (Proteintech, 15684, immunoblot 1:3,000), rabbit polyclonal anti-REEP3 (Abcam, ab106463, immunoblot 1:1,000), rabbit polyclonal anti-REEP4 (Proteintech, 26650, immunoblot 1:1,000), rabbit polyclonal anti-REEP5 (Proteintech, 14643, immunoblot 1:1,000), rabbit polyclonal anti-reticulon3 (Proteintech, 12055, immunoblot 1:2,000), rabbit polyclonal anti-reticulon4 (Proteintech, 10740, immunoblot 1:1,000), rabbit polyclonal anti-RPL3 (Proteintech, 66130, immunofluorescence 1:100), rabbit polyclonal anti-TOM20 (Santa Cruz, sc-11415, immunofluorescence 1:1,000), mouse monoclonal anti-TOM20 (BD Biosciences, 612278, Clone 29/Tom20, immunofluorescence 1:1,000), rabbit polyclonal anti-TRAP α (Proteintech, 10583, immunofluorescence 1:50), rat monoclonal anti- α -tubulin Alexa Fluor 647 (Abcam, ab195884, clone YOL1/34, immunofluorescence 1:50), mouse monoclonal anti- α -tubulin (Proteintech, 66031, clone 1E4C11, immunofluorescence 1:1,000, western blot 1:10,000), mouse monoclonal anti- β -tubulin (Proteintech, 66240, clone 1D4A4, immunofluorescence 1:1,000). Alexa Fluor 405/488/568/633 conjugated goat anti-rabbit/mouse IgG (H+L) highly cross-adsorbed secondary antibodies were from Thermo Fisher Scientific. HRP-conjugated goat anti-mouse or anti-rabbit secondary antibodies were from Santa Cruz Biotechnology.

Stable cell lines

To generate U2OS cells stably expressing mEmerald-Sec61 β , cells were transfected with the mEmerald-Sec61 β ⁷ and selected using 200–1,000 $\mu\text{g }\mu\text{l}^{-1}$ (gradually increasing) G418 for two weeks; green-positive cells were sorted into mono-clones by flow cytometry using a MoFlo Astrios cell sorter (Beckman Coulter) and cultured in the presence of 200 $\mu\text{g }\mu\text{l}^{-1}$ G418 for 2–3 weeks. Proliferated clones were verified by immunoblotting and fluorescence imaging.

CRISPR–Cas9 gene editing

All CRISPR–Cas9 knockout assays used eSpCas9(1.1)³⁵. The targets used were:
CLIMP63:GCCGCGCCGCCATGCCCTCGG; p180 in
U2OS: GGTGTCGACTTCTCCATGAAGG; p180 in COS7:
GACACCAGGAAGATGCCAATGG; KTN1: GAAAAGCCAGAAGAAGAGG and
GTTAGGGAAAGAAAAAAGAAGG.

For knock-in of CLIMP63, the same target as in CLIMP63 knockout was used, and a PCR fragment with 37 bp homology arms on each side of the mEmerald-coding sequence was used as a homologous recombination template as follows: for CLIMP63:

CCAGCCCCGGCCCGAGCCGCCGCCATGGTGAGCAAGGG
CGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCAGCTGGACGGCGAC
GTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTA
CGGCAAGCTGACCCCTGAAGTTCATCTGCACCACCGGAAGCTGCCCGTGCCT
GGCCCACCCCTCGTGACCACCTGACCTACGGCGTGCAGTGCTCGCCCGCTAC
CCCGACCACATGAAGCAGCACGACTTCAAGTCCGCCATGCCCGAAGGCTAC
GTCCAGGAGCGCACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCGC
CGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCGCATCGAGCTGAAGGGC
ATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAGCTGGAGTACAAC
AACAGCCACAAGGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGTG
AACTTCAAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTGCCGACCA
CTACCAGCAGAACACCCCCATCGGCACGGCCCGTGCTGCTGCCGACAACC
ACTACCTGAGCACCCAGTCCAAGCTGAGCAAAGACCCAACGAGAACGCGAT
CACATGGCCTGCTGGAGTTCGTGACCGCCGCCGGATCACTCTCGGCATGGA
CGAGCTGTACAAGtccggactcagatctcgagctcaagttcgaattctgcagtcacggtaaccggggccccgg
gatccCCCTCGGCCAAACAAAGGGGCTCCAAGGGCGGCCACG; for
calreticulin:

GGCCTCCTCGGCTTGGCCGCCGTCAGGCCGCCATGGTGAGCAAG
GGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCAGCTGGACGGCG
ACGTAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCAC
CTACGGCAAGCTGACCCCTGAAGTTCATCTGCACCACCGGAAGCTGCCCGTGC
CCTGGCCCACCCCTCGTGACCACTTGAACCTACGGCGTGCAGTGCTGCCCGC
TACCCCGACCACATGAAGCAGCACGACTTCAAGTCCGCCATGCCGAAGGC
TACGTCCAGGAGCGCACCATCTTCAAGGACGACGGCAACTACAAGACCCGC
GCCGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCGCATCGAGCTGAAGG
GCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAGCTGGAGTACAAC
TACAACAGCCACAAGGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAG
GTGAACCTCAAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTGCCGA
CCACTACCAGCAGAACACCCCCATCGGCACGGCCCGTGCTGCTGCCGACA
ACCAACTACCTGAGCACCCAGTCCAAGCTGAGCAAAGACCCAACGAGAACGCGC
GATCACATGGCCTGCTGGAGTTCGTGACCGCCGCCGGATCACTCTCGGCAT

GGACGAGCTGTACAAGGAGCCGCCGTCTACTTCAAGGAGCAGTTCTG
GAC (in which bold denotes homology arms; italic denotes mEmerald coding sequence; and lowercase denotes linker). To generate mEmerald-calreticulin knock-in COS7 cells, wild-type Cas9 with a gRNA targeting the end of the signal sequence of calreticulin (*GAGCCCGCCGTCTACTTCAAGG*) was selected, and a PCR fragment with 36 bp homology arms on each side of the mEmerald-coding sequence was used as a homologous recombination template as follows:

Note that amino acids 18–20 (EPA) were appended to both sides, acting as a linker.

Centrosome depletion

To deplete the centrosome, cells were treated with 125 µM CNB for 1 week as described¹⁵ before further analysis.

Western blotting

Cells were quickly rinsed with PBS, directly lysed with sample buffer (50 mM Tris, pH 6.8, 1 mM DTT, 10% glycerol, 2% SDS, 0.1% Bromophenol blue), and boiled for 5 min. Proteins were then resolved by SDS-PAGE using Mini-PROTEAN TGX Precast Protein Gels (Bio-Rad Laboratories) and transferred to nitrocellulose membranes using the Trans-Blot Turbo RTA Midi Nitrocellulose Transfer Kit (Bio-Rad Laboratories) following the manufacturer's instructions. Membranes were blocked with 4% milk in TBST (20 mM Tris, pH 7.4, 150 mM NaCl, 0.1% Tween-20), and incubated with primary antibody (diluted in blocking buffer) at 4 °C overnight. After washing with TBST, membranes were incubated with secondary antibody at room temperature for 2 h, followed by intensive washing with TBST. Immunoreactive proteins were visualized with GE Healthcare LS ECL Prime Western Blotting Detection Reagent (RPN2236) and imaged using a ChemiDoc XRS+ (Bio-Rad). Band intensities were quantified using Fiji software (NIH).

Immunofluorescence and imaging

Cells were fixed with 4% paraformaldehyde in PBS (Lonza) for 30 min at room temperature and permeabilized with 0.1% Triton X-100 in PBS for 10 min.

Alternatively, for immunostaining of glutamylation (GT335) and polyglutamylation (polyE), cells were fixed and permeabilized with cold methanol for 5 min at -20 °C. Then, after blocking with 3% BSA for 30 min, cells were immunostained with polyE antibody at 4 °C overnight, then with polyE and GT335 together at 4 °C overnight, followed by secondary antibody staining at room temperature for 1 h, and finally with anti- α -tubulin Alexa Fluor 647 at room temperature for 2.5 h. For staining of lipid droplets with LD540 dye, cells were incubated with 0.1 µg ml⁻¹ LD540 in PBS for 5

min. Cells were mounted using Fluoromount-G (SouthernBiotech) and imaged using a Zeiss LSM880 confocal microscope in Airyscan mode equipped with a 63×1.4 NA Plan-Apochromat oil objective (Carl Zeiss). Images were acquired using ZEN software (Carl Zeiss) and processed with ZEN software or Fiji (NIH).

Quantification of ER distribution

Three-dimensional images were acquired using a Zeiss LSM880 confocal microscope in Airyscan mode and reconstructed using ZEN software (Zeiss Microscopy). Summed intensity projections were generated using floating point notation to carry precision. A custom macro in Fiji-ImageJ was used to define the centre of the nucleus and remove the signal of neighbouring cells to avoid perturbing the results. From the manually defined centre, a radius was drawn out past the furthest point on the cell and swept through 360° in 0.1° steps, taking a line profile each time and rescaling the data to correct for artifacts generated by the square shape of the pixels. The resulting data represents an (r,θ) -space representation of the cell's fluorescence distribution. For analysis referring to 'normalized' data, we account for the shape of the cytoplasm by finding the radius at each angle where the nuclear envelope and the edge of the cell are located. The fluorescence data were then rescaled to a normalized axis with the cytoplasm between the nuclear envelope and the cell periphery scaled from 0 to 100%. The nucleoplasm is scaled to stretch between -25 and 0, as a control. (Note that, in this 2D implementation, the nucleoplasm also contains the regions of cytoplasm and nuclear envelope above and below the nucleus).

The MDR and asymmetry of each compartment were calculated using custom Matlab scripts as described in the [Supplementary Text](#). Where true values are given by integrals over space, the value was estimated at the resolution limit of the microscope using a sum over the pixels.

Microtubule co-sedimentation assay

To test the microtubule-binding affinities of CLIMP63, p180 and KTN1, cells were lysed in PIPES buffer (80 mM PIPES, pH 6.8, 1 mM MgCl₂, 1 mM EGTA, 100 mM NaCl, 1% Triton X-100, plus Complete protease inhibitors) for 30 min on ice. Cell lysates were centrifugated twice at 20,000g for 20 min at 4 °C. The supernatant was supplemented with 1 mM GTP and 40 µM Taxol and incubated at 4 °C or 37 °C for 30 min for tubulin polymerization before centrifugation at 20,000g for 30 min at 4 °C or 37 °C, respectively. The resulting pellets (P) and supernatants (S) were collected and subjected to immunoblot analysis. In some experiments, only the pellets and supernatants of the 37 °C samples are shown.

Proximity ligation assay

PLA (Sigma-Aldrich, DUO92101) was performed according to the manufacturer's instructions. Samples were observed under a Zeiss LSM880 confocal microscope with a 20×1.0 NA objective using the Airyscan function. The total intensity of the PLA signal per cell was quantified using Fiji software.

Protein purification

Deletion fragments of p180 (short isoform NM_001042576, residues 29–381) and KTN1 (NM_001079521, residues 29–400) as well as full-length CLIMP63 were expressed as fusions with mNeonGreen-2 \times Strep in HEK 293T cells. 48 h post-transfection, cells were lysed in PBS (Lonza) plus 500 mM NaCl, 1% Triton X-100, and protease inhibitors and then centrifugated at 30,000g at 4 °C for 30 min.

Supernatants were combined with Strep-Tactin XT beads (IBA Lifesciences) and rotated gently for 3 h. After extensive washing with lysis buffer (PBS plus 500 mM NaCl and 1% Triton X-100) and then wash buffer (IBA Lifesciences), bound proteins were eluted with Strep-Tactin XT Elution Buffer (IBA Lifesciences). Eluted proteins were subjected to multiple rounds of PBS dilution and concentration using 10 kDa protein concentrators (Sigma-Aldrich), before being aliquoted and frozen in liquid nitrogen.

TIRF-based assays for protein binding to differentially glutamylated microtubules in vitro

Unmodified human tubulin was purified from tsA201 cells as described previously³⁶. TTLL4 and TTLL6 were expressed in *Escherichia coli* and purified as previously described¹⁸. TTLL7 was also expressed in *E. coli* and purified as previously described¹⁹. Taxol-stabilized microtubules were polymerized out of 98.5% unmodified tubulin and 1.5% biotinylated brain tubulin^{36,29} (Cytoskeleton T333P). Unmodified microtubules were modified using TTLL4, TTLL7 or TTLL6 at 1:10 molar ratio of enzyme to tubulin at room temperature in 20 mM HEPES (pH 7.0), 50 mM NaCl, 10 mM MgCl₂, 1 mM glutamate, 1 mM ATP, 0.5 mM TCEP, and 10 μM Taxol for 4.5 h for TTLL4, between 20 min and 2 h for TTLL7, and between 7.5 and 22 h for TTLL6. Control microtubules were incubated with the enzymes under the same conditions but with aspartate, which is not a substrate for TTLL glutamylases, instead of glutamate. Enzymes were removed through a high-salt wash as previously described²⁹. The extent of glutamylation was determined by liquid chromatography–electrospray mass spectrometry^{36,29} (LC–MS). The spectra display the characteristic distributions of masses with peaks separated by 129 Da, which corresponds to one glutamate (Extended Data Fig. 6o,p). The extent of tubulin glutamylation on α- or β-tubulin was determined by calculating the weighted average of peak intensities for each tubulin species present²⁹.

For microtubule-binding assays, microtubules were immobilized in chambers made of silanized glass³⁷ using Neutravidin (Thermo Fisher Scientific). Next, a solution containing 60 mM Pipes (pH 6.8), 0.7 mM MgCl₂, 0.7 mM EGTA, 50 mM KCl, 10 mM 2-mercaptoethanol, 10 µM Taxol, 1% F127 Pluronic, 1.4 mg/ml casein, 20 mM glucose, glucose oxidase, and catalase was flushed into the chamber, followed by the same solution containing 4.7 nM mNeon-labeled p180, KTN1 or CLIMP63. Images were acquired after allowing for equilibration for 5 min at room temperature using total internal reflection fluorescence (TIRF) microscopy at an exposure of 100 ms for the GFP channel. Unlabelled microtubules were visualized using interference reflection microscopy³⁸. Multiple fields of view were imaged. Background corrected line scan average intensities were measured using Fiji software. Multiple chambers were quantified for each condition.

Real-time PCR

Total mRNA were extracted using TRIzol (Thermo Fisher Scientific 15596018) and Direct-zol RNA Miniprep (Zymo Research, R2052), then reverse-transcribed using the SuperScript IV First-Strand Synthesis System (Thermo Fisher Scientific, 1809105). Real-time PCR primers, designed by a free online tool developed by Integrated DNA technologies, were as follows: CCP5-RT: GACTGCCAGGAAGTGCTAAA and AGGAGCTCCCGATGGTAATA; GAPDH-RT: GGTGTGAACCATGAGAAGTATGA and GAGTCCTTCCACGATACCAAAG.

Real-time PCR was performed using Applied Biosystems PowerUp SYBR Green Master Mix (Thermo Fisher Scientific, A25780) with Applied Biosystems QuantStudio 6 Flex real-time PCR instrument. Data were collected and analysed in QuantStudio Real-time PCR Software and Microsoft Excel using the $\Delta\Delta C_t$ method.

Multispectral imaging

Multispectral imaging was performed as described previously²². Images were acquired with a Zeiss LSM880 confocal microscope equipped with a 32-channel multi-anode spectral detector (Carl Zeiss) using a 63×/1.4 NA objective lens, at 37 °C and with 5% CO₂. Fluorophores were excited simultaneously using 458, 514 and 594 nm lasers and a 458/514/594 nm beam splitter, with images collected onto a linear array of 32 photomultiplier tube elements in λ mode at 9.7 nm bins from 468 to 687 nm. Spectra were defined by imaging singly labelled cells for each of the fluorophore reporters, using the same acquisition and laser settings as for multiply labeled cells. Multispectral images were unmixed using the linear unmixing package in ZEN (Carl Zeiss).

Measurements of autophagosome–lysosome fusion and lysosome activity

For autophagosome–lysosome fusion assessments, U2OS cells were transfected with GFP-mCherry-LC3 for 24 h, treated with EBSS for 2 h before fixation with 4% paraformaldehyde in PBS, and imaged using a Zeiss LSM880 confocal microscope in Airyscan mode equipped with a 63×1.4 NA Plan-APOCHROMAT oil objective (Carl Zeiss). A z-projection was performed using maximum projection before quantification. The mCherry-positive vesicles indicate autophagosomes already fused with lysosomes, as the GFP signal would be quenched by the acidic environment of lysosomes; vesicles with both GFP and mCherry fluorescence indicate autophagosomes not yet fused with lysosomes. Quantifications of these two types of vesicles were performed manually using Fiji software.

For lysosome acidification assays, U2OS cells were labeled with 1 μM LysoSensor Green DND 189 for 4 min and immediately imaged within one minute with a Zeiss Axio microscope using a $20\times/0.4$ NA objective. Images were captured with ZEN software, and total intensities of each cell were quantified in Fiji.

Cathepsin L activity assays were carried out using the Abcam Cathepsin L Activity Assay kit (Fluorometric; ab65306) following the manufacturer's instructions; 1×10^6 cells were assayed in each sample.

Statistics and reproducibility

No statistical method was used to predetermine sample size. All groups were randomly assigned and every group represents a distinct treatment or condition. Data were not analysed in a double-blinded manner. All comparisons were performed using Graphpad Prism or Microsoft Excel software. Data are expressed as means \pm s.d., P values are shown on top of the corresponding columns, as determined by one-way ANOVA followed by Dunnett's multiple comparisons test, Mann–Whitney test, Kruskal–Wallis test or by unpaired two-sided t -test as indicated in the figure legends. When representative images are shown, at least three repeats were performed except for Extended Data Figs. [1b](#), [3a–g](#), [5a](#), d, [6a](#), m, n, [8a](#), [e](#), for which repeats are not necessary because they represent sequential sequence mapping data that build upon one another or else they show representative knockdown or knockout efficiencies that can be further established by the resulting cellular phenotypes.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

All research materials are available upon request. [Source data](#) are provided with this paper.

Code availability

Computer algorithms can be accessed at
<https://github.com/cjobara/ProbabilityDensityIntegrator>.

References

1. 1. Barlan, K. & Gelfand, V. I. Microtubule-based transport and the distribution, tethering, and organization of organelles. *Cold Spring Harb. Perspect. Biol.* **9**, a025817 (2017).
2. 2. Roll-Mecak, A. The tubulin code in microtubule dynamics and information encoding. *Dev. Cell* **54**, 7–20 (2020).
3. 3. Goyal, U. & Blackstone, C. Untangling the web: mechanisms underlying ER network formation. *Biochim. Biophys. Acta* **1833**, 2492–2498 (2013).
4. 4. Phillips, M. J. & Voeltz, G. K. Structure and function of ER membrane contact sites with other organelles. *Nat. Rev. Mol. Cell Biol.* **17**, 69–82 (2016).
5. 5. Blackstone, C. Hereditary spastic paraparesis. *Handb. Clin. Neurol.* **148**, 633–652 (2018).
6. 6. Zheng, P. et al. DNA damage triggers tubular endoplasmic reticulum extension to promote apoptosis by facilitating ER–mitochondria signaling. *Cell Res.* **28**, 833–

854 (2018).

7. 7.

Nixon-Abell, J. et al. Increased spatiotemporal resolution reveals highly dynamic dense tubular matrices in the peripheral ER. *Science* **354**, aaf3928 (2016).

8. 8.

Rowland, A. A., Chitwood, P. J., Phillips, M. J. & Voeltz, G. K. ER contact sites define the position and timing of endosome fission. *Cell* **159**, 1027–1041 (2014).

9. 9.

Friedman, J. R. et al. ER tubules mark sites of mitochondrial division. *Science* **334**, 358–362 (2011).

10. 10.

Raiborg, C. et al. Repeated ER-endosome contacts promote endosome translocation and neurite outgrowth. *Nature* **520**, 234–238 (2015).

11. 11.

Shibata, Y. et al. Mechanisms determining the morphology of the peripheral ER. *Cell* **143**, 774–788 (2010).

12. 12.

Shen, B. et al. Calumenin-1 interacts with Climp63 to cooperatively determine the luminal width and distribution of endoplasmic reticulum sheets. *iScience* **22**, 70–80 (2019).

13. 13.

Vedrenne, C., Klopfenstein, D. R. & Hauri, H.-P. Phosphorylation controls CLIMP-63-mediated anchoring of the endoplasmic reticulum to microtubules. *Mol. Biol. Cell* **16**, 1928–1937 (2005).

14. 14.

Ogawa-Goto, K. et al. p180 is involved in the interaction between the endoplasmic reticulum and microtubules through a novel microtubule-binding and bundling domain. *Mol. Biol. Cell* **18**, 3741–3751 (2007).

15. 15.

Wong, Y. L. et al. Reversible centriole depletion with an inhibitor of Polo-like kinase 4. *Science* **348**, 1155–1160 (2015).

16. 16.

Wu, J. et al. Molecular pathway of microtubule organization at the Golgi apparatus. *Dev. Cell* **39**, 44–60 (2016).

17. 17.

Gavilan, M. P. et al. The dual role of the centrosome in organizing the microtubule network in interphase. *EMBO Rep.* **19**, e45942 (2018).

18. 18.

Mahalingan, K. K. et al. Structural basis for polyglutamate chain initiation and elongation by TTLL family enzymes. *Nat. Struct. Mol. Biol.* **27**, 802–813 (2020).

19. 19.

Garnham, C. P. et al. Multivalent microtubule recognition by tubulin tyrosine ligase-like family glutamylases. *Cell* **161**, 1112–1123 (2015).

20. 20.

Garnham, C. P., Yu, I., Li, Y. & Roll-Mecak, A. Crystal structure of tubulin tyrosine ligase-like 3 reveals essential architectural elements unique to tubulin monoglycylases. *Proc. Natl Acad. Sci. USA* **114**, 6545–6550 (2017).

21. 21.

Rogowski, K. et al. A family of protein-deglutamylating enzymes associated with neurodegeneration. *Cell* **143**, 564–578 (2010).

22. 22.

Valm, A. M. et al. Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature* **546**, 162–167 (2017).

23. 23.

Korolchuk, V. I. et al. Lysosomal positioning coordinates cellular nutrient responses. *Nat. Cell Biol.* **13**, 453–460 (2011).

24. 24.

Jia, R. & Bonifacino, J. S. Lysosome positioning influences mTORC2 and AKT signaling. *Mol. Cell* **75**, 26–38 (2019).

25. 25.

Voeltz, G. K., Prinz, W. A., Shibata, Y., Rist, J. M. & Rapoport, T. A. A class of membrane proteins shaping the tubular endoplasmic reticulum. *Cell* **124**, 573–586 (2006).

26. 26.

Hu, J. et al. A class of dynamin-like GTPases involved in the generation of the tubular ER network. *Cell* **138**, 549–561 (2009).

27. 27.

Friedman, J. R., Webster, B. M., Mastronarde, D. N., Verhey, K. J. & Voeltz, G. K. ER sliding dynamics and ER-mitochondrial contacts occur on acetylated microtubules. *J. Cell Biol.* **190**, 363–375 (2010).

28. 28.

Park, S. H., Zhu, P.-P., Parker, R. L. & Blackstone, C. Hereditary spastic paraplegia proteins REEP1, spastin, and atlastin-1 coordinate microtubule interactions with the tubular ER network. *J. Clin. Invest.* **120**, 1097–1110 (2010).

29. 29.

Valenstein, M. L. & Roll-Mecak, A. Graded control of microtubule severing by tubulin glutamylation. *Cell* **164**, 911–921 (2016).

30. 30.

Zhu, Y. et al. Sec61 β facilitates the maintenance of endoplasmic reticulum homeostasis by associating microtubules. *Protein Cell* **9**, 616–628 (2018).

31. 31.

Farias, G. G. et al. Feedback-driven mechanisms between microtubules and the endoplasmic reticulum instruct neuronal polarity. *Neuron* **102**, 184–201 (2019).

32. 32.

Blackstone, C. Converging cellular themes for the hereditary spastic paraplegias. *Curr. Opin. Neurobiol.* **51**, 139–146 (2018).

33. 33.

Shashi, V. et al. Loss of tubulin deglutamylase CCP1 causes infantile-onset neurodegeneration. *EMBO J.* **37**, e100540 (2018).

34. 34.

Spandl, J., White, D. J., Peychl, J. & Thiele, C. Live cell multicolor imaging of lipid droplets with a new dye, LD540. *Traffic* **10**, 1579–1584 (2009).

35. 35.

Slaymaker, I. M. et al. Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2016).

36. 36.

Vemu, A., Garnham, C. P., Lee, D. Y. & Roll-Mecak, A. Generation of differentially modified microtubules using in vitro enzymatic approaches. *Methods Enzymol.* **540**, 149–166 (2014).

37. 37.

Ziolkowska, N. E. & Roll-Mecak, A. In vitro microtubule severing assays. *Methods Mol. Biol.* **1046**, 323–334 (2013).

38. 38.

Mahamdeh, M., Simmert, S., Luchniak, A., Schaffer, E. & Howard, J. Label-free high-speed wide-field imaging of single microtubules using interference reflection microscopy. *J. Microsc.* **272**, 60–66 (2018).

Acknowledgements

We thank J. S. Bonifacino at NICHD for providing the GFP-mCherry-LC3 vector; D. Maric (NINDS Flow Cytometry Facility) for cell sorting; C. Smith and V. Schram at

NICHD Microscopy and Imaging Core for imaging assistance; V. Bolduc at NINDS for assistance with real-time PCR; E. He and A. Hoofring at NIH Medical Arts for drawing the model figures; and T. Fadero (UNC-Chapel Hill) for discussions and suggestions regarding radial density calculations. All codes are written in Matlab or ImageJ Macro language. This work was supported by the Intramural Research Programs of the National Institute of Neurological Disorders and Stroke and National Heart, Lung and Blood Institute, National Institutes of Health, as well as the Howard Hughes Medical Institute Janelia Research Campus.

Author information

Author notes

1. Jonathon Nixon-Abell

Present address: Cambridge Institute for Medical Research, Cambridge, UK

Affiliations

1. Cell Biology Section, Neurogenetics Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA

Pengli Zheng, Jonathon Nixon-Abell & Craig Blackstone

2. Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA, USA

Christopher J. Obara, Jonathon Nixon-Abell & Jennifer Lippincott-Schwartz

3. Cell Biology and Biophysics Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA

Ewa Szczesna, Kishore K. Mahalingan & Antonina Roll-Mecak

4. Biochemistry and Biophysics Center, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD, USA

Antonina Roll-Mecak

5. MassGeneral Institute for Neurodegenerative Disease, Massachusetts General Hospital, Charlestown, MA, USA

Craig Blackstone

6. Department of Neurology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

Craig Blackstone

Contributions

P.Z. and C.B. conceived and designed the experiments. P.Z. performed most of the experiments. C.J.O. developed and performed the ER morphology analysis. E.S. designed, performed, and interpreted in vitro microtubule binding assays together with P.Z. J.N.-A. performed multispectral imaging of organelles. K.K.M. provided purified TTLL proteins. C.B., A.R.-M. and J.L.-S. supervised the study. P.Z. and C.B. wrote the manuscript with contributions from all authors.

Corresponding authors

Correspondence to [Pengli Zheng](#) or [Craig Blackstone](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

[Extended Data Fig. 1 Knockout of CLIMP63, p180 and KTN1, and resulting ER phenotypes.](#)

a, Schematic illustration of CLIMP63, p180 and KTN1 protein domains. Purple numbers indicate key amino acids. Shorter isoform of p180 (p180s, Uniprot Q9P2E9.5) is also shown. **b**, Western blotting (WB) of the indicated wild-type (WT)

or knockout (KO) cells. The lower band in the KTN1 blots (indicated with an asterisk) corresponds to the shorter cytosolic isoform of KTN1. See [Supplementary Information](#) for uncropped western blots. **c**, Representative images of three patterns of ER distribution in U2OS cells. “Dispersed” (left) is characterized by dominant sheets or matrices at the cell periphery; “Clustered” (right) is characterized by asymmetric dense accumulation of perinuclear ER at one side of the nucleus; all other ER types are considered “Perinuclear”. **d**, Proportion of wild-type or indicated KO cells with different patterns of ER distribution. $n = 3$ experiments with at least 200 cells counted in each experiment. **e**, ER distribution of wild-type or CLIMP63, p180 or KTN1 KO cells treated with 5 μM etoposide or 100 nM camptothecin for 24 h to synchronize cells in S/G2 phase. $n = 3$ experiments with at least 200 cells counted in each experiment. **f**, ER distributions in wild-type or CLIMP63, p180 or KTN1 KO cells treated with 10 μM nocodazole for 24 h and released for 6 h to synchronize cells in G1 phase. $n = 3$ experiments with at least 200 cells counted in each experiment. **g**, Representative images of perinuclear ER in wild-type or CLIMP63, p180 and KTN1 triple-KO cells, showing LUT color grading according to intensity of the ER marker mEmerald-Sec61 β . Scale bars, 10 μm . All bars represent mean \pm s.d.

[Source data](#)

[Extended Data Fig. 2 Methods for quantifying organelle distribution, and quantifications of TRAP \$\alpha\$ and microtubule distribution in knockout cells.](#)

a, Summed projections generated from three-dimensional Airyscan images (left). Fluorescence from neighboring cells is removed and the center of the nucleus is manually selected to function as the origin (yellow dot in right image). Fluorescence intensities are converted to probabilities (right image, see [Supplementary Text](#)). Scale bar, 10 μm . **b**, A radius is drawn out from the center of the nucleus past the farthest point on the cell and swept through 360° in 0.1° intervals, taking a line profile each time. The nuclear envelope and edge of the cell are identified at each radius. **c**, Resulting probabilities of each channel in r- and θ -space, represented as fluorescence intensities. Red and yellow dashed lines indicate the approximate location of the nuclear envelope and the cell edge, respectively (left panel), as shown in (b). The probability distributions of nuclear and ER signals are normalized to correct for cell shape (right panel). Dashed lines indicate the location of the nuclear envelope and cell edge after normalization. **d**, Associated radial distributions of probabilities as measured in terms of distance across the cytoplasm as in (c). Relative probability (y-axis) indicates any single molecule of DAPI (nucleus) or mEmerald-Sec61 β (ER) falling at a specific proportion of the distance between the nuclear envelope and the edge of the cell (x-axis). ER MDR represents the average distance of the ER on this scale and can be used to quantify the propensity of the ER to penetrate the cellular

periphery; higher MDRs indicate a larger proportion of the ER in the periphery. **e–g**, A radius is drawn out past the farthest point on both sides of the cell and swept through 180° in 0.1° intervals, taking a line profile each time (e). The edge of the cell is identified at each radius. Resulting intensity distribution across all radii, with the red line indicating the center of the nucleus (f). For each radius, the difference between two sides of the center (ΔF) is calculated and plotted as a function of θ . The asymmetry value is then calculated as a sum of the exact values of ΔF (g). **h, i**, Quantifications of TRAP α (rough ER) distributions in WT or the indicated KO cells. $n = 31$ cells. **j, k**, Quantifications of microtubule (labeled with anti- α -tubulin) distribution for wild-type or the indicated KO cells. $n = 41, 19, 25, 21, 26, 24, 46, 25, 27, 24, 23, 30, 25, 30, 31$ cells (left to right) for j; $n = 22, 22, 30, 30, 30, 27, 31, 30, 30, 23, 26, 29, 24, 30, 31$ cells for k. **l, m**, Quantifications of ER and microtubule MDR for more cells to show differences in microtubule MDR. $n = 210$ cells for l, $n = 197, 161, 162$ cells for m. All bars represent mean \pm s.d. P values are shown on top; differences without labeling are not significant, comparisons are with the wild-type group using two-tailed t -tests.

[Source data](#)

[**Extended Data Fig. 3 Microtubule sedimentation assays of ER proteins.**](#)

a, Microtubule co-sedimentation assays of U2OS cells. The pellet (P) of 37 °C incubation indicates the microtubule-bound fraction, and supernatant (S) indicates the unbound fraction. 4 °C incubation acts as a microtubule-free control. **b**, Microtubule co-sedimentation assays of U2OS cells with exogenous CLIMP63-HA, p180s-myc or KTN1-myc expression. Proteins were expressed in corresponding knockout cells. Note that only one representative α -tubulin blot (from the CLIMP63 assay) is shown. **c**, p180 is very unstable after cell lysis. The input sample was collected by directly adding sample buffer (50 mM Tris, pH 6.8, 1 mM DTT, 10% glycerol, 2% SDS, 0.1% Bromophenol Blue) onto the plate followed by immediate boiling. Other samples were incubated in lysis buffer (50 mM Tris, pH7.4, 150 mM NaCl, 1% Triton X-100, 1 mM DTT, and protease inhibitor cocktail) at room temperature or on ice for the indicated times before adding sample buffer and boiling. **d**, Western blotting of WT or p180 knockout (KO) U2OS or COS7 cells, showing that only the long isoform is detectable in these cell lines. **e**, Detailed mapping of microtubule-binding domains of CLIMP63. **f**, Mapping of microtubule-binding domains of p180. Amino acid sequences around key microtubule-binding sites are shown at the bottom. Note that this part of the sequence is present in both long and short isoforms of p180. Positively charged amino acids are shown in red. Segments (amino acids 51-80) necessary for microtubule binding are underlined. **g**, Mapping of microtubule-binding domains of KTN1. Amino acid sequences around key microtubule-binding sites are shown at the bottom. Positively charged amino acids are in red. Segments (amino acids 112-120) necessary

for microtubule binding are underlined. See [Supplementary Information](#) for uncropped western blots.

[Extended Data Fig. 4 Rescue assays.](#)

a, Western blots of wild-type or CLIMP63 knockout cells transfected with empty vector or the indicated CLIMP63 constructs. **b**, Representative fluorescence images and quantifications of CLIMP63 knockout cells transfected with empty vector or the indicated CLIMP63 constructs. **c**, Western blots of wild-type or CLIMP63 knockout cells transfected with empty vector or the indicated CLIMP63 constructs. CLIMP63-3SE indicates S3E, S17E and S19E triple mutation of CLIMP63. **d**, Quantifications of ER morphology in wild-type or CLIMP63 KO cells expressing the indicated CLIMP63 constructs. $n = 43, 31, 46$ cells (left to right). **e, f**, Western blots and representative images of wild-type or p180 KO cells expressing the indicated p180 constructs. **g–i**, Western blots and representative images of wild-type or KTN1 KO cells expressing the indicated KTN1 constructs. **j–l**, p180 knockout U2OS cells were transfected with mApple, p180s-mApple, or p180s mutants lacking kinesin-binding domain (KBD), then imaged and quantified for ER morphology. Scale bars, 10 μm . $n = 82, 53, 53$ cells for both (k) and (l) (left to right). Data are mean \pm s.d. P values shown on top, two-tailed t -tests. See [Supplementary Information](#) for uncropped western blots.

[Source data](#)

[Extended Data Fig. 5 PLA of CLIMP63, p180 and KTN1 with \$\alpha\$ -tubulin, and depleting centrosome or Golgi-derived microtubules affects ER distribution.](#)

a, N-terminal mEmerald tag was knocked into endogenous CLIMP63 using CRISPR/Cas9 to facilitate PLA, as no anti-CLIMP63 antibody targeting its cytosolic region was available. Western blots of wild-type or knock-in cells are shown. **b**, Representative PLA images of p180 or KTN1 with α -tubulin. Note that all PLA dots (a dot indicates one binding event) are localized adjacent to both ER and microtubules. **c**, Representative PLA images of endogenous mEmerald-CLIMP63 with α -tubulin. **d**, Western blot analysis of cells transfected with control or AKAP450 siRNAs. **e–h**, Representative data and quantifications for PLA of CLIMP63, p180, or KTN1 with α -tubulin, indicative of microtubule binding. Cells are either untreated, treated with centrinone B (CNB) to deplete the centrosome, or transfected with control siRNA or siAKAP450 (to deplete Golgi-derived microtubules). $n = 130, 134, 142, 143$ cells (left to right) for (f), $n = 116, 139, 140, 150$ cells for (g), $n = 119, 117, 139, 166$ cells for (h). **i, j**, Asymmetry and MDR quantifications for PLA signals between CLIMP63, KTN1, or p180 with α -tubulin. $n = 83, 74, 72, 40, 38$ cells (left to right) for (i), $n = 81$,

79, 82, 43, 44 cells for (j). **k**, Representative images of U2OS cells with or without CNB treatment. Cells stably expressing mEmerald-Sec61 β (green, ER) are stained with DAPI (blue, DNA) and immunolabeled with anti-pericentrin (red) and anti- α -tubulin (magenta) antibodies. Perinuclear and peripheral regions in the cells (boxed) are enlarged at the right. **l**, Quantifications of ER MDR in wild-type or p180 knockout (KO1 and KO2) cells, with or without CNB treatment. $n = 31, 30, 31, 43, 31, 39$ cells (left to right). **m, n**, Representative images and quantifications of ER asymmetry for wild-type or CLIMP63/KTN1 double-knockout cells. Cells were transfected with control siRNA or AKAP450 siRNA to deplete Golgi-derived microtubules. $n = 31, 31, 31, 32, 31, 32$ cells (left to right). Scale bars, 10 μm . Data are mean \pm s.d. P values shown on top, two-tailed t -tests. See [Supplementary Information](#) for uncropped western blots.

[Source data](#)

[**Extended Data Fig. 6 Microtubule binding of CLIMP63, p180 and KTN1 are differently affected by microtubule glutamylation.**](#)

a, Western blots of wild-type or the indicated knockout U2OS cells. **b, c**, Western blots and quantifications of U2OS cells either treated with CNB to deplete the centrosome or else transfected with control siRNA or siAKAP450 (to deplete Golgi-derived microtubules). $n = 6$ experiments. **d**, Representative images of U2OS cells transfected with CLIMP63-mEmerald, p180-mEmerald, or KTN1-mEmerald, with or without co-expression of TTLL7. Note that CLIMP63 overexpression leads to dramatic ER-microtubule alignment (~80% of cells), while p180 and KTN1 require co-expression of TTLL7 (which polyglutamylates microtubules) for robust ER-microtubule alignment. Scale bar, 10 μm . **e**, Quantifications of microtubule-ER alignments in cells transfected with the indicated expression plasmids; $n = 4$ experiments, with at least 100 cells counted in each experiment. **f**, Quantifications of microtubule-ER alignments in cells treated with or without CNB, and overexpressing CLIMP63-mEmerald. $n = 4$ experiments, with 100 cells counted per experiment. **g**, Schematic diagram depicting how microtubule glutamylation levels are modulated by actions of tubulin glutamylases (TTLLs) and deglutamylases (CCPs). **h**, Western blots of U2OS cells overexpressing TTLL4 or TTLL7. polyE detects microtubule polyglutamylation (at least 2 glutamates in the side chain); GT335 reacts with branch points of microtubule glutamylation and thus detects both mono- and polyglutamylation. **i**, Western blots of U2OS cells overexpressing CCP1 or CCP5. Cells were subjected to microtubule sedimentation, and pellets (P, microtubule fraction) and supernatants (S, soluble fraction) at 37 °C were analyzed. **j–l**, Relative PLA intensities for CLIMP63, p180, or KTN1 with α -tubulin (indicative of microtubule binding) in U2OS cells overexpressing the indicated TTLLs or CCPs. $n = 138, 103, 112, 99, 102$ cells (left to right) for (j); $n = 112, 113, 107, 131, 121$ cells for (k); $n = 188, 108, 113, 123, 109$

cells for (l). **m**, Ponceau S staining of p180 and KTN1 proteins purified from HEK293T cells. **n**, Coomassie Brilliant Blue staining of purified CLIMP63 from HEK293T cells. **o, p**, Mass spectra of microtubules glutamylated by TTLL6 for different times, or by else TTLL4 or TTLL7, and used for in vitro microtubule-binding assays. Spectra display characteristic distributions of masses with peaks separated by 129 Da (corresponding to one glutamate). Peak labels show number of glutamates on α - and β -tubulin. Glutamate numbers are indicated in green, dark green, orange, and red for α 1B, α 1C, β I and β IVb isoforms, respectively. **q**, Representative micrographs of CLIMP63-mNeonGreen-2 \times Strep (CLIMP63, cyan) showing binding to unmodified microtubules or else microtubules (magenta) glutamylated in vitro by TTLL7. Interference reflection microscopy images for the microtubule channel were background subtracted and inverted. Average numbers of glutamates added to microtubules as quantified from mass spectrometry data for each group are shown on top. Scale bar, 5 μ m. **r**, Affinities of CLIMP63 for unmodified microtubules and microtubules glutamylated by TTLL7. X-axis indicates the weighted average of glutamate residues attached to α - and β -tubulin. $n = 107, 61, 127$, and 136 microtubules with unmodified, $\alpha+0E/\beta+1.4E$, $\alpha+0E/\beta+2.7E$, and $\alpha+0.1E/\beta+3.8E$ microtubules, respectively. Data are mean \pm s.d. P values shown on top, Mann-Whitney test for panel r, two-tailed t -tests for others. See [Supplementary Information](#) for uncropped western blots.

[Source data](#)

[Extended Data Fig. 7 Modulating tubulin glutamylation affects ER distribution.](#)

a, Representative images of U2OS and COS7 cells sequentially immunolabeled for polyE (polyglutamylation), polyE plus GT335 (glutamylation) and Alexa 647 conjugated α -tubulin. GT335 signal indicates monoglutamylation in this case because polyglutamylation is pre-saturated with polyE antibody. **b**, Quantifications of signal distributions for cells as in (a). Data points from the same cell are linked by solid lines. $n = 20$ cells for U2OS, 35 cells for COS7. **c**, Western blot analysis of U2OS, COS7, HeLa and RPE1 cells. **d**, MDR ratios of monoE (indicated by GT335 labeling) to polyE, or microtubules (indicated by α -tubulin labeling) to polyE. $n = 20$ cells for U2OS, 35 cells for COS7. **e**, Representative images of U2OS cells overexpressing the indicated TTLLs or CCPs. **f, g**, Quantifications of ER and microtubule MDRs of cells as in e; $n = 37, 40, 34, 40, 40, 41, 39$ cells (left to right) for (f); $n = 29, 38, 37, 51, 38, 46, 39$ cells for (g). **h**, Quantifications of ER MDR of cells in p180/KTN1 double knockout cells overexpressing the indicated TTLLs or CCPs; $n = 63, 61, 64, 63, 61$ cells (left to right). **i**, Quantifications of ER MDR of wild-type or KTN1 knockout cells overexpressing TTLL7; $n = 40, 64$ cells (left to right). **j**, U2OS cells transfected with the indicated siRNAs for 72 h were analyzed by western blotting. **k**, U2OS cells

transfected with the indicated siRNAs for 48 h were analyzed by real time PCR. $n = 4$ repeats. **l, m**, Representative image and quantification of ER in control or CCP5 siRNA transfected cells. Scale bars, 10 μm . $n = 64, 70, 69, 39, 39, 39$ cells (left to right). Data are mean \pm s.d. P values shown on top, two-tailed t -test.

See [Supplementary Information](#) for uncropped western blots.

[Source data](#)

[**Extended Data Fig. 8 Knockout of CLIMP63, p180 and KTN1, or overexpression of CCPs in COS7 cells changes ER distribution.**](#)

a, Western blots of wild-type or else CLIMP63, p180, or KTN1 knockout cells. **b**, Representative images of wild-type or CLIMP63, p180, KTN1 knockout cells. ER is labeled by endogenous mEmerald-calreticulin. Scale bar, 10 μm . **c, d**, Quantifications of ER distribution. $n = 60, 60, 58, 96, 88, 96, 51$ cells (left to right) for both (c) and (d). **e–h**, Wild-type or p180 knockout COS7 cells with endogenous mEmerald-calreticulin (ER, green) were transfected with CCP5-mApple or CCP6-mApple for 24 h and analyzed by western blotting (e) and confocal imaging (f). ER distributions are quantified (g, h). Scale bar, 10 μm . $n = 39, 49, 45, 41, 40$ cells (left to right) for both (g) and (h). Data are mean \pm s.d. P values shown on top, two-tailed t -tests.

See [Supplementary Information](#) for uncropped western blots.

[Source data](#)

[**Extended Data Fig. 9 Organelle distribution in knockout cells.**](#)

a, Simultaneous live imaging of six organelles with spectral unmixing. Two representative cells are shown. Note that where there is more ER, there tends to be more of the other organelles. **b**, Representative images of organelle distributions in CLIMP63 or p180 knockout cells. Markers used: mEmerald-Sec61 β for ER; anti-TOM20 for mitochondria; anti-GM130 for Golgi apparatus; anti-EEA1 for endosomes; anti-LC3 for autophagosomes; anti-catalase for peroxisomes, anti-Lamp1 for lysosomes; and LD540 for lipid droplets. Cells labeled for autophagosome distribution were starved in EBSS for 2 h. ER is shown in green, while other organelles are in red. **c, d**, Quantifications of the distributions of different organelles labeled with specific markers in wild-type or else CLIMP63, KTN1, or p180 knockout cells. $n = 30, 31, 30, 46$ cells (left to right) for mitochondria MDR, $n = 83, 86, 87, 46$ cells for lipid droplet MDR, $n = 30, 31, 31, 46$ cells for peroxisome MDR, $n = 29$ cells for Golgi MDR, $n = 30$ cells for Autophagosome and lysosome MDR; $n = 51, 46$ cells for mitochondria asymmetry, $n = 29, 31$ cells for Golgi asymmetry, $n = 49, 50$ cells for lipid droplet asymmetry, $n = 31$ cells for autophagosome, lysosome and peroxisome asymmetry. **e–g**, Quantifications of lysosome, mitochondria, and peroxisome

distributions in wild-type or p180/KTN1 double knockout cells transfected with control vector or CCP1. $n = 32, 36, 36, 37$ cells (left to right) for (e) and (f), $n = 41, 48, 44, 40$ cells for (g). Scale bars, 10 μm . All bars represent mean \pm s.d. with individual data points shown. P values are shown along the top, two-tailed t -tests.

[Source data](#)

Extended Data Fig. 10 CLIMP63 and p180 regulate autophagic flux.

a, U2OS cells were pre-treated with the indicated compounds for 2 h (except etoposide, which was added for 16 h), then with EBSS plus the same compound for 30 min, and then subjected to western blotting. Compounds used: 1 μM lysosome degradation inhibitor DC661, 1 μM proteasome degradation inhibitor MG132, 6 $\mu\text{g/mL}$ ER translation inhibitor puromycin, 100 $\mu\text{g/mL}$ protein translation inhibitor cycloheximide (CHX), 50 μM etoposide that causes DNA damage and also blocks protein expression, and 1 μM N-linked glycosylation inhibitor tunicamycin. **b**, Western blot analysis of wild-type, CLIMP63 or p180 knockout U2OS cells (Parental, no stable mEmerald-Sec61 β expression). **c**, Representative images and quantifications for wild-type, CLIMP63 or p180 knockout U2OS cells expressing GFP-mCherry-LC3. The GFP signal is quenched by the acidic environment of lysosomes, so only autophagosomes that have not yet fused with lysosomes have green GFP signal, while autolysosomes only exhibit mCherry signal. $n = 40$ cells. Scale bar, 10 μm . **d**, Wild-type or CLIMP63 knockout cells were starved in EBSS for 0 or 8 h, with or without brefeldin A treatment to block lysosomal degradation, and then immunoblotted. Relative amounts of autophagic substrate p62 are shown in the lower panel. $n = 4$ experiments. **e**, Cathepsin L activity as determined by substrate reaction. $n = 4$ experiments. **f**, Lysosome acidification analysis using Lysosensor Green. $n = 200$ cells. **g**, Wild-type or p180 knockout cells were starved in EBSS for 0 or 8 h, with or without brefeldin A treatment to block lysosomal degradation, and western blotted. **h**, Wild-type or p180 knockout cells were starved in EBSS for 2 h, then re-supplemented with regular medium for 0–30 min. Cells were immunoblotted for phosphorylated S6K (p-S6K) and total S6K to indicate activity of mTOR signaling. $n = 4$ experiments. **i**, Western blots reveal the indicated microtubule modifications with or without EBSS starvation for 0.5 and 2 h. Relative intensities of GT335 and polyE immunoreactive signals are quantified at the right. $n = 5$ experiments. **j**, Representative images for PLA of p180 and the ribosomal marker RPL3 at 0 or 2 h of EBSS starvation, or else for cells treated with 2 $\mu\text{g/mL}$ puromycin for 2 h. Scale bar, 50 μm . **k**, Representative images of U2OS cells transfected with mEmerald-Sec61 β , p180s-mEmerald, or p180L-mEmerald with or without co-transfection of TTLL7-3 \times flag. Scale bar, 10 μm . **l**, Quantifications of microtubule alignments in cells transfected with the indicated constructs; $n = 4$ experiments, with at least 100 cells counted per experiment. **m**, Microtubule alignments of p180L-mEmerald in cells treated with EBSS or 2 $\mu\text{g/mL}$

puromycin for 2 h. $n = 6$ experiments, with at least 100 cells counted per experiment. Data are mean \pm s.d. P values are shown along the top, two-tailed t -tests. See [Supplementary Information](#) for uncropped western blots.

[Source data](#)

Supplementary information

[Supplementary Information](#)

This file contains Supplementary Text and Supplementary Figs 1–7.

[Reporting Summary](#)

[Peer Review File](#)

[Supplementary Video 1](#)

Wild-type or CLIMP63, p180 KO U2OS cells stably expressing mEmerald–Sec61 β were labelled with Lysotracker Red, starved with EBSS, and recorded for 2.5 h with imaging at 1 min intervals. Scale bar, 10 μm .

Source data

[Source Data Fig. 1](#)

[Source Data Fig. 2](#)

[Source Data Fig. 3](#)

[Source Data Extended Data Fig. 1](#)

[Source Data Extended Data Fig. 2](#)

[Source Data Extended Data Fig. 4](#)

[Source Data Extended Data Fig. 5](#)

[Source Data Extended Data Fig. 6](#)

Source Data Extended Data Fig. 7

Source Data Extended Data Fig. 8

Source Data Extended Data Fig. 9

Source Data Extended Data Fig. 10

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Reprints and Permissions

About this article

Cite this article

Zheng, P., Obara, C.J., Szczesna, E. *et al.* ER proteins decipher the tubulin code to regulate organelle distribution. *Nature* **601**, 132–138 (2022).
<https://doi.org/10.1038/s41586-021-04204-9>

- Received: 09 March 2021
- Accepted: 03 November 2021
- Published: 15 December 2021
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04204-9>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04204-9>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 08 December 2021](#)

Structure of pathological TDP-43 filaments from ALS with FTLD

- [Diana Arseni](#) ORCID: [orcid.org/0000-0001-7585-288X¹](https://orcid.org/0000-0001-7585-288X),
- [Masato Hasegawa](#) ORCID: [orcid.org/0000-0001-7415-8159²](https://orcid.org/0000-0001-7415-8159),
- [Alexey G. Murzin¹](#),
- [Fuyuki Kametani](#) ORCID: [orcid.org/0000-0001-9125-7001²](https://orcid.org/0000-0001-9125-7001),
- [Makoto Arai](#) ORCID: [orcid.org/0000-0003-3400-9815³](https://orcid.org/0000-0003-3400-9815),
- [Mari Yoshida⁴](#) &
- [Benjamin Ryskeldi-Falcon](#) ORCID: [orcid.org/0000-0002-8176-2618¹](https://orcid.org/0000-0002-8176-2618)

[Nature](#) volume **601**, pages 139–143 (2022)

- 7868 Accesses
- 1 Citations
- 192 Altmetric
- [Metrics details](#)

Subjects

- [Amyotrophic lateral sclerosis](#)
- [Cryoelectron microscopy](#)
- [Dementia](#)
- [Neurodegeneration](#)

Abstract

The abnormal aggregation of TAR DNA-binding protein 43 kDa (TDP-43) in neurons and glia is the defining pathological hallmark of the neurodegenerative disease amyotrophic lateral sclerosis (ALS) and multiple forms of frontotemporal lobar degeneration (FTLD)^{1,2}. It is also common in other diseases, including Alzheimer's and Parkinson's. No disease-modifying therapies exist for these conditions and early diagnosis is not possible. The structures of pathological TDP-43 aggregates are unknown. Here we used cryo-electron microscopy to determine the structures of aggregated TDP-43 in the frontal and motor cortices of an individual who had ALS with FTLD and from the frontal cortex of a second individual with the same diagnosis. An identical amyloid-like filament structure comprising a single protofilament was found in both brain regions and individuals. The ordered filament core spans residues 282–360 in the TDP-43 low-complexity domain and adopts a previously undescribed double-spiral-shaped fold, which shows no similarity to those of TDP-43 filaments formed in vitro^{3,4}. An abundance of glycine and neutral polar residues facilitates numerous turns and restricts β-strand length, which results in an absence of β-sheet stacking that is associated with cross-β amyloid structure. An uneven distribution of residues gives rise to structurally and chemically distinct surfaces that face external densities and suggest possible ligand-binding sites. This work enhances our understanding of the molecular pathogenesis of ALS and FTLD and informs the development of diagnostic and therapeutic agents that target aggregated TDP-43.

This is a preview of subscription content

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Fig. 1: TDP-43 pathology of ALS with FTLD.

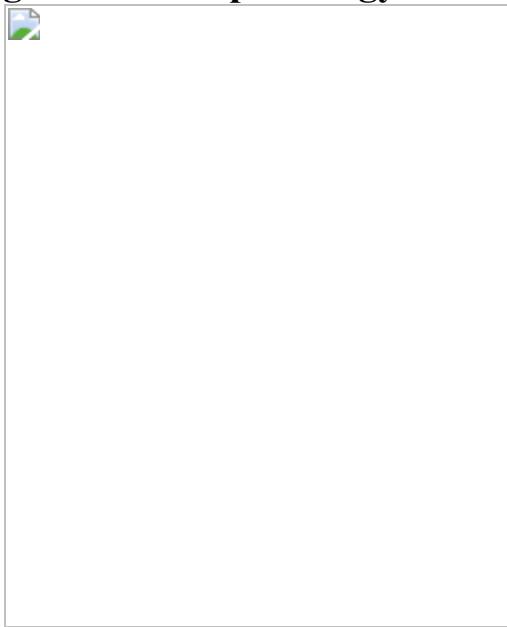


Fig. 2: Cryo-EM maps and atomic model of TDP-43 filaments from ALS with FTLD.

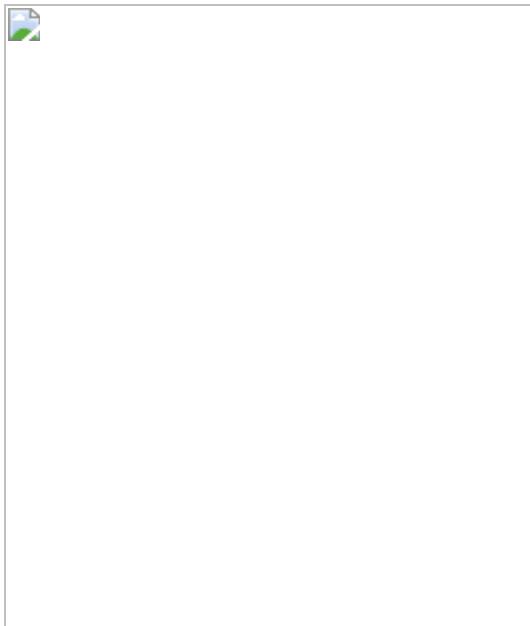


Fig. 3: Surfaces of TDP-43 filaments from ALS with FTLD.



Data availability

Cryo-EM datasets have been deposited to the Electron Microscopy Public Image Archive (EMPIAR) under the following accession numbers: [10830](#) for individual one, frontal cortex; [10831](#) for individual one, motor cortex; and [10832](#) for individual two, frontal cortex. Cryo-EM maps have been deposited to the Electron Microscopy Data Bank (EMDB) under the following accession numbers: [13708](#) for individual one, frontal cortex; [13710](#) for individual one, motor cortex; and [13712](#) for individual two, frontal cortex. The atomic model of the double-spiral fold has been

deposited to the Protein Data Bank (PDB) under accession number [7PY2](#). The atomic model of the α -subunit of glutamate synthase was obtained from the PDB (accession number 1EA0). The atomic model of the α -helices formed in solution by the hydrophobic region of TDP-43 was obtained from the PDB (accession number 2N3X). The atomic models of TDP-43 filaments formed in vitro were obtained from the PDB (accession numbers 7KWZ, 6N3C, 6N3A, 6N37 and 6N3B). Mass spectrometry data have been deposited to the Japan Proteome Standard Repository (jPOSTrepo) under accession number PXD029001. Any other data are available from the corresponding author upon request.

References

1. 1.

Neumann, M. et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130–133 (2006).

2. 2.

Arai, T. et al. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem. Biophys. Res. Commun.* **351**, 602–611 (2006).

3. 3.

Cao, Q., Boyer, D. R., Sawaya, M. R., Ge, P. & Eisenberg, D. S. Cryo-EM structures of four polymorphic TDP-43 amyloid cores. *Nat. Struct. Mol. Biol.* **26**, 619–627 (2019).

4. 4.

Li, Q., Babinchak, W. M. & Surewicz, W. K. Cryo-EM structure of amyloid fibrils formed by the entire low complexity domain of TDP-43. *Nat. Commun.* **12**, 1620 (2021).

5. 5.

Brettschneider, J. et al. Stages of pTDP-43 pathology in amyotrophic lateral sclerosis. *Ann. Neurol.* **74**, 20–38 (2013).

6. 6.

Nonaka, T. et al. Prion-like properties of pathological TDP-43 aggregates from diseased brains. *Cell Rep.* **4**, 124–134 (2013).

7. 7.

Phukan, J., Pender, N. P. & Hardiman, O. Cognitive impairment in amyotrophic lateral sclerosis. *Lancet Neurol.* **6**, 994–1003 (2007).

8. 8.

Sreedharan, J. et al. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science* **319**, 1668–1672 (2008).

9. 9.

Kabashi, E. et al. *TARDBP* mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nat. Genet.* **40**, 572–574 (2008).

10. 10.

Van Deerlin, V. M. et al. *TARDBP* mutations in amyotrophic lateral sclerosis with TDP-43 neuropathology: a genetic and histopathological analysis. *Lancet Neurol.* **7**, 409–416 (2008).

11. 11.

Hergesheimer, R. C. et al. The debated toxic role of aggregated TDP-43 in amyotrophic lateral sclerosis: a resolution in sight? *Brain* **142**, 1176–1194 (2019).

12. 12.

Ayala, Y. M. et al. Structural determinants of the cellular localization and shuttling of TDP-43. *J. Cell Sci.* **121**, 3778–3785 (2008).

13. 13.

Afroz, T. et al. Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation. *Nat. Commun.* **8**, 45 (2017).

14. 14.

Colombrita, C. et al. TDP-43 is recruited to stress granules in conditions of oxidative insult. *J. Neurochem.* **111**, 1051–1061 (2009).

15. 15.

Hasegawa, M. et al. Phosphorylated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Ann. Neurol.* **64**, 60–70 (2008).

16. 16.

Lin, W.-L. & Dickson, D. W. Ultrastructural localization of TDP-43 in filamentous neuronal inclusions in various neurodegenerative diseases. *Acta Neuropathol.* **116**, 205–213 (2008).

17. 17.

Mori, F. et al. Maturation process of TDP-43-positive neuronal cytoplasmic inclusions in amyotrophic lateral sclerosis with and without dementia. *Acta Neuropathol.* **116**, 193–203 (2008).

18. 18.

Tsuji, H. et al. Molecular analysis and biochemical classification of TDP-43 proteinopathy. *Brain* **135**, 3380–3391 (2012).

19. 19.

Mackenzie, I. R. A. et al. A harmonized classification system for FTLD-TDP pathology. *Acta Neuropathol.* **122**, 111–113 (2011).

20. 20.

Jiang, L.-L. et al. Structural transformation of the amyloidogenic core region of TDP-43 protein initiates its aggregation and cytoplasmic inclusion. *J. Biol. Chem.* **288**, 19614–19624 (2013).

21. 21.

Kametani, F. et al. Mass spectrometric analysis of accumulated TDP-43 in amyotrophic lateral sclerosis brains. *Sci. Rep.* **6**, 23281 (2016).

22. 22.

Sunde, M. et al. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* **273**, 729–739 (1997).

23. 23.

Binda, C. et al. Cross-talk and ammonia channeling between active centers in the unexpected domain arrangement of glutamate synthase. *Structure* **8**, 1299–1308 (2000).

24. 24.

Parkinson, G. N., Cuenca, F. & Neidle, S. Topology conservation and loop flexibility in quadruplex–drug recognition: crystal structures of inter- and intramolecular telomeric DNA quadruplex–drug complexes. *J. Mol. Biol.* **381**, 1145–1156 (2008).

25. 25.

Shi, Y. et al. Structure-based classification of tauopathies. *Nature* **598**, 359–363 (2021).

26. 26.

Schweighauser, M. et al. Structures of α -synuclein filaments from multiple system atrophy. *Nature* **585**, 464–469 (2020).

27. 27.

Kollmer, M. et al. Cryo-EM structure and polymorphism of A β amyloid fibrils purified from Alzheimer's brain tissue. *Nat. Commun.* **10**, 4760 (2019).

28. 28.

Krebs, M. R. H. et al. The binding of thioflavin-T to amyloid fibrils: localisation and implications. *J. Struct. Biol.* **149**, 30–37 (2005).

29. 29.

Shi, Y. et al. Cryo-EM structures of tau filaments from Alzheimer's disease with PET ligand APN-1607. *Acta Neuropathol.* **141**, 697–708 (2021).

30. 30.

Bigio, E. H. et al. Inclusions in frontotemporal lobar degeneration with TDP-43 proteinopathy (FTLD-TDP) and amyotrophic lateral sclerosis (ALS), but not FTLD with FUS proteinopathy (FTLD-FUS), have properties of amyloid. *Acta Neuropathol.* **125**, 463–465 (2013).

31. 31.

Robinson, J. L. et al. TDP-43 skeins show properties of amyloid in a subset of ALS cases. *Acta Neuropathol.* **125**, 121–131 (2013).

32. 32.

Sun, Y. et al. Physiologically important electrolytes as regulators of TDP-43 aggregation and droplet-phase behavior. *Biochemistry* **58**, 590–607 (2019).

33. 33.

Marquié, M. et al. Validating novel tau positron emission tomography tracer [F-18]-AV-1451 (T807) on postmortem brain tissue. *Ann. Neurol.* **78**, 787–800 (2015).

34. 34.

Lowe, V. J. et al. An autoradiographic evaluation of AV-1451 tau PET in dementia. *Acta Neuropathol. Commun.* **4**, 58 (2016).

35. 35.

Sander, K. et al. Characterization of tau positron emission tomography tracer [^{18}F]AV-1451 binding to postmortem tissue in Alzheimer's disease, primary tauopathies, and other dementias. *Alzheimers Dement.* **12**, 1116–1124 (2016).

36. 36.

Zhang, W. et al. Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *eLife* **8**, e43584 (2019).

37. 37.

Murray, D. T. et al. Structure of FUS protein fibrils and its relevance to self-assembly and phase separation of low-complexity domains. *Cell* **171**, 615-627.e16 (2017).

38. 38.

Lu, J. et al. CryoEM structure of the low-complexity domain of hnRNPA2 and its conversion to pathogenic amyloid. *Nat. Commun.* **11**, 4090 (2020).

39. 39.

Sun, Y. et al. The nuclear localization sequence mediates hnRNPA1 amyloid fibril formation revealed by cryoEM structure. *Nat. Commun.* **11**, 6349 (2020).

40. 40.

Lee, M., Ghosh, U., Thurber, K. R., Kato, M. & Tycko, R. Molecular structure and interactions within amyloid-like fibrils formed by a low-complexity protein sequence from FUS. *Nat. Commun.* **11**, 5735 (2020).

41. 41.

Chen, A. K.-H. et al. Induction of amyloid fibrils by the C-terminal fragments of TDP-43 in amyotrophic lateral sclerosis. *J. Am. Chem. Soc.* **132**, 1186–1187 (2010).

42. 42.

Guo, W. et al. An ALS-associated mutation affecting TDP-43 enhances protein aggregation, fibril formation and neurotoxicity. *Nat. Struct. Mol. Biol.* **18**, 822–830 (2011).

43. 43.

Laferrière, F. et al. TDP-43 extracted from frontotemporal lobar degeneration subject brains displays distinct aggregate assemblies and neurotoxic effects reflecting disease progression rates. *Nat. Neurosci.* **22**, 65–77 (2019).

44. 44.

Porta, S. et al. Patient-derived frontotemporal lobar degeneration brain extracts induce formation and spreading of TDP-43 pathology in vivo. *Nat. Commun.* **9**, 4220 (2018).

45. 45.

Yoshida, M. Amyotrophic lateral sclerosis with dementia: the clinicopathological spectrum. *Neuropathology* **24**, 87–102 (2004).

46. 46.

Taniguchi-Watanabe, S. et al. Biochemical classification of tauopathies by immunoblot, protein sequence and mass spectrometric analyses of Sarkosyl-insoluble and trypsin-resistant tau. *Acta Neuropathol.* **131**, 267–280 (2016).

47. 47.

Tsuji, H. et al. Epitope mapping of antibodies against TDP-43 and detection of protease-resistant fragments of pathological TDP-43 in amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Biochem. Biophys. Res. Commun.* **417**, 116–121 (2012).

48. 48.

Tarutani, A. & Hasegawa, M. in *Experimental Models of Parkinson's Disease* (ed. Imai, Y.) 17–25 (Springer, 2021).

49. 49.

Kametani, F. et al. Comparison of common and disease-specific post-translational modifications of pathological tau associated with a wide range of tauopathies. *Front. Neurosci.* **14**, 581936 (2020).

50. 50.

Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* **6**, 5–17 (2019).

51. 51.

Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

52. 52.

He, S. & Scheres, S. H. W. Helical reconstruction in RELION. *J. Struct. Biol.* **198**, 163–176 (2017).

53. 53.

Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

54. 54.

Scheres, S. H. W. Amyloid structure determination in RELION-3.1. *Acta Crystallogr. D Struct. Biol.* **76**, 94–101 (2020).

55. 55.

Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ* **7**, 253–267 (2020).

56. 56.

Chen, S. et al. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryo-microscopy. *Ultramicroscopy* **135**, 24–35 (2013).

57. 57.

Casañal, A., Lohkamp, B. & Emsley, P. Current developments in Coot for macromolecular model building of electron cryo-microscopy and crystallographic data. *Protein Sci.* **29**, 1055–1064 (2020).

58. 58.

Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.* **74**, 519–530 (2018).

59. 59.

Brown, A. et al. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta*

Crystallogr. D Biol. Crystallogr. **71**, 136–153 (2015).

60. 60.

Yamashita, K., Palmer, C. M., Burnley, T. & Murshudov, G. N. Cryo-EM single -particle structure refinement and map calculation using *Servalcat*. *Acta Crystallogr. D Struct. Biol.* **77**, 1281–1291 (2021).

61. 61.

Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).

62. 62.

Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

Acknowledgements

We thank the individuals and their families for donating the brain tissue; K. Yamada and Y. Itoh for the clinical information; A. Akagi for dissecting the brain tissue; T. Arai and H. Akiyama for helpful comments on the clinical history and neuropathology; R. Otani and M. Takase for technical support with immunohistochemistry; K. Yamashita and G. N. Murshudov for help with Servalcat; staff at the MRC Laboratory of Molecular Biology Electron Microscopy Facility for access to and support with EM; staff at the MRC Laboratory of Molecular Biology Scientific Computing Facility for access to and support with computing; and M. Goedert, S. H. W. Scheres, S.W. Davies, S. Tetter, T. S. Behr and R. R. Chen for discussions. D.A. is a Fellow at Darwin College, University of Cambridge. This work was supported by the Medical Research Council (MRC) grant MC_UP_1201/25 and Alzheimer’s Research UK award ARUK-RS2019-001 to B.R.-F.; the Japan Agency for Medical Research and Development (AMED) grants JP20ek0109392 and JP20ek0109391 to M.Y. and JP20dm0207072 to M.H;

and the Japan Science and Technology Agency (JST) Core Research for Evolutional Science and Technology (CREST) grant JPMJCR18H3 to M.H.

Author information

Affiliations

1. MRC Laboratory of Molecular Biology, Cambridge, UK

Diana Arseni, Alexey G. Murzin & Benjamin Ryskeldi-Falcon

2. Department of Brain and Neurosciences, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan

Masato Hasegawa & Fuyuki Kametani

3. Department of Psychiatry and Behavioural Sciences, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan

Makoto Arai

4. Institute for Medical Science of Aging, Aichi Medical University, Aichi, Japan

Mari Yoshida

Contributions

M.Y. identified the individuals and performed the neuropathological examinations. M.A. performed the genetic analysis. M.H. extracted the TDP-43 filaments and performed immunohistochemistry, immunoblotting and immuno-EM. F.K. performed the mass spectrometry. D.A. performed cryo-EM. D.A. and B.R.-F. analysed the cryo-EM data. D.A., B.R.-F. and A.G.M. built and analysed the atomic model. B.R.-F. supervised the study. B.R.-F., M.Y. and M.H. are the senior authors. All authors contributed to writing the manuscript.

Corresponding author

Correspondence to [Benjamin Ryskeldi-Falcon](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks James Shorter, Henning Stahlberg and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

[Extended Data Fig. 1 TDP-43 pathology of ALS with FTLD continued.](#)

a, Staining of TDP-43 (brown) in the frontal and motor cortices and spinal cords of individuals 1 and 2 with anti-TDP-43 (amino acids 203–209) antibody. Nuclei were counterstained in blue. Scale bars, 50 µm. **b**, Staining of TDP-43 glial cytoplasmic inclusions (brown) in the motor cortices of individuals 1 and 2 with anti-phosphorylated S409 and S410 TDP-43 antibody. Nuclei were counterstained in blue. Scale bars, 50 µm. **c, d**, Immunoblots of the total homogenate (T), Sarkosyl-soluble fraction (S) and Sarkosyl-insoluble fraction (I) from the motor cortices of individuals 1 and 2 with anti-TDP-43 (amino acids 203–209) (**c**) and anti-phosphorylated S409 and S410 TDP-43 (**d**) antibodies. The original, uncropped blots are shown in Supplementary Fig. 1. **e**, Immuno-electron microscopy of the

Sarkosyl-insoluble fraction from the frontal cortex of individual 2 using anti-phosphorylated S409 and S410 TDP-43 antibody. Scale bar, 100 nm. a–e, Similar results were obtained in at least three independent experiments.

Extended Data Fig. 2 Cryo-EM and helical reconstruction.

a, Representative cryo-EM images of TDP-43 filaments from the frontal and motor cortices of individual 1 and from the frontal cortex of individual 2. Scale bars, 100 nm. Similar results were obtained in at least three independent experiments. **b**, Fourier shell correlation (FSC) curves for two independently-refined cryo-EM half-maps. The FSC threshold of 0.143 is shown with red dashed lines. **c**, Local resolution estimates for the Cryo-EM 3D reconstructions. Scale bars, 10 Å. **d**, Cryo-EM density maps viewed along the helical axis. Scale bars, 10 Å.

Extended Data Fig. 3 Cryo-EM density map and atomic model comparisons.

a–c, Views of the cryo-EM density map of TDP-43 filaments from the frontal cortex of individual 1 and the corresponding atomic model showing representative densities for amino acid side chains (**a**), peptide group oxygen atoms (**b**) and ordered solvent molecules (red arrows) (**c**). **d**, Fourier shell correlation (FSC) curves for the refined atomic model against the cryo-EM density map (black); for the atomic model shaken and refined using the first half-map against the first half-map (magenta); and for the same atomic model against the second half-map (green). The FSC threshold of 0.5 is shown with a red dashed line.

Extended Data Fig. 4 Double-spiral fold of TDP-43 filaments from ALS with FTLD.

a, Schematic view of the double-spiral fold. The two hydrophobic clusters in the hydrophobic nucleus, composed of the side chains of M307, M311, F313, M322, M323, A326, A329 and L330 (cluster 1); and A328, W334, M336 and L340 (cluster 2), are numbered. **b**, Hydrophobicity of the double-spiral fold from most hydrophilic (cyan) to most hydrophobic (yellow). **c**,

Secondary structure of the double-spiral fold, depicted as five successive rungs. The glycine-rich (G282-G310, magenta), hydrophobic (M311-S342, white) and Q/N-rich (Q343-Q360, green) regions are highlighted. **d**, Views of the double-spiral fold, depicted as three successive rungs, showing hydrogen bonds (blue dashed lines) between buried polar side chains and main chain peptide groups. **e**, Superposition of residues N319-A326 of the double-spiral fold (cyan), depicted as two successive rungs, with residues G88-M95 and G107-A114 of the β -helix domain of glutamate synthase (magenta, PDB ID 1EA0). **f**, Unmasked cryo-EM 3D reconstructions of TDP-43 filaments from the frontal and motor cortices of individual 1 and from the frontal cortex of individual 2, shown as central slices perpendicular to the helical axis. Additional density within the prominent groove on the filament surface formed by the main chain of G282–Q286 and the side chain of Q286 is indicated with a magenta arrow. Additional densities adjacent to flat strips on the surface of the glycine-rich spiral branch between R293 and A315 are indicated with yellow arrows. Additional less well-resolved density projecting from a polar patch formed by the side chains of N352, Q354 and Q356 on the surface of the Q/N-rich spiral branch is indicated with a cyan arrow. Scale bars, 25 Å.

Extended Data Fig. 5 Comparison of the double-spiral fold with recombinant TDP-43 structures.

a, Amino acid sequence alignment of the secondary structure elements of the double-spiral fold and recombinant TDP-43 structures with the TDP-43 LC domain. PDB IDs are given for the recombinant TDP-43 structures. The glycine-rich (G282-G310, magenta), hydrophobic (M311-S342, white) and Q/N-rich (Q343-Q360, green) regions are highlighted on the double-spiral fold secondary structure elements. **b**, Secondary structure of the double-spiral fold and recombinant TDP-43. PDB IDs are given for the recombinant TDP-43 structures. The glycine-rich (G282-G310, magenta), hydrophobic (M311-S342, white) and Q/N-rich (Q343-Q360, green) regions are highlighted.

Extended Data Fig. 6 Heat stability of TDP-43 filaments from ALS with FTLD.

Representative cryo-EM images of TDP-43 filaments from the frontal cortex of individual 1 with and without heating at 65 °C for 10 min. Scale bars, 100 nm. Similar results were obtained in at least three independent experiments.

Extended Data Table 1 Clinicopathological evaluation and TDP-43 inclusion pathology

Extended Data Table 2 Cryo-EM data collection, refinement and validation statistics

Extended Data Table 3 Mass spectrometry analyses of post-translational modifications

Extended Data Table 4 Compatibility of ALS-associated *TARDBP* mutations with the double-spiral fold

Supplementary information

Supplementary Figure

Uncropped images of immunoblots presented in this study.

Reporting Summary

Peer Review File

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Arseni, D., Hasegawa, M., Murzin, A.G. *et al.* Structure of pathological TDP-43 filaments from ALS with FTLD. *Nature* **601**, 139–143 (2022).
<https://doi.org/10.1038/s41586-021-04199-3>

- Received: 03 August 2021
- Accepted: 02 November 2021
- Published: 08 December 2021
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04199-3>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

Further reading

- [Aggregates of TDP-43 protein spiral into view](#)

- Hana M. Odeh
- James Shorter

Nature (2022)

[Aggregates of TDP-43 protein spiral into view](#)

- Hana M. Odeh
- James Shorter

News & Views 08 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04199-3>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 23 December 2021](#)

Time-resolved structural analysis of an RNA-cleaving DNA catalyst

- [Jan Borggräfe](#)^{1,2},
- [Julian Victor](#)¹,
- [Hannah Rosenbach](#)¹,
- [Aldino Viegas](#) [ORCID: orcid.org/0000-0003-1733-136X](#)¹ nAff10,
- [Christoph G. W. Gertzen](#) [ORCID: orcid.org/0000-0002-9562-7708](#)^{3,4},
- [Christine Wuebben](#) [ORCID: orcid.org/0000-0002-7860-3149](#)⁵,
- [Helena Kovacs](#) [ORCID: orcid.org/0000-0002-5246-9439](#)⁶,
- [Mohanraj Gopalswamy](#)³,
- [Detlev Riesner](#)¹,
- [Gerhard Steger](#)¹,
- [Olav Schiemann](#) [ORCID: orcid.org/0000-0001-6346-9779](#)⁵,
- [Holger Gohlke](#)^{2,3,7},
- [Ingrid Span](#)^{1,8} &
- [Manuel Etzkorn](#) [ORCID: orcid.org/0000-0002-9796-3246](#)^{1,2,9}

[Nature](#) volume **601**, pages 144–149 (2022)

- 6745 Accesses
- 150 Altmetric
- [Metrics details](#)

Subjects

- [Biocatalysis](#)

- [DNA](#)
- [Molecular conformation](#)
- [Solution-state NMR](#)

Abstract

The 10–23 DNAzyme is one of the most prominent catalytically active DNA sequences^{1,2}. Its ability to cleave a wide range of RNA targets with high selectivity entails a substantial therapeutic and biotechnological potential². However, the high expectations have not yet been met, a fact that coincides with the lack of high-resolution and time-resolved information about its mode of action³. Here we provide high-resolution NMR characterization of all apparent states of the prototypic 10–23 DNAzyme and present a comprehensive survey of the kinetics and dynamics of its catalytic function. The determined structure and identified metal-ion-binding sites of the precatalytic DNAzyme–RNA complex reveal that the basis of the DNA-mediated catalysis is an interplay among three factors: an unexpected, yet exciting molecular architecture; distinct conformational plasticity; and dynamic modulation by metal ions. We further identify previously hidden rate-limiting transient intermediate states in the DNA-mediated catalytic process via real-time NMR measurements. Using a rationally selected single-atom replacement, we could considerably enhance the performance of the DNAzyme, demonstrating that the acquired knowledge of the molecular structure, its plasticity and the occurrence of long-lived intermediate states constitutes a valuable starting point for the rational design of next-generation DNAzymes.

This is a preview of subscription content

Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Fig. 1: The precatalytic Dz–RNA complex adopts an unexpected but effective structure that is in dynamic exchange with divalent metal ions.

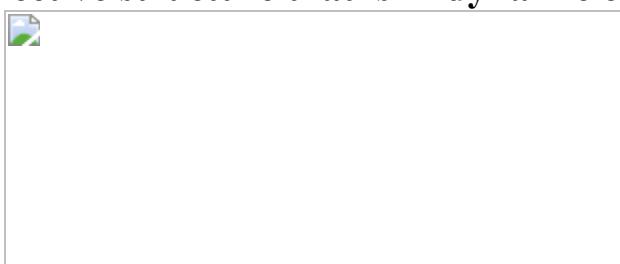


Fig. 2: Metal-ion interactions determine the activity of Dz by condensing the structure and bringing the catalytic loop into an active conformation.

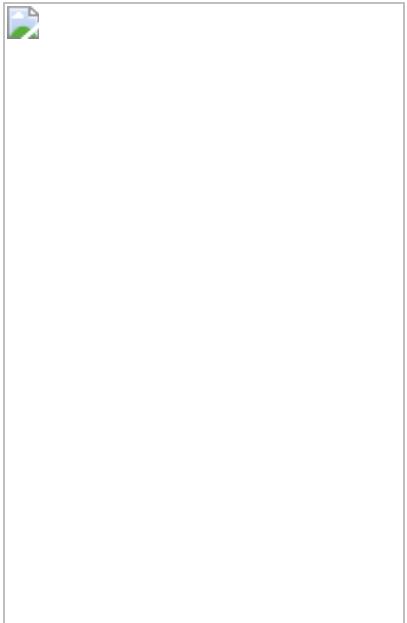
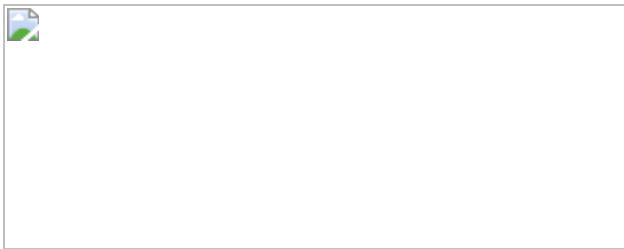


Fig. 3: The catalytic reaction involves multiple rate-limiting transient intermediate states that offer suitable targets to increase the activity of Dz.



Data availability

NMR resonance assignments of Dz^{5C}-RNA^{2'F} are available via the BMRB with the accession code: [34654](#). Atomic coordinates of the Dz^{5C}-RNA^{2'F} ensemble (cluster I) are deposited in the PDB under the accession code: [7PDU](#). Coordinates of additional cluster and scripts for data analysis are available from the corresponding author upon reasonable request.

Code availability

The codes used for structure calculation, MD simulations and simulation of real-time NMR data are available from the corresponding author upon reasonable request.

References

1. 1.

Santoro, S. W. & Joyce, G. F. A general purpose RNA-cleaving DNA enzyme. *Proc. Natl Acad. Sci. USA* **94**, 4262–4266 (1997).

2. 2.

Silverman, S. K. Catalytic DNA: scope, applications, and biochemistry of deoxyribozymes. *Trends Biochem. Sci* **41**, 595–609 (2016).

3. 3.

Rosenbach, H. et al. Molecular features and metal ions that influence 10-23 DNAzyme activity. *Molecules* **25**, 3100 (2020).

4. 4.

Cuenoud, B. & Szostak, J. W. A DNA metalloenzyme with DNA ligase activity. *Nature* **375**, 611–614 (1995).

5. 5.

Pradeepkumar, P. I., Höbartner, C., Baum, D. A. & Silverman, S. K. DNA-catalyzed formation of nucleopeptide linkages. *Angew. Chemie Int. Ed.* **47**, 1753–1757 (2008).

6. 6.

Shen, Z. et al. A catalytic DNA activated by a specific strain of bacterial pathogen. *Angew. Chemie Int. Ed.* **55**, 2431–2434 (2016).

7. 7.

Torabi, S.-F. et al. In vitro selection of a sodium-specific DNAzyme and its application in intracellular sensing. *Proc. Natl Acad. Sci. USA* **112**, 5903–5908 (2015).

8. 8.

Ponce-Salvatierra, A., Wawrzyniak-Turek, K., Steuerwald, U., Höbartner, C. & Pena, V. Crystal structure of a DNA catalyst. *Nature* **529**, 231–234 (2016).

9. 9.

Liu, H. et al. Crystal structure of an RNA-cleaving DNAzyme. *Nat. Commun.* **8**, 2006 (2017).

10. 10.

Victor, J., Steger, G. & Riesner, D. Inability of DNAzymes to cleave RNA *in vivo* is due to limited Mg²⁺ concentration in cells. *Eur. Biophys. J.* **47**, 333–343 (2018).

11. 11.

Young, D. D., Lively, M. O. & Deiters, A. Activation and deactivation of DNAzyme and antisense function with light for the photochemical regulation of gene expression in mammalian cells. *J. Am. Chem. Soc.* **132**, 6183–6193 (2010).

12. 12.

Wang, Y., Nguyen, K., Spitale, R. C. & Chaput, J. C. A biologically stable DNAzyme that efficiently silences gene expression in cells. *Nat. Chem.* **13**, 319–326 (2021).

13. 13.

Liu, C. et al. Self-assembly of copper–DNAzyme nanohybrids for dual-catalytic tumor therapy. *Angew. Chemie Int. Ed.* **60**, 14324–14328 (2021).

14. 14.

Wang, Y., Liu, E., Lam, C. H. & Perrin, D. M. A densely modified M^{2+} -independent DNAzyme that cleaves RNA efficiently with multiple catalytic turnover. *Chem. Sci.* **9**, 1813–1821 (2018).

15. 15.

Nowakowski, J., Shim, P. J., Prasad, G. S., Stout, C. D. & Joyce, G. F. Crystal structure of an 82-nucleotide RNA-DNA complex formed by the 10-23 DNA enzyme. *Nat. Struct. Biol.* **6**, 151–156 (1999).

16. 16.

Zaborowska, Z., Fürste, J. P., Erdmann, V. A. & Kurreck, J. Sequence requirements in the catalytic core of the “10–23” DNA enzyme. *J. Biol. Chem.* **277**, 40617–40622 (2002).

17. 17.

Vögeli, B., Olsson, S., Güntert, P. & Riek, R. The exact NOE as an alternative in ensemble structure determination. *Biophys. J.* **110**, 113–126 (2016).

18. 18.

Elstner, M., Hobza, P., Frauenheim, T., Suhai, S. & Kaxiras, E. Hydrogen bonding and stacking interactions of nucleic acid base pairs: a density-functional-theory based treatment. *J. Chem. Phys.* **114**, 5149–5155 (2001).

19. 19.

Kisseleva, N., Khvorova, A., Westhof, E. & Schiemann, O. Binding of manganese(II) to a tertiary stabilized hammerhead ribozyme as studied by electron paramagnetic resonance spectroscopy. *RNA* **11**, 1–6 (2005).

20. 20.

Rosenbach, H. et al. Influence of monovalent metal ions on metal binding and catalytic activity of the 10–23 DNAzyme. *Biol. Chem.*

402, 99–111 (2020).

21. 21.

Balbach, J. et al. Protein folding monitored at individual residues during a two-dimensional NMR experiment. *Science* **274**, 1161–1163 (1996).

22. 22.

Etzkorn, M., Böckmann, A. & Baldus, M. Kinetic analysis of protein aggregation monitored by real-time 2D solid-state NMR spectroscopy. *J. Biomol. NMR* **49**, 121–129 (2011).

23. 23.

Nawrot, B. et al. Mapping of the functional phosphate groups in the catalytic core of deoxyribozyme 10-23. *FEBS J.* **274**, 1062–1072 (2007).

24. 24.

Horton, T. E., Clardy, D. R. & DeRose, V. J. Electron paramagnetic resonance spectroscopic measurement of Mn²⁺ binding affinities to the hammerhead ribozyme and correlation with cleavage activity. *Biochemistry* **37**, 18094–18101 (1998).

25. 25.

Hunsicker-Wang, L., Vogt, M. & Derose, V. J. EPR methods to study specific metal-ion binding sites in RNA. *Methods Enzymol.* **468**, 335–367 (2009).

26. 26.

Rinnenthal, J., Klinkert, B., Narberhaus, F. & Schwalbe, H. Modulation of the stability of the *Salmonella* fourU-type RNA thermometer. *Nucleic Acids Res.* **39**, 8258–8270 (2011).

27. 27.

Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).

28. 28.

Keller, R. L. J. *Computer Aided Resonance Assignment Tutorial* (Cantina, 2004).

29. 29.

Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).

30. 30.

Schubert, S. RNA cleaving ‘10-23’ DNAzymes with enhanced stability and activity. *Nucleic Acids Res.* **31**, 5982–5992 (2003).

31. 31.

Thakur, C. S., Luo, Y., Chen, B., Eldho, N. V. & Dayie, T. K. Biomass production of site selective ¹³C/¹⁵N nucleotides using wild type and a transketolase *E. coli* mutant for labeling RNA for high resolution NMR. *J. Biomol. NMR* **52**, 103–114 (2012).

32. 32.

Orts, J., Vögeli, B. & Riek, R. relaxation matrix analysis of spin diffusion for the NMR structure calculation with eNOEs. *J. Chem. Theory Comput.* **8**, 3483–3492 (2012).

33. 33.

Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Marius Clore, G. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73 (2003).

34. 34.

Valafar, H. & Prestegard, J. H. REDCAT: a residual dipolar coupling analysis tool. *J. Magn. Reson.* **167**, 228–241 (2004).

35. 35.

Huang, K. et al. Solution structure of the MEF2A–DNA complex: structural basis for the modulation of DNA bending and specificity by MADS-box transcription factors. *EMBO J.* **19**, 2615–2628 (2000).

36. 36.

Pimenta, J. et al. NMR solution structure and SRP54M predicted interaction of the N-terminal sequence (1-30) of the ovine Doppel protein. *Peptides* **49**, 32–40 (2013).

37. 37.

Case, D. A. et al. *Amber 2020* (Univ. California, 2020).

38. 38.

Wang, J., Cieplak, P. & Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **21**, 1049–1074 (2000).

39. 39.

Bas, D. C., Rogers, D. M. & Jensen, J. H. Very fast prediction and rationalization of pKa values for protein-ligand complexes. *Proteins* **73**, 765–783 (2008).

40. 40.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

41. 41.

Joung, I. S. & Cheatham, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **112**, 9020–9041 (2008).

42. 42.

Schott-Verdugo, S. & Gohlke, H. PACKMOL-Memgen: a simple-to-use, generalized workflow for membrane-protein-lipid-bilayer system building. *J. Chem. Inf. Model.* **59**, 2522–2528 (2019).

43. 43.

Li, P., Roberts, B. P., Chakravorty, D. K. & Merz, K. M. Rational design of particle mesh Ewald compatible Lennard-Jones parameters for +2 metal cations in explicit solvent. *J. Chem. Theory Comput.* **9**, 2733–2748 (2013).

44. 44.

Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: an N ·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).

45. 45.

Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).

46. 46.

Hanke, C. A. & Gohlke, H. In *Computational Methods for Understanding Riboswitches* Vol. 553 (eds. Chen, S.-J. & Burke-Aguero, D. H. B.) 163–191 (Academic Press, 2015).

47. 47.

Roe, D. R. & Cheatham, T. E. PTraj and CPPtraj: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).

48. 48.

Richter, C. et al. Determination of sugar conformation in large RNA oligonucleotides from analysis of dipole–dipole cross correlated relaxation by solution NMR spectroscopy. *J. Biomol. NMR* **15**, 241–250 (1999).

49. 49.

Steger, G. Thermal denaturation of double-stranded nucleic acids: prediction of temperatures critical for gradient gel electrophoresis and polymerase chain reaction. *Nucleic Acids Res.* **22**, 2760–2768 (1994).

50. 50.

Sugimoto, N. et al. Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry* **34**, 11211–11216 (1995).

51. 51.

Poland, D. Recursion relation generation of probability profiles for specific-sequence macromolecules with long-range correlations. *Biopolymers* **13**, 1859–1871 (1974).

Acknowledgements

We acknowledge access to the Jülich-Düsseldorf Biomolecular NMR Center. We thank H. Arthanari, A. Boeszoermenyi and T. Viennet for valuable support regarding the ^{19}F -NMR experiments; R. Biehl for helpful discussions; and R. Kümmerle for data acquisition at 1.1 GHz. H.G. is grateful for computational support and infrastructure provided by the "Zentrum für Informations-und Medientechnologie" (ZIM) at the Heinrich Heine University Düsseldorf and the John von Neumann Institute for Computing (NIC) (user ID: HKF7, VSK33). This work was supported by

the German Research Foundation (DFG) (ET 103/2-1, ET 103/2-2, ET 103/4-1, and ET 103/5-1) to M.E., the Chemical Industry Fund (Li 196/05 to I.S. and Hoe 700080 to H.R.), the German Academic Scholarship Foundation (to H.R.), the Bayer AG (Grants4Ag program to I.S.), and the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement number 660258 to A.V. The Center for Structural Studies is funded by the DFG (grant number 417919780).

Author information

Author notes

1. Aldino Viegas

Present address: UCIBIO, Department of Chemistry, NOVA School of Science and Technology, Universidade Nova de Lisboa, Caparica, Portugal

Affiliations

1. Institut für Physikalische Biologie, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

Jan Borggräfe, Julian Victor, Hannah Rosenbach, Aldino Viegas, Detlev Riesner, Gerhard Steger, Ingrid Span & Manuel Etzkorn

2. Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich, Jülich, Germany

Jan Borggräfe, Holger Gohlke & Manuel Etzkorn

3. Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

Christoph G. W. Gertzen, Mohanraj Gopalswamy & Holger Gohlke

4. Center for Structural Studies (CSS), Heinrich Heine University
Düsseldorf, Düsseldorf, Germany

Christoph G. W. Gertzen

5. Institute of Physical and Theoretical Chemistry, University of Bonn,
Bonn, Germany

Christine Wuebben & Olav Schiemann

6. Bruker Switzerland AG, Fällanden, Switzerland

Helena Kovacs

7. John von Neumann Institute for Computing (NIC), Jülich
Supercomputing Centre (JSC) and Institute of Bio- and Geosciences
(IBG-4: Bioinformatics), Forschungszentrum Jülich, Jülich, Germany

Holger Gohlke

8. Department of Chemistry and Pharmacy, Bioinorganic Chemistry,
Friedrich Alexander University Erlangen-Nürnberg, Erlangen,
Germany

Ingrid Span

9. Jülich Center for Structural Biology (JuStruct), Forschungszentrum
Jülich, Jülich, Germany

Manuel Etzkorn

Contributions

J.B., J.V., H.R., A.V., C.G.W.G., C.W., M.G., H.K. and M.E. conducted the experiments and/or performed simulations, and, together with G.S., O.S., H.G. and I.S., analysed the data. NMR data acquisition and processing were carried out by J.B., A.V., M.G. and M.E. Activity assays were performed by J.V., H.R., H.K. and M.E. MD simulations were performed by C.G.W.G.

EPR spectra were recorded by C.W. J.B., J.V., H.R., D.R., H.G. and M.E. wrote the manuscript. All authors designed the experiments and commented on the manuscript.

Corresponding author

Correspondence to [Manuel Etzkorn](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Harald Schwalbe, Scott Silverman and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

[Extended Data Fig. 1 Properties of used constructs and validation of applied modifications.](#)

a) Table of used constructs. b) Gel electrophoretic analysis of substrate cleavage of indicated Dz variants (top, using fluorescein-labelled RNA) and of effects of 2'F stabilization (bottom, using GelRed staining). c) Schematic picture of 2'F RNA modification (left) and affected protons in [¹H,¹H]-TOCSY fingerprint spectra (Dz^{5C}-RNA^{2'F}, blue and Dz^{5C}-RNA, black). The 2'F modification at rG₀ induces CSP only in its direct proximity, i.e., at

the H5 and H6 position of rU₋₁. While the absence of ¹³C enrichments in the RNA substrate impedes accurate determination of potential effects of the 2'F modification on the pseudorotation phase around the cleavage site⁴⁸, analysis of the CSP pattern induced by the 2'F modification at all resolved ¹H positions within the complex (d) confirms that the substrate stabilization does not alter the overall structure of the precatalytic complex. e) Nucleotide-specific ratio of peak intensities in the presence and absence of 1 mM Mg²⁺ for Dz^{5A}-RNA^{2'F} (red) and Dz^{5C}-RNA^{2'F} (blue). To enable a reliable comparison between different nucleotides, the changes of the cross-peak intensities of the correlation between H1' and H6/H8 are shown for each Dz nucleotide. The peak disappearing in the loop region is linked to exchange processes occurring in the NMR intermediate-exchange regime, whereas the otherwise observed CSPs reveal exchange processes in the NMR fast-exchange regime. It can be concluded that Mg²⁺ resides longer within the catalytic loop of the 5A variant, possibly facilitating cleavage. f) Validation of labelling efficiency and cleavage capabilities of the click chemistry approach used for PRE spin labelling. A schematic of the used copper-catalysed click reaction using TEMPO-azide (2) and 5-ethynyl-2'-deoxyuridine (EdU) (1) (f, top). The EdU was used to replace one selected thymine nucleotide in the Dz sequence. To test labelling efficiency and cleavage capabilities, a FAM-azide was used, enabling direct detection via SDS-PAGE (f, bottom). To evaluate the labelling efficiency, identical amounts of a commercial FAM-labelled Dz were loaded on the indicated lane. g) Validation of activity of Dz^{6xF}. Denaturing SDS-PAGE results of time-dependent substrate cleavage of Dz^{5C} without and with the six 2'-¹⁹F substitutions (Dz^{6xF}). The data demonstrate that the fluorine atoms do neither affect the Dz's cleavage activity nor Mg²⁺-dependency.

Extended Data Fig. 2 NMR resonance assignment of Dz^{5C}-RNA^{2'F}.

- a) Extracts of a [¹H, ¹H]-NOESY spectrum showing sequential correlations of the Dz's sequence dG₊₂ to dT₊₉ as indicated in b). c) [¹H, ¹³C]-HSQC spectrum showing assignments for 1'-CH groups (see [methods](#) for more details).

Extended Data Fig. 3 Molecular properties of the precatalytic state.

a) Presence of amino peaks as indicators of stable hydrogen-bond formation (e.g., Watson-Crick base pairing). The spectral extract of a 2D NOESY spectrum (right) containing cross correlations between cytosine H5 and the amino group as schematically highlighted (left). The data show clear signals for all base-paired cytosines present in the binding arms, whereas no signal was detected for any of the five cytosines of the Dz's loop, indicating the absence of stable hydrogen bonding of the respective amino groups. (b,c,d) Temperature-dependent TOCSY NMR peak intensities at increasing temperatures for indicated nucleotides in either single-stranded Dz^{5C} (b) or in the precatalytic Dz^{5C}-RNA^{2'F} complex (c,d). Deviation from a continuous intensity increase indicates denaturing of a locally stabilized structure (see Supplementary Discussion for more details). T4 shows a different behaviour than its neighbours (d). e) Comparison of [¹H,¹H]-TOCSY fingerprint spectra of DNA^c-RNA (cyan) and DNA^{5C}-RNA^{2'F}(black). Assignments for most affected nucleotides are indicated and show that in the absence of the catalytic loop, the respective peaks around the cleavage site shift into the region characteristic for a regular duplex structure as found for the rest of the binding arms. f) Distribution of chemical shift perturbations (CSPs) induced by the presence of the catalytic loop. CSPs for all assigned protons of the binding arm region between Dz^{5C}-RNF^{2'F} and DNA^c-RNA are shown as contour plot (left) or mapped on the determined structure of the precatalytic complex (right). The latter demonstrates that the regions with strong CSPs agree very well with the regions that differ from the classical A-form helix in the precatalytic complex. g) Schematic representation summarizing the NMR-based structure determination of the precatalytic complex. Indicated distance and homology restraints were combined in an *ab initio* structure calculation, followed by a cluster analysis and conformational filtering using RDCs, resulting in a structural ensemble (cluster I) that best resembles all experimental data (see [Methods](#) for details and Extended Data Fig. 4 for representative experimental data).

Extended Data Fig. 4 NMR-based structure determination of Dz^{5C}-RNA^{2'F}.

a,b) Experimental data and the resulting analysis of the PRE experiments. a) [¹H,¹H]-NOESY spectral extracts of a sample with a single TEMPO label at position T8 before (red) and after reducing the nitroxide spin label via the addition of ascorbic acid (black). b) The resulting intensity ratios between the two spectra for each resolved proton in the three samples. The marker size (in I/I₀ direction) indicates the error margin (SD) obtained from the analysis of values for all resolved cross peaks of the respective proton. Only protons with at least two different resolved cross peaks were considered. c) Subset of NOE-buildup curves used for the eNOE approach. Normalized NOE intensities recorded with mixing times between 40 and 800 ms (filled circles) and buildup curves (lines) determined by eNORA³². [¹H,¹H]-NOESY cross peaks for the same spin system occurring above (red) or below (green) the diagonal were considered for most NOE contacts. Notably, even short and geometrically fixed distances show considerable differences in their buildup behaviour, demonstrating the necessity to incorporate cross-relaxation effects. Calculated cross-relaxation rates (σ) and extracted eNOE distance range (within 20% error margin) (blue) are indicated. d-i) ¹⁹F-based NMR experiments. d) 1D ¹⁹F-NMR spectra of Dz^{5C}-RNA^{2'F} (top) and of Dz^{5C} variant containing six 2'-¹⁹F modifications at position G₋₆, G₋₅, G2, C7, A11, and G14 (Dz^{6xF}, bottom) in complex with RNA^{2'F}. While [¹⁹F,¹⁹F]-NOESY spectra of Dz^{6xF}-RNA^{2'F} did not show any detectable peaks (data not shown), ¹H-detected (e) or ¹⁹F-detected (f) [¹H,¹⁹F]-HOESY spectra show a limited number of distinct cross correlations that were used for sequential resonance assignments as well as long-range distance restraints. g) H1' and H2' protons of the respective fluorinated ribose moieties can be identified in a [¹H,¹H]-NOESY spectrum via the peak splitting induced by the strong J_{FH} coupling. h+i) To overcome sensitivity limitations of the 2D HOESY correlations, ¹⁹F-saturated and ¹H-detected STD NMR was used. h) Resulting ¹⁹F-STD spectrum of Dz^{5C}-RNA^{2'F}. i) ¹⁹F-STD spectra of Dz^{6xF}-RNA^{2'F} using the indicated ¹⁹F saturation frequencies (color code refers to assignment in panel d). j-m) Characterization of residual dipolar couplings (RDCs). j) Example of RDC-

induced frequency shifts in the precatalytic complex. The section shows the H5-C5 cross peak of loop position C13 in a [$^1\text{H}, ^{13}\text{C}$]-HSQC spectra recorded in the absence (black) and presence (red) of Pfl phage at 20 °C using ^{13}C isotope-labelled Dz $^{5\text{C}}$ in complex with unlabelled RNA $^{2\text{'F}}$. k) 1D cross section of the cross peaks shown in j) representing experimental limitations due to linewidths, peak overlap and signal-to-noise effects. l,m) Correlation plot of observed and back-calculated RDC constants for a non-matching structure of cluster III (l) and the improved correlation of the structure in cluster I (m). Although experimental limitations introduce larger error margins in the determined values (as visible by RMSD values), an RDC effect is still apparent (as visible by a considerably increased correlation).

Extended Data Fig. 5 Characterization of NMR-derived structural ensemble.

a) Results of cluster analysis after *ab initio* structure calculation without usage of RDCs. The 20 lowest energy structures (out of 200 calculated) were sorted according to their relative orientation of the Dz loop according to: Full winding around the RNA (cluster I), positioning downstream (cluster II) or upstream (cluster III) of the cleavage site, or other (cluster IV). Numbers below the structures refer to the number of structures in the respective fold. b) Same procedure as in a) but including RDC data. While cluster I is already the most populated in the absence of RDC restraints, it becomes by far the dominant fold after RDCs are included. In general, RDCs can be seen as one of the most reliable NMR-based probes of relative molecular orientation. As such, and albeit the remaining limitation in spectral quality and accuracy (Extended Data Fig. 4j,k), the observed effects on cluster selection appear significant. c) Overlay of the four lowest energy structures (out of 100) for each cluster (cluster I-III) after a more rigorous structural refinement and energy minimization.

Extended Data Fig. 6 Conformational plasticity of the precatalytic complex.

a) Root mean square fluctuations (RMSF) of Dz^{5C}-RNA during MD simulations. The mean RMSF \pm SEM (black area) for each nucleotide over all replicas is shown. The orange area highlights the nucleotides of the catalytic loop, and the blue areas represent nucleotides surrounding the cleavage site. b) Average structure (cartoon) from MD simulations overlaid with the structural ensemble generated from the MD trajectory (semi-transparent ribbon, with structures taken every 10 ns). c) Overlay of all seven lowest energy Dz^{5C}-RNA^{2'F} structures belonging to cluster I and obtained with loose-restraining in the ab initio structure calculation. In contrast to the ensembles shown in Extended Data Fig. 5, here, the loop region was excluded from structure alignment. It can be seen that the core region, including the binding arms and the orientation of the cleavage site, is reasonably well defined with an all-atom RMSD of 2.6 Å. On the contrary, the catalytic-loop region displays pronounced variations in the position within the different structures, providing an estimate of the allowed conformational space. Noteworthy, in direct comparison to the better-defined regions, the loop region does not display largely increased dynamics in the detected time regimes (Fig. 1f) and it still shows stacking interactions of the nucleotides with their neighbours, indicating locally defined conformations (Extended Data Figs. 9d and Supplementary Figure 1). In line with the MD-derived picture, it can hence be assumed that the Dz shows increased conformational plasticity in the loop region compared to the binding arm and, particularly, the cleavage-site region.

[Extended Data Fig. 7 Metal-ion binding to the precatalytic complex.](#)

a) Mg²⁺ binding induces chemical shift perturbations (CSPs) in the fingerprint spectrum of Dz^{5C}-RNA^{2'F} indicative of binding and structural rearrangements occurring in a fast-exchange regime. b) Plot of obtained nucleotide-specific K_D values (black, fitted values \pm standard error) as well as the magnitude of Mg²⁺-induced maximal CSP (red). Note that CSPs report on direct interaction and/or relayed conformational changes. Therefore, the obtained K_D values can only serve as a reporter on effects (either direct or indirect) induced by Mg²⁺ binding. c) Mg²⁺-dependent cleavage activity as seen by FRET. d) Plot of atom-specific Mg²⁺-induced

CSPs along the Dz–RNA sequence. e) Comparison between Mn²⁺- and Mg²⁺-induced activity of Dz^{5C}. f) Mn²⁺ binding in the presence of excess Mg²⁺ does not induce CSPs but clear PRE effects allowing to distinguish between binding and structural changes. g) Different PRE rates for selected nucleotides resulting from differential distances to the Mn²⁺-binding site. h) Plot of atom-specific Mn²⁺-induced PRE rates along the Dz–RNA sequence. i) Likelihood of per-nucleotide binding of Mg²⁺ to the precatalytic complex during multiple microsecond-long MD simulations. j) Direct NOE cross peaks between [Co(NH₃)₆]³⁺ and Dz^{5C}–RNA^{2'F}. The dashed line indicates the frequency of the hexamine protons. In line with the observed low affinities and fast exchange rates these NOEs point to interactions of the hydration shell of the divalent metal ions with the nucleotides. The observed interactions clearly involve nucleotides of metal-ion binding region I and likely also metal-ion binding region III, confirming that central metal-ion binding sites of the precatalytic Dz–RNA complex adopt a conformation capable of accommodating the hydration shell of divalent metal ions. k) PAGE analysis of Dz^{5C} activity in the presence of indicated concentrations of different metal ions. Interestingly, while Co²⁺ ions can fully activate the Dz and promote substrate cleavage, [Co(NH₃)₆]³⁺ does not mediate the Dz cleavage reaction but does inhibit Mg²⁺-mediated catalysis. These inhibiting effects are in line with competition for the observed same binding sides. However, the data also suggest that a hydration shell mimic does not maintain a catalytically active state. l) same as i) but, instead of cluster 1 (as used in i), structures of cluster 2–4 were used as starting point of the MD simulations. It can be seen that the MD data using cluster 1 (i) agree considerably better with the experimental data (h) than the other cluster (l), validating the cluster selection based on RDC data.

Extended Data Fig. 8 Experimental links between structure and function of the precatalytic complex.

- a) The structural features of the scaffolding involve close contacts between the two binding arms, non-base pairing at position +1, and metal-ion binding in region I. b) The clear absence of respective peaks in the NOE

pattern confirming flip out of dG₊₁. Unlike to all nucleotides with expected in-register stacking, no cross correlations of dG₊₁ to dG₊₂ is detected (which should appear along the dashed blue line). Stacking pattern for dA₋₆ – dT₊₉ is shown as positive control (green arrows). c) Effects of mutations at positions +1 and -1 on Dz^{5C} activity confirming that position +1, unlike position -1, does not form an essential Watson-Crick base pair. d) Example of NMR data (extracts of NOESY spectra) confirming close spatial proximity of dG₋₅ and dG₊₂ after Mg²⁺ binding. e) Atom-specific PRE rates obtained from Mn²⁺ titration for DNA^c–RNA (top) compared to respective nucleotides in Dz^{5C}–RNA^{2'F} (bottom). The data demonstrate that the presence of the catalytic loop dramatically changes the M²⁺-binding behaviour of the arms from a rather diffuse pattern (top) to the defined binding region that forms the basis of the scaffolding step (bottom). f) Top view on the precatalytic structure focusing on the cleavage-site surroundings. g) NOE-buildup rates strongly indicate *syn*-conformation of dA₋₁. The schematic model shown on top visualizes the considered inter-proton interactions with either fixed distances (H1'-H2', purple, and H1'-H2'', red) or distances strongly depending on the χ -angle (H1' to indicated base proton, blue). h) NOE pattern indicative of the Mg²⁺-induced flip out of T4. Note the same behaviour is observed for all ten resolved inter-nucleotide correlations of T4 to its neighbours. i) Changes of cleavage activity by mutations in the 5' side of the catalytic loop (metal ion-binding site II). Mutations at position 5 serve as a reference; all other mutations are variants of Dz^{5C}. Data are presented as mean values +/- SD of triplicate experiments. j) Comparison of the cleavage activity of the variants Dz^{5A} (black), Dz^{5C} (blue) and Dz^{5G} (red) in the presence of 3 mM Mg²⁺ (data for Dz^{5A} and Dz^{5C} are identical to respective conditions shown in Fig. [1b](#)). k) Simplified schematic model highlighting the central features of the 3D structure. l-n) cw EPR characterization of Mn²⁺ binding. l) Double integrated EPR signal of Mn²⁺ in the absence (blue) and presence of 40 μ M Dz^{5C}–RNA^{2'F} complexes (black). Full range and zoom into lower concentrations are shown. m,n) Fit of experimentally determined binding behaviour (black data points) for either full Mn²⁺ concentration range (left panels) or only the higher affinity binding sites (right panels). Data were fitted to different binding models (red curves, see [Methods](#) and

Supplementary Discussion for more details on the applied data fitting). n) Three cooperative binding sites were assumed, resulting in high- and low-affinity binding sites and overall the best fit to the obtained data (left). In the right panel, a model with only one cooperative binding site was assumed, and only the shown data range was considered. The shown fit captures the key features of the binding behaviour with minimal number of parameters.

Extended Data Fig. 9 Properties of central Dz states in the catalytic cycle.

a) NMR fingerprint region of indicated states of Dz^{5C} and its substrate in the catalytic cycle. Ellipses as shown in Fig. 1d are overlaid as a reference in each spectrum. The spectra allow to follow complex formation, effects of Na⁺ and Mg²⁺, RNA cleavage and RNA dissociation. Schematic models below spectra illustrate the respective (simplified) conditions and structural features. b) Comparison of the Dz fingerprint region before (red) and after (black) addition of RNA^{2'F}. c-d) Analysis of sequential contacts occurring in NOESY spectra of free (red) and complexed Dz (black). c) Schematic of analyzed sequential contacts. d) Extract of spectral region representative of (sequential) cross correlations of the nucleotides' base protons (H6/H8). The data indicate that a central region of the catalytic loop already adopts a (minimally structured) conformation that resembles features found for the same region after complex formation with its substrate (see Supplementary Discussion for more details). e-k) Effects of monovalent (Na⁺) metal ions. In the absence of (mono- and divalent) metal ions additional complex conformations are found in the NMR fingerprint spectrum (cyan arrows in e, top). These conformations disappear after addition of either 100 mM Na⁺ (e, middle) or 1 mM Mg²⁺ (e, bottom) and do not represent a catalytically relevant state (f, as visible from extrapolation of data from Extended Data Fig. 7a, grey arrow). It is tempting to speculate that this off-pathway state could represent Dz–RNA complexes in an inactive conformation characterized by hybridized binding arms lacking the additional turn of the loop around the substrate. However, other conformations are also possible. In addition to different Dz–RNA complex conformations, the spectrum in the absence of metal ions also shows the occurrence of free RNA and Dz. g)

Analysis of peak volumes representative of signal from free RNA (grey), free Dz (purple), and additional complex conformation (cyan). Data were acquired in the absence of NaCl and the presence of indicated concentrations of MgCl₂. A rather strong effect on complex formation (population and conformations) is found for the initial (low) Mg²⁺ concentrations. As visible in the spectrum recorded at 100 mM NaCl and absence of MgCl₂ (e, middle), both features can also be saturated by NaCl. In other words, under conditions that are more relevant in respect to physiological ionic-strength conditions (e.g. at 100 mM NaCl) complex formation appears to not be a limiting step. However, the data strongly suggest that at low-ionic-strength conditions, strong effects in respect to complex formation and conformation (the latter likely not relevant for catalysis) can be induced by the addition of small amounts of metal ions. This observation may help to clarify previous findings involving a transition from low-ionic strength conditions. h) Results of MgCl₂ NMR-titration for fingerprint region of Dz^{5C}-RNA^{2'F} in the absence (top) or presence (bottom) of 100 mM NaCl (the latter is identical to data in Extended Data Fig. 7a). i) Magnification of the behaviour of two indicated nucleotides. While the endpoints of the titrations are very similar in the absence and presence of 100 mM NaCl, the absence of NaCl leads to an extended trend with increased maximal CSPs (j). A roughly 2-fold increased K_D values for Mg²⁺ binding is observed in the presence of NaCl. k) Correlation of the total CSPs (sum of all resolved CSPs) of the respective titration steps in the absence and presence of 100 mM NaCl. For MgCl₂ concentration above 1 mM, the presence of NaCl does not have a strong effect on the Mg²⁺-induced CSPs. By contrast, below 1 mM MgCl₂, the minimum of the correlation plot deviates from the diagonal (white arrow). The data indicate that 100 mM NaCl has a comparable effect on the Dz^{5C}-RNA^{2'F} complex structure as approx. 0.7 mM MgCl₂. l) Comparison of 2D [¹H-¹H] TOCSY spectral extracts recorded before (black) and after full cleavage (orange). m) The nucleotides with (orange) or without (green) detected multiple conformations after cleavage highlighted on the structure of the precatalytic complex (dotted representation marks nucleotides with residual uncertainties). n) Detected signal intensities and resulting relative populations of the different states occurring in the post-catalytic complex

for indicated nucleotides with increasing temperatures. o) Ratio of NMR peak intensities of the two states for best-resolved nucleotides. The values were calculated after cleavage (orange, data from spectrum in l) as well as from the real-time NMR data during the cleavage reaction (blue, data from spectrum in Extended Data Fig. 10d). The data reveal that the relative population of the respective states is stable over time, including their initial appearance during the cleavage reaction. The observed behaviour can be best explained by a stable equilibrium of slowly exchanging states (see Supplementary Discussion for more details). p) Native PAGE demonstrating cleavage of RNA substrate under NMR conditions (here, the sample after the real-time NMR experiments is shown). q) Theoretical degree of denaturation Θ at 37 °C and denaturation temperatures T_m of Dz complexes with cleaved RNA (calculated with program POLAND⁴⁹ using RNA:DNA hybrid parameters⁵⁰, loop parameters with $\Delta G = 1 \text{ kJ/mol}$, dissociation constant $K_d = 1 \text{ M}^{-1}$, and the given complex concentrations.

Extended Data Fig. 10 Time-resolved atomic-level insights into the catalysis reaction.

a) Series of 1D ^1H spectra recorded after addition of Mg^{2+} to pre-formed non-stabilized Dz $^{5\text{C}}$ -RNA complexes revealing clear time-dependent changes of the peak intensities. b) Extract of 1D spectra recorded on Dz $^{5\text{C}}$ -RNA before (black), directly after (blue), and at indicated time points (up to about 2 h) after Mg^{2+} addition. A spectrum recorded on stabilized Dz $^{5\text{C}}$ -RNA $^{2\text{F}}$, preincubated with Mg^{2+} for three days, serves as a reference for the Mg^{2+} -equilibrated precatalytic complex (red). c) Contour plot of time-resolved 1D NMR data following the cleavage reaction. A magnification of data shown in (a) highlighting two peaks representing the educt (right) and product (left) state is displayed. A clear shift of the peak maximum during the initial stage of the reaction is apparent for the educt state, which is not present for the product state (dashed lines are given as guides). The direction of the observed frequency change follows the CSP induced by Mg^{2+} binding (visible in (b)) and therefore is in line with an increasing effective $\text{Mg}^{2+}:\text{Dz}$ ratio. The data strongly suggest that the post-catalytic complex has a lower affinity for Mg^{2+} ions than the precatalytic complex

leading to an effective Mg^{2+} release before product release. d,e) Series of 2D [1H - 1H] TOCSY spectra recorded before (black) and successively after adding 1 mM $MgCl_2$. The acquisition time of each spectrum was 3 h. Clear peak position, peak shape, and/or intensity changes are present between the first (d, cyan) and second spectrum (e, brown). Only very weak intensity changes are present between second and third spectrum (e, orange). Indicated magnifications compare peak positions before (black) and during Mg^{2+} -induced transition (d, cyan/blue denote positive/negative contour levels) and after full cleavage (e, orange). f) Magnification of signal for rU₊₆ shown in (d). The peak positions from the Mg^{2+} titration obtained on Dz^{5C}-RNA^{2'F} at indicated Mg^{2+} concentrations are shown as purple dashed lines. The regions representing the educt and product peaks in the real-time spectrum are highlighted in brown and orange. The region of the Mg^{2+} -free state is highlighted in grey. The data reveal that the observed transition does not start from the Mg^{2+} -free state but instead from a state that matches the frequencies of the Mg^{2+} -equilibrated state at Mg^{2+} levels between 0.5 and 1 mM. g) Overlay of time-resolved experimental data for one cross peak (T8-H6/H7; positive/negative contours are denoted in blue/cyan, respectively) with best-fit simulated spectrum (red/yellow denote positive/negative contours). Characteristic features of the initial and final states are indicated. The simulations can reproduce the experimental data well. Note that data shown in panels a, b, d and g show enlarged versions of the respective data shown in Fig. 3. h) Difference between experimental and simulated real-time NMR data as a function of different rate constants applied in the simulations. Results are shown individually for each resolved atom. The minimum represents the best fit condition. The nuclei can be divided into two groups, which either show rates that are slower than the FRET rate determined under matching conditions (brown dashed line) and are plotted in blue, or faster rates (plotted in red). Respective nucleotides are mapped on the structure in Fig. 3d) 1D slice of T8(H6-H7) obtained from the indirect dimension of the 2D spectrum recorded during the cleavage reaction (d, cyan). Experimental data (black, representing state C₁) is compared to peak shapes simulated using either a 2-state transition model (blue) or a 3-state transition model (red). Both simulated spectra represent the best fit for the respective model. Note that negative shoulders are better represented by the 3-state model (dotted arrows). j) Difference between the

experimentally obtained data for indicated nucleotides and the simulated data as a function of the rate constant used in the simulation. While for dC₊₅ (grey) the overall fit is close to the experimental data (as visible by the rather low overall values of the difference), the peak is insensitive to changes in the rate constants (as visible by the low variations over the range of applied rate constants). The respective peak, therefore, is not a good sensor of the transition kinetics and was excluded from data interpretation. On the contrary, the data for C3 show a clear minimum and display good sensitivity in respect to changes in rate constants. The respective peak was therefore included in further data analysis. k) Likelihood of interactions of G14 or hexa-hydrated Mg²⁺ with the O2' atom of rG₀ or the O5' atom of rU₋₁ at the RNA cleavage site, when in any conformation of the cleavage site (left) and when the cleavage site is in the in-line attack conformation (angle O2'-P-O5': 130–180°, right). An interaction is considered present if a hetero-atom of G14 or Mg²⁺ is within 5 Å of the target atom. l) Effect of Mg²⁺ binding to the metal ion-binding site II on the frequency of structural features of the RNA cleavage reaction during MD simulations at 20 mM Mg²⁺. m) FRET-based activity assay in the presence of Mg²⁺ (red data points, 1 mM) or Mn²⁺ (blue data points, 0.5 mM) of Dz^{5C} with a 6-thio-modification at either G14 (left) or at G6 (right). While the measured behaviour of 6-thio-G6 is well in line with previous reports²³, the unusual behaviour of 6-thio-G14 would be in line with the acid-base mechanism shown in panel k (right, X₁ = G14).

Supplementary information

Supplementary Information

This file contains Supplementary Text, Supplementary Figs. 1, 2, Supplementary Tables 1, 2 and Supplementary Equations.

Reporting Summary

Peer Review File

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Borggräfe, J., Victor, J., Rosenbach, H. *et al.* Time-resolved structural analysis of an RNA-cleaving DNA catalyst. *Nature* **601**, 144–149 (2022). <https://doi.org/10.1038/s41586-021-04225-4>

- Received: 29 May 2021
- Accepted: 06 November 2021
- Published: 23 December 2021
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04225-4>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

| [Section menu](#) | [Main menu](#) |

Amendments & Corrections

- [**Author Correction: A bimodal burst energy distribution of a repeating fast radio burst source**](#) [15 December 2021]
Author Correction •
- [**Author Correction: Half- and quarter-metals in rhombohedral trilayer graphene**](#) [15 December 2021]
Author Correction •
- [**Publisher Correction: A solar C/O and sub-solar metallicity in a hot Jupiter atmosphere**](#) [15 December 2021]
Publisher Correction •

- Author Correction
- [Published: 15 December 2021](#)

Author Correction: A bimodal burst energy distribution of a repeating fast radio burst source

- [D. Li](#) ORCID: [orcid.org/0000-0003-3010-7661^{1,2}](https://orcid.org/0000-0003-3010-7661) [na1](#),
- [P. Wang](#) ORCID: [orcid.org/0000-0002-3386-7159¹](https://orcid.org/0000-0002-3386-7159) [na1](#),
- [W. W. Zhu](#) ORCID: [orcid.org/0000-0001-5105-4058¹](https://orcid.org/0000-0001-5105-4058) [na1](#),
- [B. Zhang](#) ORCID: [orcid.org/0000-0002-9725-2524³](https://orcid.org/0000-0002-9725-2524),
- [X. X. Zhang](#) ORCID: [orcid.org/0000-0001-9587-631X¹](https://orcid.org/0000-0001-9587-631X),
- [R. Duan](#) ORCID: [orcid.org/0000-0002-9260-1641¹](https://orcid.org/0000-0002-9260-1641),
- [Y. K. Zhang](#) ORCID: [orcid.org/0000-0002-8744-3546^{1,2}](https://orcid.org/0000-0002-8744-3546),
- [Y. Feng](#) ORCID: [orcid.org/0000-0002-0475-7479^{1,2,4}](https://orcid.org/0000-0002-0475-7479),
- [N. Y. Tang](#) ORCID: [orcid.org/0000-0002-2169-0472^{1,5}](https://orcid.org/0000-0002-2169-0472),
- [S. Chatterjee](#) ORCID: [orcid.org/0000-0002-2878-1502⁶](https://orcid.org/0000-0002-2878-1502),
- [J. M. Cordes](#) ORCID: [orcid.org/0000-0002-4049-1882⁶](https://orcid.org/0000-0002-4049-1882),
- [M. Cruces](#) ORCID: [orcid.org/0000-0001-6804-6513⁷](https://orcid.org/0000-0001-6804-6513),
- [S. Dai](#) ORCID: [orcid.org/0000-0002-9618-2499^{1,4,8}](https://orcid.org/0000-0002-9618-2499),
- [V. Gajjar](#) ORCID: [orcid.org/0000-0002-8604-106X⁹](https://orcid.org/0000-0002-8604-106X),
- [G. Hobbs](#) ORCID: [orcid.org/0000-0003-1502-100X⁴](https://orcid.org/0000-0003-1502-100X),
- [C. Jin¹](#),
- [M. Kramer](#) ORCID: [orcid.org/0000-0002-4175-2271⁷](https://orcid.org/0000-0002-4175-2271),
- [D. R. Lorimer](#) ORCID: [orcid.org/0000-0003-1301-966X^{10,11}](https://orcid.org/0000-0003-1301-966X),
- [C. C. Miao](#) ORCID: [orcid.org/0000-0002-9441-2190^{1,2}](https://orcid.org/0000-0002-9441-2190),
- [C. H. Niu](#) ORCID: [orcid.org/0000-0001-6651-7799¹](https://orcid.org/0000-0001-6651-7799),
- [J. R. Niu](#) ORCID: [orcid.org/0000-0001-8065-4191^{1,2}](https://orcid.org/0000-0001-8065-4191),
- [Z. C. Pan](#) ORCID: [orcid.org/0000-0001-7771-2864¹](https://orcid.org/0000-0001-7771-2864),
- [L. Qian](#) ORCID: [orcid.org/0000-0003-0597-0957¹](https://orcid.org/0000-0003-0597-0957),

- [L. Spitler](#) ORCID: [orcid.org/0000-0002-3775-8291⁷](https://orcid.org/0000-0002-3775-8291),
- [D. Werthimer](#)⁹,
- [G. Q. Zhang](#)¹²,
- [F. Y. Wang](#) ORCID: [orcid.org/0000-0003-4157-7714^{12,13}](https://orcid.org/0000-0003-4157-7714),
- [X. Y. Xie](#)¹⁴,
- [Y. L. Yue](#) ORCID: [orcid.org/0000-0003-4415-2148¹](https://orcid.org/0000-0003-4415-2148),
- [L. Zhang](#) ORCID: [orcid.org/0000-0001-8539-4237^{1,15}](https://orcid.org/0000-0001-8539-4237),
- [Q. J. Zhi](#)^{14,16} &
- [Y. Zhu](#)¹

[Nature](#) volume **601**, page E1 (2022)

- 396 Accesses
- 1 Altmetric
- [Metrics details](#)

The [Original Article](#) was published on 13 October 2021

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-021-03878-5>
Published online 5 August 2021

During the preparation of this manuscript, an error arose regarding the total isotropic energy described. In the last paragraph of the main text, the sentence now reading “The total isotropic energy emitted in the 1,652 bursts that we report is 3.4×10^{41} erg” originally read “ 6.4×10^{46} erg.” Similarly, in the Methods section “Energy budget constraint on the synchrotron maser magnetar model,” the first sentence now reading “The total isotropic energy emitted in the 1,652 bursts reported in this paper is 3.4×10^{41} erg,” 3.4×10^{41} now replaces “ 6.4×10^{46} .” In the fourth sentence of the same Methods section, in the text now reading “the total energy emitted … is already $\sim 6.4 \times 10^{46} (\{\eta\}_{-4}^{-1})$ erg,” 6.4×10^{46} replaces “ 2×10^{45} .” In the

Methods “Polarization characteristics” paragraph, the text now reading “The circular polarization of a few brightest bursts is consistent with noise,” the text “of a few brightest bursts” has been inserted in the sentence.

These errors have been corrected in the online version of the Article.

Author information

Author notes

1. These authors contributed equally: D. Li, P. Wang, W. W. Zhu

Affiliations

1. CAS Key Laboratory of FAST, NAOC, Chinese Academy of Sciences, Beijing, China

D. Li, P. Wang, W. W. Zhu, X. X. Zhang, R. Duan, Y. K. Zhang, Y. Feng, N. Y. Tang, S. Dai, C. Jin, C. C. Miao, C. H. Niu, J. R. Niu, Z. C. Pan, L. Qian, Y. L. Yue, L. Zhang & Y. Zhu

2. University of Chinese Academy of Sciences, Beijing, China

D. Li, Y. K. Zhang, Y. Feng, C. C. Miao & J. R. Niu

3. Department of Physics and Astronomy, University of Nevada, Las Vegas, Las Vegas, NV, USA

B. Zhang

4. CSIRO Astronomy and Space Science, Epping, New South Wales, Australia

Y. Feng, S. Dai & G. Hobbs

5. Department of Physics, Anhui Normal University, Wuhu, China

N. Y. Tang

6. Cornell Center for Astrophysics and Planetary Science and Department of Astronomy, Cornell University, Ithaca, NY, USA
S. Chatterjee & J. M. Cordes
7. Max-Planck-Institut für Radioastronomie, Bonn, Germany
M. Cruces, M. Kramer & L. Spitler
8. Western Sydney University, Penrith, New South Wales, Australia
S. Dai
9. Department of Astronomy, University of California Berkeley, Berkeley, CA, USA
V. Gajjar & D. Werthimer
10. Department of Physics and Astronomy, West Virginia University, Morgantown, WV, USA
D. R. Lorimer
11. Center for Gravitational Waves and Cosmology, West Virginia University, Morgantown, WV, USA
D. R. Lorimer
12. School of Astronomy and Space Science, Nanjing University, Nanjing, China
G. Q. Zhang & F. Y. Wang
13. Key Laboratory of Modern Astronomy and Astrophysics (Nanjing University), Ministry of Education, Nanjing, China
F. Y. Wang
14. Guizhou Normal University, Guiyang, China

X. Y. Xie & Q. J. Zhi

15. School of Physics and Technology, Wuhan University, Wuhan, China

L. Zhang

16. Guizhou Provincial Key Laboratory of Radio Astronomy and Data Processing, Guizhou Normal University, Guiyang, China

Q. J. Zhi

Corresponding authors

Correspondence to [D. Li](#) or [B. Zhang](#).

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Li, D., Wang, P., Zhu, W.W. *et al.* Author Correction: A bimodal burst energy distribution of a repeating fast radio burst source. *Nature* **601**, E1 (2022). <https://doi.org/10.1038/s41586-021-04178-8>

- Published: 15 December 2021
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04178-8>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04178-8>

| [Section menu](#) | [Main menu](#) |

- Author Correction
- [Published: 15 December 2021](#)

Author Correction: Half- and quarter-metals in rhombohedral trilayer graphene

- [Haoxin Zhou](#) ORCID: [orcid.org/0000-0003-1235-0035](#)¹ na1,
- [Tian Xie](#) ORCID: [orcid.org/0000-0002-6406-0403](#)¹ na1,
- [Areg Ghazaryan](#)²,
- [Tobias Holder](#) ORCID: [orcid.org/0000-0003-1365-4276](#)³,
- [James R. Ehrets](#)¹,
- [Eric M. Spanton](#)¹,
- [Takashi Taniguchi](#) ORCID: [orcid.org/0000-0002-1467-3105](#)⁴,
- [Kenji Watanabe](#) ORCID: [orcid.org/0000-0003-3701-8119](#)⁵,
- [Erez Berg](#) ORCID: [orcid.org/0000-0001-8956-3384](#)³,
- [Maksym Serbyn](#) ORCID: [orcid.org/0000-0002-2399-5827](#)² &
- [Andrea F. Young](#) ORCID: [orcid.org/0000-0001-5954-8028](#)¹

[Nature](#) volume 601, page E2 (2022)

- 490 Accesses
- 45 Altmetric
- [Metrics details](#)

The [Original Article](#) was published on 01 September 2021

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-021-03938-w>
Published online 1 September 2021

In the version of this Article initially published, there were errors in the author affiliations. Erez Berg (Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot, Israel) and Maksym Serbyn (Institute of Science and Technology, Klosterneuburg, Austria) now show their correct affiliations.

These errors have been corrected in the online version of the Article.

Author information

Author notes

1. These authors contributed equally: Haoxin Zhou, Tian Xie

Affiliations

1. Department of Physics, University of California, Santa Barbara, CA, USA

Haoxin Zhou, Tian Xie, James R. Ehrets, Eric M. Spanton & Andrea F. Young

2. Institute of Science and Technology, Klosterneuburg, Austria

Areg Ghazaryan & Maksym Serbyn

3. Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot, Israel

Tobias Holder & Erez Berg

4. International Center for Materials Nanoarchitectonics, National Institute for Materials Science, Tsukuba, Japan

Takashi Taniguchi

5. Research Center for Functional Materials, National Institute for Materials Science, Tsukuba, Japan

Kenji Watanabe

Corresponding author

Correspondence to [Andrea F. Young](#).

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Zhou, H., Xie, T., Ghazaryan, A. *et al.* Author Correction: Half- and quarter-metals in rhombohedral trilayer graphene. *Nature* **601**, E2 (2022). <https://doi.org/10.1038/s41586-021-04181-z>

- Published: 15 December 2021
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04181-z>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04181-z>

| [Section menu](#) | [Main menu](#) |

- Publisher Correction
- [Published: 15 December 2021](#)

Publisher Correction: A solar C/O and sub-solar metallicity in a hot Jupiter atmosphere

- [Michael R. Line](#) [ORCID: orcid.org/0000-0001-6247-8323](#)^{1,2},
- [Matteo Brogi](#)^{3,4,5},
- [Jacob L. Bean](#)⁶,
- [Siddharth Gandhi](#)^{3,4},
- [Joseph Zalesky](#) [ORCID: orcid.org/0000-0002-2259-4116](#)¹,
- [Vivien Parmentier](#)⁷,
- [Peter Smith](#)¹,
- [Gregory N. Mace](#)⁸,
- [Megan Mansfield](#)⁹,
- [Eliza M.-R. Kempton](#) [ORCID: orcid.org/0000-0002-1337-9051](#)¹⁰,
- [Jonathan J. Fortney](#)¹¹,
- [Evgenya Shkolnik](#)^{1,2},
- [Jennifer Patience](#)¹,
- [Emily Rauscher](#) [ORCID: orcid.org/0000-0003-3963-9672](#)¹²,
- [Jean-Michel Désert](#)¹³ &
- [Joost P. Wardenier](#) [ORCID: orcid.org/0000-0003-3191-2486](#)⁷

[Nature](#) volume 601, page E3 (2022)

- 422 Accesses
- 1 Altmetric
- [Metrics details](#)

The [Original Article](#) was published on 27 October 2021

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-021-03912-6>
Published online 27 October 2021

In the version of this article initially published, refs. [26,27,28](#) were missing from the print version and are provided below. In the online version, refs. 33–36 were disordered, where Coles, 2010; Rothman, 2010; Hargreaves, 2020; Li, 2015 now appear as Rothman, 2010; Hargreaves, 2020; Li, 2015; Coles, 2010 for refs. 33–36, respectively. In the last paragraph of “Extended results, Retrieved constraints” section, in the text now reading “For legacy with past works we also include the ‘classic’^{13,14,29,50} CCF analysis,” ref. 30 initially appeared in place of ref. 29. In the first paragraph of “Extended results, Physical plausibility assessment” section, the sentence now reading “We assess the physical/chemical plausibility of the retrieved quantities with a one-dimensional radiative convective thermochemical equilibrium^{53–55},” ref. 56 initially appeared in place of ref. 55. In the second paragraph of the “Physical plausibility assessment” section, the text now reading “what would be nominally predicted^{56,59}” initially appeared with refs. 59,61. In the final paragraph of the Methods section and in the Extended Data Fig. 7 caption, ref. 67 citations initially appeared as ref. 68. In the “Extended results, Elemental abundance determination and interpretation” section, fourth paragraph, “the O-based measurements” corrects “the-based O measurements.” Further, in Fig. 2a, the lower *x*-axis label “log₁₀*n*_{H2O}” was missing and has been restored. These errors have been corrected in the online version of the article.

References

1. 26.

Atreya, S. K. et al. The origin and evolution of Saturn, with exoplanet perspective. Preprint at <https://arxiv.org/abs/1606.04510> (2016).

2. 27.

Thorngren, D. & Fortney, J. J. Connecting giant planet atmosphere and interior modeling: constraints on atmospheric metal enrichment. *Astrophys. J. Lett.* **874**, L31 (2019).

3. 28.

Burrows, A. & Sharp, C. M. Chemical equilibrium abundances in brown dwarf and extrasolar giant planet atmospheres. *Astrophys. J.* **512**, 843–863 (1999).

Author information

Affiliations

1. School of Earth and Space Exploration, Arizona State University, Tempe, AZ, USA

Michael R. Line, Joseph Zalesky, Peter Smith, Evgenya Shkolnik & Jennifer Patience

2. NASA Astrobiology Institute, Virtual Planetary Laboratory Team, Seattle, WA, USA

Michael R. Line & Evgenya Shkolnik

3. Department of Physics, University of Warwick, Coventry, UK

Matteo Brogi & Siddharth Gandhi

4. Centre for Exoplanets and Habitability, University of Warwick, Coventry, UK

Matteo Brogi & Siddharth Gandhi

5. INAF—Osservatorio Astrofisico di Torino, Turin, Italy

Matteo Brogi

6. Department of Astronomy and Astrophysics, University of Chicago,
Chicago, IL, USA

Jacob L. Bean

7. Atmospheric, Oceanic, and Planetary Physics, Clarendon Laboratory,
Department of Physics, University of Oxford, Oxford, UK

Vivien Parmentier & Joost P. Wardenier

8. Department of Astronomy, University of Texas at Austin, Austin, TX,
USA

Gregory N. Mace

9. Department of Geophysical Sciences, University of Chicago, Chicago,
IL, USA

Megan Mansfield

10. Department of Astronomy, University of Maryland, College Park, MD,
USA

Eliza M.-R. Kempton

11. Department of Astronomy and Astrophysics, University of California,
Santa Cruz, CA, USA

Jonathan J. Fortney

12. Department of Astronomy, University of Michigan, Ann Arbor, MI,
USA

Emily Rauscher

13. Anton Pannekoek Institute for Astronomy, University of Amsterdam,
Amsterdam, the Netherlands

Jean-Michel Désert

Corresponding author

Correspondence to [Michael R. Line](#).

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Line, M.R., Brogi, M., Bean, J.L. *et al.* Publisher Correction: A solar C/O and sub-solar metallicity in a hot Jupiter atmosphere. *Nature* **601**, E3 (2022). <https://doi.org/10.1038/s41586-021-04192-w>

- Published: 15 December 2021
- Issue Date: 06 January 2022
- DOI: <https://doi.org/10.1038/s41586-021-04192-w>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04192-w>

| [Section menu](#) | [Main menu](#) |