

The International Journal of Science / 4 June 2020

# nature



## DRY AND DURABLE

A mechanically robust superhydrophobic surface

Premature births  
The long-term effects of being born exceptionally early

Coronavirus  
16 ethical questions for contact-tracing apps

Different views  
Neuroimage analyses can produce widely divergent results



# 打赏 - JUST FOR FUN

- 支持分享! 一杯咖啡钱, 打赏金额随意, 感谢大家~ :)



资源来自 : <https://github.com/hehonghui/the-economist-ebooks>

# Delaying COP26 is not a reason to delay climate action

**Coronavirus has pushed climate talks back by a year. But action on global warming must not be postponed.**

The first Conference of the Parties under the United Nations climate convention took place in Berlin a quarter of a century ago. By today's standards, it was a relatively modest affair, involving just a few thousand people, including government officials, scientists, environmentalists and journalists. By contrast, the event scheduled for this year in Glasgow, UK – the 26th Conference of the Parties (COP26) – would have attracted some 30,000 participants. But it has had to be postponed because of the coronavirus outbreak. Last week, we learnt that the delay will last for a year, with COP26 now due to take place in November 2021.

The decision to delay was unavoidable: a COP needs representatives of all countries to be present, which would not have been possible if those countries were at different stages of virus transmission and lockdowns. In Glasgow, the conference venue has been temporarily converted to a COVID-19 field hospital. An online meeting was considered, but rejected. Delegates rightly concluded that complex negotiations cannot be conducted using available virtual-meeting technologies.

But delay has risks, and principal among them is slower progress in the struggle against climate change. By the time COP26 was originally scheduled to begin, countries were expected to propose new commitments to bring emissions to net zero. And, at the conference, extra finance for less-wealthy countries was due to be proposed, making good on decades of promises that have not been kept.

The role of carbon markets was also to be discussed, along with nations' responsibility for damages caused by global warming. The meeting had a full agenda, and a delay of a year will have consequences – ultimately making it even harder to limit temperature increases to 1.5–2°C above pre-industrial levels, the main goal of the 2015 Paris climate agreement.

But countries do not need to wait until COP26 to take further action. Indeed, there are signs that some are planning to focus more public investments on green infrastructure and clean energy, rather than doubling down on a past marked by polluting fossil fuels. It is the least they can do.

Last week, the European Commission unveiled proposals for a €750-billion (US\$833-billion) post-coronavirus funding package that includes measures to accelerate the adoption of renewable energy technologies as part of a commitment to reduce emissions to net zero by 2050. The



**By delaying emissions reductions, all countries have done is borrow against the planet's future."**

package will, however, need to be agreed by the 27 member states and the European Parliament.

At the same time, China announced that, for the first time in decades, it will not set a target for economic growth. Not setting such a target will enable policymakers to prioritize innovative ideas for improving the environment and well-being, which can be difficult to do when a country's principal policy focus is economic growth.

When it comes to the actual climate talks, waiting another year does have one potential advantage. Many countries were unlikely to step up with ambitious climate plans this year, even before the COVID-19 crisis complicated matters. This is due, in part, to the uncertainty surrounding the United States' intentions following President Donald Trump's decision to pull out of the Paris agreement. The US position is likely to become more predictable after US voters have delivered their verdict in the presidential election in November.

If the Democrat candidate Joe Biden is elected, the United States is expected to rejoin the Paris accord and restore a more evidence-based approach to its climate-change decision-making. It will also revert to engaging more constructively with other countries. But if Trump is re-elected, that will signal to other countries – especially those in the European Union – that they should organize and press forwards with ambitious climate targets and funding pledges independently of the United States.

The latter scenario is likely to present the UK government with some challenges as it hosts COP26, along with next year's meeting of the Group of Seven large industrialized countries. As the United Kingdom exits the EU, it is loosening its ties to the EU's many collective decision-making structures. At the same time, the government is in active talks with the United States, seeking a closer relationship in areas including trade and research.

## Policy dilemma

On climate policy, the United Kingdom could find itself trapped between a White House urging it to pay due attention to the US perspective, and most other countries, which will be advocating for action on the basis of the Paris consensus. British climate negotiators must do what is best for the planet, and act according to the best available evidence.

Five years ago, countries came together in Paris on a new agreement to curb greenhouse-gas emissions. They knew then that their pledges fell short of what was needed, and agreed to return to the table this year with new commitments. A further delay of a year gives countries more time to plan for more ambitious reductions. It also gives policymakers more time to think strategically as they work to bolster jobs, improve livelihoods and, ultimately, build more-resilient economic systems.

The pressure is on and countries must continue to act with urgency. They might think that they have carved out an extra year to prepare for COP26, but, in practice, there is no extra time. The climate clock is still ticking and by delaying emissions reductions, all countries have done is borrow against the planet's future.

# SpaceX launch will boost International Space Station

**NASA's partnership with SpaceX and its founder Elon Musk should recognize the global nature of space exploration and research.**

**O**n 30 May, tens of millions of space enthusiasts were glued to their screens as SpaceX's Dragon capsule soared into the air above Cape Canaveral, Florida, aboard a Falcon 9 rocket. The following day, as the capsule docked with the International Space Station (ISS), some 422 kilometres above China's border with Mongolia, Robert Behnken and Douglas Hurley made history as the first astronauts to ride a commercial craft into orbit.

This development – a decade in the planning – is undoubtedly an achievement for NASA, and for SpaceX and its reusable rockets. But it is equally a boost for space science and innovation, and especially the enduring value of global cooperation in space research and technology. Amid the jubilation, this aspect of the achievement should be highlighted more.

For NASA, the launch means, among other things, some more money in the bank. Since 2011, when the agency retired the Space Shuttle, NASA has paid Russia up to US\$90 million per person to ferry crews to the ISS aboard the Soyuz craft. Seats on the SpaceX capsule are around two-thirds of this cost, which means that NASA can channel the savings into other priorities, including its ambition to return astronauts to the Moon by 2024.

The weekend's launch also consolidates the position of SpaceX, a company that has mushroomed from start-up to major aerospace player in 18 years. Corporations have been entwined with national space agencies from early on – Grumman (now Northrop Grumman) famously designed and built the lunar module that carried the Apollo astronauts to the Moon's surface. More recently, other companies have flown humans to space. Virgin Galactic, founded by entrepreneur Richard Branson, has pulled off sub-orbital flights and is planning to offer short trips for passengers to experience a few minutes of weightlessness before returning to Earth. But SpaceX has succeeded at the more ambitious goal of carrying people all the way into orbit.

The company has achieved this through nimbleness, an outstanding team of engineers and product designers, and the determination of its founder, Elon Musk. Musk – who is never far from controversy – has had a hand in disrupting two established industries, first as one of the early developers of online payment systems such as PayPal, and later as chief executive of Tesla, the electric-vehicle manufacturer. But few thought he would succeed when he announced

 **Find a new hashtag – something that will resonate with millions around the world."**

his intention to compete with much larger and more-established corporations in space technology. Arguably, SpaceX's most important innovation has been to engineer the Falcon rocket so that it can be reused after launch. Once it has jettisoned its payload, the Falcon returns to Earth and lands, vertically, which other rockets do not do.

Although attention is understandably focused on the launch and docking, the reason for the SpaceX mission to the ISS should not be forgotten – the astronauts' mission is ultimately in the service of science and international research cooperation. Behnken and Hurley will take part in installing a new hardware platform called Bartolomeo, designed by the European Space Agency and Airbus to enable the ISS to host extra science experiments from teams from all over the world.

When big launches grab everyone's attention, it is hard for research to get a hearing. Earlier this year, two NASA astronauts, Christina Koch and Jessica Meir, completed a challenging upgrade of a fundamental physics experiment on the station, the Cold Atom Laboratory – doing in zero gravity what physicists on Earth might have struggled to do. And, last month, the agency's Human Research Program announced plans for extended flights to the ISS, designed to simulate the effects on the human body of a journey to Mars.

## A global endeavour

It's unfortunate that those following the weekend's events did not see or hear much about the ISS's research contributions, or the fact that astronauts have visited the space station from 19 nations – among them Malaysia, the United Arab Emirates and Kazakhstan. They did, however, see SpaceX and NASA promote the #LaunchAmerica hashtag, and they heard NASA's Administrator Jim Bridenstine say: "It's been nine years since we've launched American astronauts on American rockets from American soil."

New space launches – regardless of their country of origin – are often accompanied by a heavy display of national symbols. But it would have been much more powerful, and more uplifting, had the launch also recognized the contributions made by other nations, including Russia, which has been reliably carrying astronauts to the ISS all this time.

From Yuri Gagarin's orbit of Earth in 1961 to the Moon landings of 1969, space has always been an arena of fierce superpower competition – and newer players, not least China, have since come onto the scene. But, in space research, such competition has not prevented nations from cooperating, and that needs to be recognized and celebrated.

There is plenty of opportunity to do so. SpaceX will make its next run to the ISS as early as August. Bridenstine and Musk should use this next mission to demonstrate that space exploration and research are global. At the very least, they should find a new hashtag – something that will resonate with the millions around the world who watched the weekend's launch with awe, and will inspire them to join the next generation of researchers, engineers and astronauts.

# World view



By Diane Coyle

## Economists must collaborate courageously

### To navigate pandemic trade-offs, policymakers need syntheses.

**E**conomists, like researchers in many disciplines, are responding to the urgency of the COVID-19 pandemic. The immediate priorities are understanding the consequences of the crisis for public finances and international trade.

Scholars are scrambling to collect data on how many jobs are lost, what people can afford to buy and what shortages will emerge. Even constructing basic economic statistics such as inflation and gross domestic product is challenging when one-third, say, of activity in the economy has halted. Do we count a furloughed person as in work? What comprises a standard basket of goods when no one is going shopping? We need these measures to understand which groups of people will be intolerably affected so that governments can direct help to them. That's hard to do when standard metrics are having to be rebuilt on the fly.

There are many other pressing questions. When will the health toll of isolation, unemployment or delayed surgery outweigh that caused directly by COVID-19? What are the implications for next year's supplies of staple foods or of higher levels of long-term disability? How quickly can vaccine manufacture be scaled up? What release-from-lockdown strategies are behaviourally and hence politically feasible? Can national governments negotiate with each other to arrive at cooperative, mutually beneficial policies? What can international agencies do to encourage this when geopolitical tensions are rising?

Addressing these questions requires collaboration across many disciplines to synthesize new findings with old – fast. It's time to deliver on the benefits of public investment in research.

The courage to step cautiously into other domains must be welcomed. Economists are notoriously less likely than other social scientists to look outside their own discipline, and medical and natural scientists are not accustomed to looking to the social sciences for insight. The pandemic is changing all that. It has become obvious that the search for viable exit strategies needs biomedical science, epidemiology, public health, behavioural and social psychology, engineering, economics, law, ethics, international relations and political science. Without contributions from all these, navigating toward less-than-disastrous outcomes for well-being – human and planetary – will be impossible.

We economists bring essential insight, namely our habit of thinking about how to improve the terms of unavoidable trade-offs. Every action, even inaction, has an opportunity cost. How can that be minimized, and what are the distributional consequences? These, as much as medical

**Standard metrics are having to be rebuilt on the fly."**

factors, determine political and policy choices.

There is still a tendency to suggest a facile trade-off between lives and livelihoods. To be clear: whatever lockdown policies governments impose, or not, the pandemic will have a devastating economic impact. Every nation is groping for an approach that will save lives, mitigate the economic harm and prove feasible. This is not simple, and there is much we do not know across all fronts.

To share findings fast, the economics community has set up light-touch peer-review outlets, such as the European Economic Association's COVID-19 resource ([see go.nature.com/2zgzezy](http://go.nature.com/2zgzezy)). The United Kingdom's Economic and Social Research Council (ESRC) is kick-starting an observatory to collate and translate research. Funders such as the ESRC and the European Commission's corona platform are rapidly processing research proposals. Academics who have long studied what previously seemed like niche topics, such as the links between financial uncertainty and stress or knots in supply chains, are producing research at extraordinary pace and providing public commentary to communicate their work.

But for any fruits of all this discovery to be a guide to policy, they must be used. I am concerned that government ministers and officials are having to judge for themselves – at a time when they are massively overstretched and under pressure – how to combine insights from various disciplines. Some COVID-19 advisory groups, such as that of the UK government, have too narrow a range of experience, excellent as the individual members might be. This challenge, like other global challenges looming, is the moment for the research community to prioritize synthesizing knowledge.

Sadly, academic incentives work against people who are brave enough to cross into another discipline's territory. Career, funding and publishing structures reward research into small, narrow questions, when the world has big, complex problems. Forbidding argot is prized; accessibility is viewed with suspicion. Universities, research institutes and laboratories are condemning themselves to irrelevance in future – and worse, now – if they do not break the shackles of departments and disciplines, and reward academics for policy relevance, as well as for basic research. This requires institutional reform, which is never easy and too slow for this crisis. High-profile funders such as Wellcome or the Gates Foundation, as well as those in the public sector, could make a powerful point by insisting that programmes integrate disciplines.

Getting good at interdisciplinarity will pay dividends long-term. The specifics will vary, but the need for coordinated research and policy applies to building a post-pandemic social order and to crafting a net-zero economy that limits climate change as far as possible.

**Diane Coyle** is Bennett Professor of Public Policy at the University of Cambridge, UK, and senior independent member of the United Kingdom's Economic and Social Research Council. e-mail: dc700@cam.ac.uk

# News in brief

## SPACEX LAUNCHES ASTRONAUTS — AND A NEW ERA OF HUMAN SPACE FLIGHT

Two NASA astronauts are the first people to travel to orbit in a spacecraft built by a private company. The Crew Dragon capsule, built by SpaceX of Hawthorne, California, is the first new spacecraft in 17 years to carry humans to orbit, and its launch on 30 May was the first to send a crew to space from US soil since NASA retired the Space Shuttle in 2011 (see ‘How humans have reached orbit’). “This is a whole new way of sending people to space,” says Robert Cabana, a former NASA astronaut who is director of the Kennedy Space Center in Florida.

Astronauts Robert Behnken and Douglas Hurley left Florida’s space centre at 3.22 p.m. US Eastern Time in the 8.1-metre-long bullet-shaped capsule, and docked with the International Space Station (ISS) 19 hours later, joining three spacefarers who have been living there since

April. The pair will spend up to four months on the ISS.

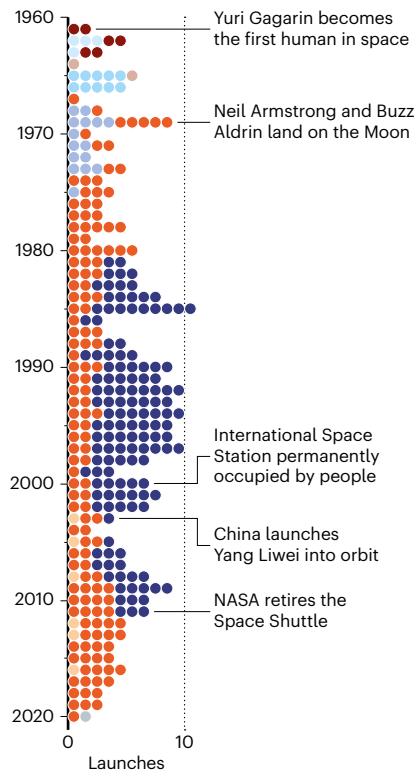
The flight is the culmination of NASA’s push to transition from using its own vehicles to ferry astronauts to the ISS to using spaceships provided by private companies. Since 2011, NASA and other space agencies have relied on Russian Soyuz craft to take people to orbit. The agency will now use the Crew Dragon.

SpaceX has been taking cargo to and from the ISS since 2012, using spacecraft with the same basic design as the Crew Dragon. Its Falcon 9 rockets, which propel the capsules to orbit, have slashed the cost of such launches, in part because they reuse expensive components. NASA is thought to be paying about US\$60 million for each seat on the Crew Dragon, compared with the \$90 million it has been paying for seats on the Soyuz.

## HOW HUMANS HAVE REACHED ORBIT

The Crew Dragon launch marks the ninth time humans have rocketed into orbit on a brand-new spaceship, following four US craft, three Soviet and one Chinese. The Crew Dragon, built by SpaceX for NASA, is the first privately developed spaceship to take people to orbit.

- Mercury
- Gemini
- Apollo
- Space Shuttle
- Vostok
- Voskhod
- Soyuz
- Shenzhou
- Crew Dragon



## CORONAVIRUS DASHES ETHIOPIA’S GOAL OF HOSTING MAJOR AI MEETING

Computer scientists had planned to converge on Addis Ababa in April for the first major artificial-intelligence (AI) conference to be held in an African country. But, like most scientific gatherings this year, the conference ended up being virtual, depriving Ethiopia of a powerful opportunity to boost its research community.

Organizers had hoped that hosting the International Conference on Learning Representations (ICLR) in Ethiopia would make it more accessible to researchers who can’t readily get visas for Western countries, where such meetings are often held. The conference, which focuses on the AI technique deep learning, would have also been important in making the field more diverse in terms of geography, race, gender and sexual orientation – while inspiring the region’s youth to pursue it.

The virtual venue had benefits: online workshops can be easier to join for people who can’t afford to travel, or for those with children or disabilities. But researchers hope that the ICLR can be held in Addis Ababa when the pandemic subsides. “Once big conferences are in-person again, we’ll push for 2022,” says Esube Bekele, a computer-vision architect who is part of Black in AI, a group that advocated holding the meeting in Africa.

## US HALTS PIONEERING CORONAVIRUS-TESTING PROJECT

The research team that first uncovered COVID-19 spreading in US communities has been asked to stop testing for the disease. The decision by the US Food and Drug Administration (FDA) to prevent the SCAN project in Seattle, Washington, from analysing nose swabs sent from people’s homes is likely to be temporary. But it has frustrated public-health initiatives across the country.

Working with the local public-health department, the programme processed 20,000 tests and helped to reveal which communities in Seattle were being hit hardest by COVID-19.

SCAN, which stands for the Seattle Coronavirus Assessment Network, also navigated a regulatory thicket and won key approvals from state authorities. It seemed to be in line with FDA guidelines for emergency use, but the agency then clarified its recommendations, saying that they did not cover tests for which samples need to be transported. The halt is a concern for other researchers developing diagnostics for use outside hospital settings.

SCAN was the first US group to roll out home-sampled tests and to partner effectively with health authorities. As businesses begin to reopen in the country, many argue that such an approach is needed more than ever.



# News in focus



FERMIN RODRIGUEZ/NURPHOTO/GETTY

Many academic laboratories are instituting strict cleaning protocols in response to the coronavirus pandemic.

## RETURN TO THE LAB: SCIENTISTS FACE SHIFTWORK, MASKS AND DISTANCING AS LOCKDOWNS EASE

As scientists around the world return to work, they're encountering new safety rules and awkward restrictions – and sometimes writing the protocols themselves.

By Nidhi Subbaraman

**A**fter her university closed in March, Jeannine Randall sat down to adapt her research plan for a pandemic. Her project to monitor tree swallows through the spring and summer with a team of three scientists would now require travelling to the nesting sites in separate vehicles, using individual work kits, staying 2 metres apart and, of course, sanitizing regularly. When she realized hand sanitizer was in short supply, she made her own batch using ethanol from her lab.

Now, as the university resumes some services, she is putting the plan into action: counting eggs, waiting for hatchlings and watching the birds from daybreak to sundown.

"I think scientists are very well placed in some ways to come up with a protocol that makes sense and then follow it," says Randall, an avian ecologist at the University of Northern British Columbia in Prince George, Canada.

As countries around the world begin lifting pandemic lockdowns, researchers are entering a new phase of work – donning masks with their lab coats, staggering hours in laboratory

spaces and taking shifts on shared instruments. Some universities have created detailed plans to track and test staff, and many have limited the capacity of indoor spaces and the flow of people through hallways and entrances. For others, plans are still taking shape. And whereas some universities have worked in lockstep with governments to formulate safety plans, others have charted their own paths.

University associations in the United States estimate that research there will need a boost of at least US\$26 billion to get back up to speed. Among them, the Association of

## News in focus

American Universities (AAU) in Washington DC is drafting a set of priorities for university leaders to consider as they tackle reopening.

Of more than 3,000 researchers who responded to an online *Nature* reader poll last month, just under half said that they are still on lockdown. Following national policy trends, scientists in the United Kingdom, United States and Brazil were among the most likely to report restrictions, whereas only about 7% of respondents in Germany did so (see 'Getting back to business'). "Now, we are mostly back to normal, apart from wearing face masks and having to maintain a certain distance between each other," says Boyan Garvalov, a cancer researcher at Heidelberg University, who juggles his career with monitoring his children's online learning.

### Gloves and masks

In Italy, which was hit particularly hard early in the global pandemic, nearly 30% of roughly 90 respondents to the *Nature* poll said that they had returned to work, and another 18% reported having worked throughout lockdowns. Cell biologist Paolo Bernardi went into the University of Padua nearly every working day to oversee a skeleton staff and teach his pathophysiology class over Zoom. "Now we are seeing better days," Bernardi says. His lab is at about 50% capacity. University guidelines for resuming work, in effect since 26 April, require distances of 1 metre between people in brief contact, or 2 metres for those in the same room for more than 15 minutes; masks are to be worn at all times and gloves are compulsory in the lab. Capacity is limited to three people to a room, conference rooms are closed and meetings must still take place through calls or video-conferencing. Bernardi is comfortable with the university's balance of safety and flexibility.

At the University of Groningen in the Netherlands, researchers have been asked

to avoid working with hazardous chemicals when they can, to minimize the risk of spills that would require medical attention, says Jana Volaric, a synthetic organic chemist. But for her, the biggest impact is the diminished conference schedule. She had hoped to be networking in anticipation of being on the job market next year, and she says that meaningful interactions are harder to come by at online conferences. "This is the most disappointing part."

### 'Singing from the same song sheet'

Organic chemist Kirsty Anderson lost about four weeks of work when the University of Auckland in New Zealand closed down along with the rest of the country. It reopened

instructions through a shared document.

Winston Byblow, a neuroscientist at the University of Auckland who studies motor function after strokes, says the government and university are united in their safety messaging and pandemic response. "Everyone is singing from the same song sheet," he says. He's worried, however, that experiments with human participants might take a hit for a while. People are going to be wary about taking part in trials for many months, because of concerns about being in enclosed spaces with others, he says. "If recruitment rates drop because of uncertainty or fear, then it just means that it's going to take a lot longer to complete the research, and that comes at a cost."

### Charting their own path

Other universities are striking out on their own. Jorge Huete-Pérez is a molecular biologist and vice-president at the University of Central America in Managua, a private university with about 8,000 students. The institution formulated its own lockdown plan – reducing the number of people on campus by about 90% – independently of the Nicaraguan government, which put in place no measures to mitigate the virus's spread. The pandemic arrived late in Central America, and reports put the total number of cases in Nicaragua at less than 800. But Huete-Pérez thinks the real number is higher, more in step with or perhaps higher than the 1,000 cases reported in neighbouring Costa Rica, and the roughly 5,000 in Honduras. "We don't really know the real dimensions of the COVID situation," he says. When it is time reopen, he anticipates that he will look to independent medical organizations and the World Health Organization to inform the university's plan.

In the United States, as the presidential administration and the Centers for Disease Control and Prevention clashed over the plan for a post-pandemic return to work, the University of California, San Diego, drafted a detailed one of its own, including an ambitious screening and testing regime for staff and students. The plan will come into effect in the autumn, and relies on campus labs to process samples. The pilot phase, begun on 11 May, tested more than 1,000 students and ended last week.

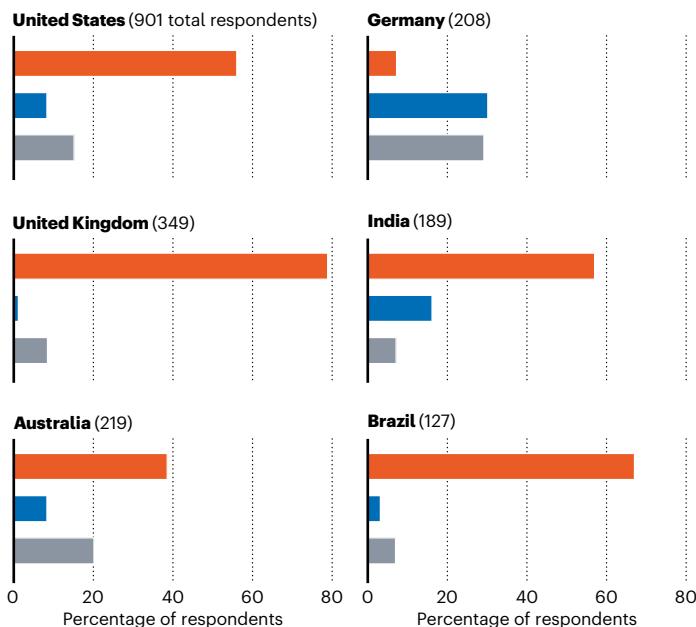
But that approach is something of an anomaly. In the United States, policies and priorities can differ drastically between states. "The situation is changing pretty rapidly and the stakes are very high," says Peter Schiffer, a physicist and vice-provost for research at Yale University in New Haven, Connecticut, who is a fellow at the AAU. A number of universities have adjusted their schedules or delayed students' return to campus until next year, but have yet to firm up plans for researchers.

The logistics are unparalleled says Tobin Smith, vice-president for policy at the AAU. "This is all new territory."

### GETTING BACK TO BUSINESS

In an online poll, more than 3,000 researchers shared their current working status. Less than 13% said that they had returned to the laboratory; others reported that they had been going to the lab throughout their lockdowns or could do most of their work from home. Among the countries with the most respondents, Germany had the highest proportion back in the lab.

■ On lockdown  
■ Already returned  
■ Still going in during pandemic





WHO director-general Tedros Adhanom Ghebreyesus.

## WHAT US EXIT FROM THE WHO MEANS FOR GLOBAL HEALTH

Experts foresee troubles ahead as Donald Trump ends US relationship with the agency.

By Amy Maxmen

**G**lobal health-policy experts reeled after an announcement by US President Donald Trump on 29 May that he would be “terminating” the country’s relationship with the World Health Organization (WHO).

The announcement follows a ramping up of blame levelled at the WHO by Trump. Ten days earlier, the president had sent a sharply worded letter to the agency’s director-general, Tedros Adhanom Ghebreyesus, threatening to permanently freeze funding and quit the organization if it didn’t “demonstrate independence from China” within 30 days. Now, WHO leaders and others are predicting repercussions that could range from a resurgence of polio and malaria to barriers in the flow of information on COVID-19. Scientific partnerships around the world could be damaged, and the United States might lose influence over global health initiatives. “This will hurt,” says Kelley Lee, a global health-policy researcher at Simon Fraser University in Burnaby, Canada.

Proposals for new US-led initiatives for pandemic preparedness abroad do little to quell researchers’ concerns. They say these efforts might complicate the world’s response

to COVID-19, and global health more generally. “It’s surreal to even be having this conversation, since it’s so hard to get one’s head around the massive implications,” says Rebecca Katz, director of the Center for Global Health Science and Security at Georgetown University in Washington DC.

The split is poorly timed, given the need for international coordination to contend with the coronavirus. “In this pandemic, people have said we’re building the plane while flying,” Katz says. “This proposal is like removing the windows while the plane is mid-air.”

### Balance due

Trump does not need congressional approval to withhold funds from the WHO, and it is unclear whether he will require it for withdrawal from the agency. Last year, the US government gave the WHO roughly US\$450 million. The country provides 27% of the WHO’s budget for polio eradication; 19% of its budget for tackling tuberculosis, HIV, malaria and vaccine-preventable diseases such as measles; and 23% of its budget for emergency health operations. If these initiatives shrink, researchers say, death and suffering will surge. David Heymann, an epidemiologist at the London School of Hygiene and

Tropical Medicine, says this will also amount to squandered investment for the United States, particularly for polio. Gains won through vaccination campaigns that cost hundreds of millions of dollars will be lost, he says.

### New initiatives

Trump says that the US government will continue to fund global health through aid groups and its own agencies. And proposed legislation suggests that the government might be considering alternative methods. Devex, an online platform focused on global development, reported that the US state department is circulating a proposal for a \$2.5-billion initiative that would oversee national and international pandemic responses. And late last month, a proposed bill for a ‘Global Health Security and Diplomacy Act of 2020’ was introduced to the Senate. The bill, obtained by *Nature*, would authorize \$3 billion for an international initiative to contain epidemics at home and abroad, to be overseen by a presidential appointee from the US state department.

Amanda Glassman, a senior fellow at the Center for Global Development, a think tank based in Washington DC, says that she and her colleagues welcome a US effort dedicated to fighting pandemics worldwide. But she doesn’t expect such parallel efforts to be very effective if they don’t work alongside a strong WHO. That’s because it takes years to build partnerships with countries, and the WHO works in some regions that the United States does not. Lee agrees. “You can’t just show up in Afghanistan and start vaccinating people.”

“The US relies on multilaterals to work in countries where diplomatic ties are almost non-existent,” says Suerie Moon, a global-health researcher at the Graduate Institute of International and Development Studies in Geneva, Switzerland.

Even in countries where the United States has long run programmes to tackle HIV, malaria and other health issues, the WHO still coordinates efforts. A rift between the WHO and researchers at US agencies could weaken long-standing collaborations. Furthermore, around 180 epidemiologists, health-policy specialists and other staff members at the WHO are from the United States, and dozens of Americans work at the organization as visiting scholars and interns.

The WHO will survive a US funding freeze over the short term, say researchers, because other donors will help. For example, Chinese President Xi Jinping has pledged \$2 billion to the coronavirus response.

Over time, the United States could lose its influence abroad. Ironically, that is exactly what the Trump administration is complaining about. “If the US pulls out and leaves a vacuum, it will be filled by other countries, like China,” says Lee. “You’ll see a self-fulfilling prophecy.”

# HYDROXYCHLOROQUINE SAFETY FEARS SPARK GLOBAL CONFUSION

Study suggesting the drug could be dangerous to people with coronavirus has slowed clinical trials.

By Heidi Ledford

**A** study suggesting that the malaria drug hydroxychloroquine could be harmful to people with severe COVID-19 is sowing confusion among researchers – and halting clinical trials that are crucial to finding out whether the medication is effective at treating people infected with the new coronavirus.

In light of the study, the World Health Organization (WHO) has paused enrolment in its trial of the drug as a treatment for COVID-19, as have regulators in the United Kingdom, France and Australia, where similar trials are under way. But the picture isn't clear-cut: on 28 May, 120 researchers signed a letter to *The Lancet*, which published the study (M. R. Mehra *et al.* *Lancet* <http://doi.org/ggwzsb>; 2020), highlighting concerns about the quality of the data and its analysis.

Hydroxychloroquine has been controversial during the pandemic, with politicians such as US President Donald Trump endorsing the medication as a COVID-19 treatment despite scant evidence that it works. Researchers have been eagerly awaiting results from clinical trials, but now they fear that the *Lancet* study, and the negative press coverage that followed, might dissuade people from joining trials.

"There's so much swirling around it, people won't want to enter those trials," says David Smith, an infectious-disease specialist at the University of California, San Diego. "In which case, it will be an open question that won't get answered."

Currently, only one treatment has shown promise against COVID-19 in a rigorous clinical trial. That drug, remdesivir, is in short supply. In the early days of the pandemic, many hoped that hydroxychloroquine, a relatively cheap and widely available drug, might offer some benefits. Early laboratory studies suggested that the compound, as well as a similar medicine named chloroquine, might interfere with replication of the coronavirus, but human trials have been inconclusive.

Researchers have been waiting for results from a number of randomized, controlled clinical trials testing hydroxychloroquine as a COVID-19 treatment. The *Lancet* study was not based on such trials, but was an observational study cataloguing the effects of

hydroxychloroquine in about 96,000 people hospitalized with COVID-19 around the world. It found that the drug had no benefit, and instead recorded a worryingly higher mortality rate among those who received it: 18%, compared with 9% in those who did not.

Data from observational studies can

**"We never called for halting clinical trials. In fact, we called for the opposite."**

be useful for giving a sense of how a drug performs outside the confines of a clinical trial, but they can also be prone to real-world bias: one concern, for example, is that physicians might be more likely to give a drug to patients who are sicker and more likely to die.

Nevertheless, the results have prompted action from regulators and clinical-trial sponsors. The WHO has paused enrolment in the hydroxychloroquine treatment group of its Solidarity trial, an effort to test potential COVID-19 therapies in 35 countries. The UK Medicines and Healthcare Regulatory Agency has similarly paused most of the country's

hydroxychloroquine trials for COVID-19, pending safety assessments. And in France, doctors have been forbidden to prescribe the drug to people with COVID-19 outside clinical trials.

## Trial difficulties

The *Lancet* study's lead author, Mandeep Mehra, a cardiologist at Brigham and Women's Hospital in Boston, Massachusetts, agrees with the WHO's decision to take a hard look at the safety data from the Solidarity trial. But he does not support ending clinical trials of hydroxychloroquine altogether. "We never called for halting clinical trials," he says. "In fact, we called for the opposite."

Enrolments in some clinical trials might resume, says cardiologist Joseph Cherian of the University of Cambridge, UK, after regulators have been assured that the data do not indicate that participants are being harmed by hydroxychloroquine. Within a day of the *Lancet* paper's publication, the UK RECOVERY trial, which has enrolled 10,000 people to test 6 potential COVID-19 treatments, consulted its safety board. The board analysed the trial's interim data and found no sign of harm to study participants from hydroxychloroquine. As a result, the trial can continue enrolment.

At the moment, hydroxychloroquine is nearly the only drug being evaluated in COVID-19 prevention studies, which aim to determine whether a drug could lessen the risk of coronavirus infection. Cherian worries that negative publicity surrounding hydroxychloroquine could affect enrolment in his prevention trial among healthy people, even though the studies suggesting possible harm were in people who were already ill with COVID-19.

In the middle of it all, the *Lancet* analysis itself has come under fire. Some researchers



Hydroxychloroquine has been controversial during the coronavirus pandemic.

have pointed to oddities in the data, such as smoking and obesity rates that were reported as nearly equal across all parts of the world studied in the analysis. And because the work relied on proprietary data gathered from medical centres, the raw data were not made available to other researchers.

Following criticisms, Mehra's team has tracked down some errors – the table reporting smoking and obesity rates, for example, erroneously listed modelled data rather than raw data. *The Lancet* will soon issue a correction, Mehra says, but the errors did not affect the conclusions. And the company that gathered the proprietary data, Surgisphere of Chicago, Illinois, says that it will accept proposals from other researchers to analyse its data.

But some remain sceptical about the study's conclusion that hydroxychloroquine is a safety risk to some people with COVID-19. Smith says that he and his collaborators have analysed the data that are available and concluded that people who received hydroxychloroquine in the study were often sicker in the first place than were those who did not. But Mehra says that his team accounted for those differences and still found a small difference in mortality. Furthermore, he says, his study merely highlights an association between the treatment and increased death, but because it is an observational study and not a clinical trial, it cannot exclude the possibility that other factors contributed to that difference in mortality.

sequence SARS-CoV-2 genomes.

So far, the team has sequenced samples from three-quarters of the state's roughly 1,700 cases<sup>3</sup>, thought to be the most comprehensive sequencing coverage in the world for an infectious-disease outbreak.

The data will be used to help identify the probable origin of new cases that arise as Victoria's social restrictions ease. Public-health officials will be able to make decisions to control outbreaks much faster than was possible in the first few weeks of the pandemic, when there were fewer genomes to compare, says Torsten Seemann, a bioinformatician at the Microbiological Diagnostic Unit Public Health Laboratory based at the Peter Doherty Institute for Infection and Immunity in Melbourne.

For example, the sequence data helped to resolve the true source of exposure for one health-care worker, proving that they contracted the virus at a social event and not from a patient in hospital. That information prevented the need for an investigation into a possible outbreak at the hospital, he says.

Genomic data will be particularly important when regional travel resumes. Borders of all states in Australia have been closed since March, but new infections are expected when they reopen. Over the past two months, viral genomes will have mutated just enough to tell whether they come from outside the state, researchers say.

Scientists in New Zealand have so far sequenced 25% of the country's 1,154 reported cases. They're aiming for more than 70% to get the most complete picture practically possible, says Joep de Ligt, lead bioinformatician at the Institute of Environmental Science and Research near Wellington, which is sequencing the country's cases. But the sequence data are already proving useful in responding to outbreaks, he says. Genomic data have identified links between cases that were missed by conventional contact tracing, and have untangled two clusters that were thought to be one.

But using genomics as part of the response to an outbreak has limitations, says de Ligt. With SARS-CoV-2, people who experience asymptomatic infections are unlikely to be tested, creating gaps in the genomic data, he says.

The use of genomic analysis to help contact tracing is also largely restricted to high-income countries, says Meru Sheel, an epidemiologist at the Australian National University in Canberra. She would like to see genomics considered as a tool for outbreak responses in resource-limited countries in the Asia-Pacific region, as it was in the Democratic Republic of the Congo, Sierra Leone and Guinea during the 2014–16 Ebola outbreak.

1. Zhang, Y.-Z. & Holmes, E. C. *Cell* **181**, 223–227 (2020).

2. Stevens, E. et al. *Front. Microbiol.* **8**, 808 (2017).

3. Seemann, T. et al. Preprint at MedRxiv <https://doi.org/10.1101/2020.05.12.20099929> (2020).

## GENOMICS USED TO HELP AVOID A SECOND CORONAVIRUS WAVE

Scientists in New Zealand and elsewhere are using sequence data to track infections as lockdowns ease.

By Clare Watson

**A**s many countries emerge from lockdowns, researchers are poised to use genome sequencing to avoid an expected second wave of COVID-19 infections.

Since the first whole-genome sequence of the new coronavirus, SARS-CoV-2, was shared online on 11 January, scientists have sequenced and shared some 32,000 viral genomes from around the world. The data have helped researchers to trace the origin of their countries' COVID-19 outbreaks and pinpoint when community transmission occurred<sup>1</sup>.

Now, countries that have successfully suppressed infections are entering the next phase of the COVID-19 pandemic – where there's a risk of new cases appearing as social restrictions ease. Researchers say that genomics will be crucial to quickly track and control these outbreaks. Studies already show that disease outbreaks tend to be shorter and smaller when genomics is used to help contact tracing<sup>2</sup>.

"When there are few cases, genomics can very quickly tell you what you're dealing with and therefore guide precision interventions," says Gytis Dudas, a consulting bioinformatician at the Gothenburg Global Biodiversity Centre in Sweden.

Several places are particularly well placed to do that because they invested in genome

sequencing early in the pandemic and have a relatively small numbers of cases. Researchers in New Zealand, and at least one state in Australia, decided that they would aim to sequence most coronavirus genomes in their region.

As SARS-CoV-2 spread around the world and viruses circulating in different regions gradually evolved, distinct lineages began to form. By comparing sequences, researchers can quickly rule out possible lines of transmis-

**"When there are few cases, genomics can very quickly tell you what you're dealing with."**

sion if two sequences don't match, and can link together cases that do.

Scientists in the United Kingdom, the United States and other countries are also sequencing SARS-CoV-2 from a large proportion of cases there, but because their epidemics are still ongoing and case numbers are high, genomics is being used to monitor spread and help identify the source of some cases where contact tracing fails.

Before the first reported case arrived in Australia from Wuhan, China, in early January, researchers at a laboratory in Melbourne, Victoria, that usually investigates outbreaks of food-borne illness started preparing to

# SURVIVAL OF THE LITTLEST



Babies born before 28 weeks of gestation are surviving into adulthood at higher rates than ever. What are the consequences, in later life, of being born so early? **By Amber Dance**

**T**hey told Marcelle Girard her baby was dead.

Back in 1992, Girard, a dentist in Gatineau, Canada, was 26 weeks pregnant and on her honeymoon in the Dominican Republic.

When she started bleeding, physicians at the local clinic assumed the baby had died. But Girard and her husband felt a kick. Only then did the doctors check for a fetal heartbeat and realize the baby was alive.

The couple was medically evacuated by air to Montreal, Canada, then taken to the Sainte-Justine University Hospital Center. Five hours later, Camille Girard-Bock was born,

weighing just 920 grams (2 pounds).

Babies born so early are fragile and underdeveloped. Their lungs are particularly delicate: the organs lack the slippery substance, called surfactant, that prevents the airways from collapsing upon exhalation. Fortunately for Girard and her family, Sainte-Justine had recently started giving surfactant, a new treatment at the time, to premature babies.

After three months of intensive care, Girard took her baby home.

Today, Camille Girard-Bock is 27 years old and studying for a PhD in biomedical sciences at the University of Montreal. Working with

researchers at Sainte-Justine, she's addressing the long-term consequences of being born extremely premature – defined, variously, as less than 25–28 weeks in gestational age.

Families often assume they will have grasped the major issues arising from a premature birth once the child reaches school age, by which time any neurodevelopmental problems will have appeared, Girard-Bock says. But that's not necessarily the case. Her PhD advisers have found that young adults of this population exhibit risk factors for cardiovascular disease – and it may be that more chronic health conditions will show up with time.

Girard-Bock doesn't let these risks preoccupy



**Scientists are watching out for the health of premature babies as they reach adulthood and beyond.**

her. "As a survivor of preterm birth, you beat so many odds," she says. "I guess I have some kind of sense that I'm going to beat those odds also."

She and other against-the-odds babies are part of a population which is larger now than at any time in history: young adults who are survivors of extreme prematurity. For the first time, researchers can start to understand the long-term consequences of being born so early. Results are pouring out of cohort studies that have been tracking kids since birth, providing data on possible long-term outcomes; other studies are trialling ways to minimize the consequences for health.

These data can help parents make difficult

decisions about whether to keep fighting for a baby's survival. Although many extremely premature infants grow up to lead healthy lives, disability is still a major concern, particularly cognitive deficits and cerebral palsy.

Researchers are working on novel interventions to boost survival and reduce disability in extremely premature newborns. Several compounds aimed at improving lung, brain and eye function are in clinical trials, and researchers are exploring parent-support programmes, too.

Researchers are also investigating ways to help adults who were born extremely prematurely to cope with some of the long-term

health impacts they might face: trialling exercise regimes to minimize the newly identified risk of cardiovascular disease, for example.

"We are really at the stage of seeing this cohort becoming older," says neonatologist Jeanie Cheong at the Royal Women's Hospital in Melbourne, Australia. Cheong is the director of the Victorian Infant Collaborative Study (VICS), which has been following survivors for four decades. "This is an exciting time for us to really make a difference to their health."

The late twentieth century brought huge changes to neonatal medicine. Lex Doyle, a paediatrician and previous director of VICS, recalls that when he started caring for preterm

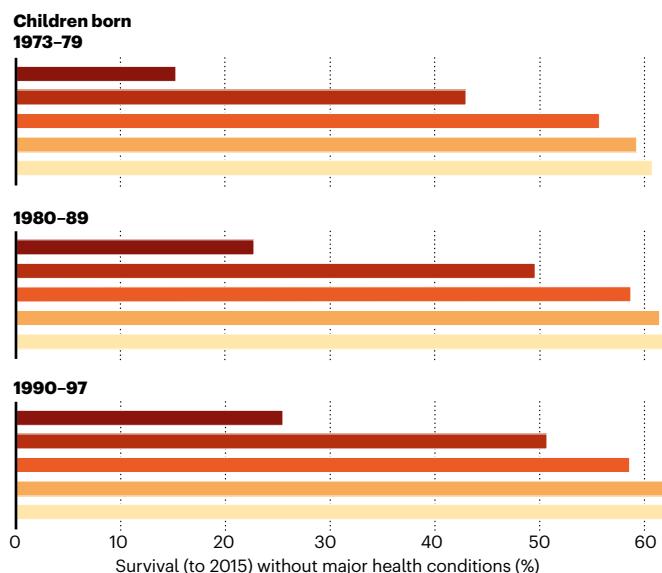
# Feature

## THE EFFECTS OF BEING EARLY

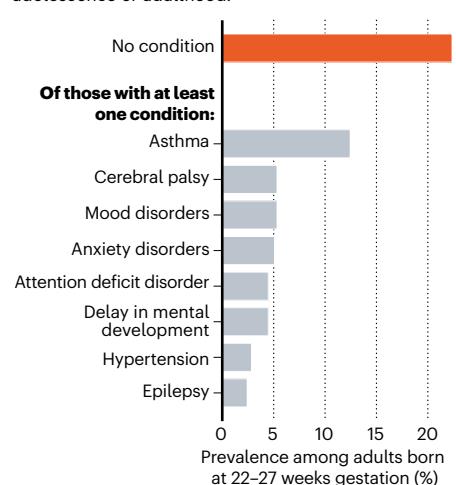
Data from a large Swedish cohort show the proportion of babies that live beyond age 18 without major health problems. Survival rates increased from the 1970s to the 1990s, especially for babies born extremely prematurely.

Gestational age at birth (weeks)

- 22–27
- 28–33
- 34–36
- 37–38
- 39–41



More than three-quarters of the babies born extremely preterm — at or before 27 weeks — had at least one chronic health condition when they reached adolescence or adulthood.



infants in 1975, very few survived if they were born at under 1,000 grams — a birthweight that corresponds to about 28 weeks' gestation. The introduction of ventilators, in the 1970s in Australia, helped, but also caused lung injuries, says Doyle, now associate director of research at the Royal Women's Hospital. In the following decades, doctors began to give corticosteroids to mothers due to deliver early, to help mature the baby's lungs just before birth. But the biggest difference to survival came in the early 1990s, with surfactant treatment.

"I remember when it arrived," says Anne Monique Nuyt, a neonatologist at Sainte-Justine and one of Girard-Bock's advisers. "It was a miracle." Risk of death for premature infants dropped to 60–73% of what it was before<sup>1,2</sup>.

Today, many hospitals regularly treat, and often save, babies born as early as 22–24 weeks. Survival rates vary depending on location and the kinds of interventions a hospital is able to provide. In the United Kingdom, for example, among babies who are alive at birth and receiving care, 35% born at 22 weeks survive, 38% at 23 weeks, and 60% at 24 weeks<sup>3</sup>.

For babies who survive, the earlier they are born, the higher the risk of complications or ongoing disability (see 'The effects of being early'). There is a long list of potential problems — including asthma, anxiety, autism spectrum disorder, cerebral palsy, epilepsy and cognitive impairment — and about one-third of children born extremely prematurely have one condition on the list, says Mike O'Shea, a neonatologist at the University of North Carolina School of Medicine in Chapel Hill, who co-runs a study tracking children born between 2002 and 2004. In this cohort, another one-third have multiple disabilities, he says, and the rest have none.

"Preterm birth should be thought of as a chronic condition that requires long-term follow-up," says Casey Crump, a family physician and epidemiologist at the Icahn School

of Medicine at Mount Sinai in New York, who notes that when these babies become older children or adults, they don't usually get special medical attention. "Doctors are not used to seeing them, but they increasingly will."

### Outlooks for earlies

What should doctors expect? For a report in the *Journal of the American Medical Association* last year<sup>4</sup>, Crump and his colleagues scraped data from the Swedish birth registry. They looked at more than 2.5 million people born from 1973 to 1997, and checked their records for health issues up until the end of 2015.

Of the 5,391 people born extremely preterm, 78% had at least one condition that manifested in adolescence or early adulthood, such as a psychiatric disorder, compared with 37% of those born full-term. When the researchers looked at predictors of early mortality, such as heart disease, 68% of people born extremely prematurely had at least one such predictor, compared with 18% for full-term births — although these data include people born before surfactant and corticosteroid use were widespread, so it's unclear if these data reflect outcomes for babies born today. Researchers have found similar trends in a UK cohort study of extremely premature births. In results published earlier this year<sup>5</sup>, the EPICure study team, led by neonatologist Neil Marlow at University College London, found that 60% of 19-year-olds who were extremely premature were impaired in at least one neuropsychological area, often cognition.

Such disabilities can impact education as well as quality of life. Craig Garfield, a paediatrician at the Northwestern University Feinberg School of Medicine and the Lurie Children's Hospital of Chicago, Illinois, addressed a basic question about the first formal year of schooling in the United States: "Is your kid ready for kindergarten, or not?"

To answer it, Garfield and his colleagues

analysed standardized test scores and teacher assessments on children born in Florida between 1992 and 2002. Of those born at 23 or 24 weeks, 65% were considered ready to start kindergarten at the standard age, 5–6 years old, with the age adjusted to take into account their earlier birth. In comparison, 85.3% of children born full term were kindergarten-ready<sup>6</sup>.

Despite their tricky start, by the time they reach adolescence, many people born prematurely have a positive outlook. In a 2006 paper<sup>7</sup>, researchers studying individuals born weighing 1,000 grams or less compared these young adults' perceptions of their own quality of life with those of peers of normal birth-weight — and, to their surprise, found that the scores were comparable. Conversely, a 2018 study<sup>8</sup> found that children born at less than 28 weeks did report having a significantly lower quality of life. The children, who did not have major disabilities, scored themselves 6 points lower, out of 100, than a reference population.

As Marlow spent time with his participants and their families, his worries about severe neurological issues diminished. Even when such issues are present, they don't greatly limit most children and young adults. "They want to know that they are going to live a long life, a happy life," he says. Most are on track to do so. "The truth is, if you survive at 22 weeks, the majority of survivors do not have a severe, life-limiting disability."

### Breathless

But scientists have only just begun to follow people born extremely prematurely into adulthood and then middle age and beyond, where health issues may yet lurk. "I'd like scientists to focus on improving the long-term outcomes as much as the short-term outcomes," says Tala Alsadik, a 16-year-old high-school student in Jeddah, Saudi Arabia.

When Alsadik's mother was 25 weeks

pregnant and her waters broke, doctors went so far as to hand funeral paperwork to the family before consenting to perform a caesarean section. As a newborn, Alsadik spent three months in the neonatal-intensive-care unit (NICU) with kidney failure, sepsis and respiratory distress.

The complications didn't end when she went home. The consequences of her prematurity are on display every time she speaks, her voice high and breathy because the ventilator she was put on damaged her vocal cords. When she was 15, her navel unexpectedly began leaking yellow discharge, and she required surgery. It turned out to be caused by materials leftover from when she received nutrients through a navel tube.

That certainly wasn't something her physicians knew to check for. In fact, doctors don't often ask if an adolescent or adult patient was born prematurely – but doing so can be revealing.

Charlotte Bolton is a respiratory physician at the University of Nottingham, UK, where she specializes in patients with chronic obstructive pulmonary disease (COPD). People coming into her practice tend to be in their 40s or older, often current or former smokers. But in around 2008, she began to notice a new type of patient being referred to her owing to breathlessness and COPD-like symptoms: 20-something non-smokers.

Quizzing them, Bolton discovered that many had been born before 32 weeks. For more insight, she got in touch with Marlow, who had also become concerned about lung function as the EPICure participants aged. Alterations in lung function are a key predictor of cardiovascular disease, the leading cause of death around the world. Clinicians already knew that after extremely premature birth, the lungs often don't grow to full size. Ventilators, high oxygen levels, inflammation and infection can further damage the immature lungs, leading to low lung function and long-term breathing problems, as Bolton, Marlow and their colleagues showed in a study of 11-year-olds<sup>9</sup>.

VICS research backs up the cardiovascular concerns: researchers have observed diminished airflow in 8-year-olds, worsening as they aged<sup>10</sup>, as well as high blood pressure in young adults<sup>11</sup>. "We really haven't found the reason yet," says Cheong. "That opens up a whole new research area."

At Sainte-Justine, researchers have also noticed that young adults who were born at 28 weeks or less are at nearly three times the usual risk of having high blood pressure<sup>12</sup>. The researchers figured they would try medications to control it. But their patient advisory board members had other ideas – they wanted to try lifestyle interventions first.

The scientists were pessimistic as they began a pilot study of a 14-week exercise programme. They thought that the cardiovascular risk

factors would be unchangeable. Preliminary results indicate that they were wrong; the young adults are improving with exercise.

Girard-Bock says the data motivate her to eat healthily and stay active. "I've been given the chance to stay alive," she says. "I need to be careful."

### From the start

For babies born prematurely, the first weeks and months of life are still the most treacherous. Dozens of clinical trials are in progress for prematurity and associated complications, some testing different nutritional formulas or improving parental support, and others tar-

they leave the hospital. For parents, it can be nerve-racking to go it alone after depending on a team of specialists for months, and lack of parental confidence has been linked to parental depression and difficulties with behaviour and social development in their growing children.

At Women & Infants Hospital of Rhode Island in Providence, Betty Vohr is director of the Neonatal Follow-Up Program. There, families are placed in private rooms, instead of sharing a large bay as happens in many NICUs. Once they are ready to leave, a programme called Transition Home Plus helps them to prepare and provides assistance such as regular check-ins by phone and in person in the first few days at home, and a 24/7 helpline. For mothers with postnatal depression, the hospital offers care from psychologists and specialist nurses.

The results have been significant, says Vohr. The single-family rooms resulted in higher milk production by mothers: 30% more at four weeks than for families in more open spaces. At 2 years old, children from the single-family rooms scored higher on cognitive and language tests<sup>16</sup>. After Transition Home Plus began, babies discharged from the NICU had lower health-care costs and fewer hospital visits – issues that are of great concern for premature infants<sup>17</sup>. Other NICUs are developing similar programmes, Vohr says.

With these types of novel intervention, and the long-term data that continue to pour out of studies, doctors can make better predictions than ever before about how extremely premature infants will fare. Although these individuals face complications, many will thrive.

Alsadik, for one, intends to be a success story. Despite her difficult start in life, she does well academically, and plans to become a neonatologist. "I also want to improve the long-term outcomes of premature birth for other people."

**Amber Dance** is a freelance journalist in Los Angeles, California.

1. Soll, R. & Özek, E. *Cochrane Database Syst. Rev.* <https://doi.org/10.1002/14651858.CD000511> (1997).
2. Soll, R. *Cochrane Database Syst. Rev.* <https://doi.org/10.1002/14651858.CD001149> (1998).
3. British Association of Perinatal Medicine. *Perinatal Management of Extreme Preterm Birth Before 27 Weeks of Gestation: A BAPM Framework for Practice* (BAPM, 2019).
4. Crump, C., Winkleby, M. A., Sundquist, J. & Sundquist, K. *J. Am. Med. Assoc.* **322**, 1580–1588 (2019).
5. O'Reilly, H., Johnson, S., Ni, Y., Wolke, D. & Marlow, N. *Pediatrics* **145**, e20192087 (2020).
6. Garfield, C. F. et al. *JAMA Pediatr.* **171**, 764–770 (2017).
7. Saigal, S. et al. *Pediatrics* **118**, 1140–1148 (2006).
8. Gire, C. et al. *Arch. Dis. Child.* **104**, 333–339 (2018).
9. Bolton, C. E. et al. *J. Pediatr.* **161**, 595–601.e2 (2012).
10. Doyle, L. W. et al. *Thorax* **72**, 712–719 (2017).
11. Halkarwal, A. et al. *Hypertension* **75**, 211–217 (2019).
12. Flahault, A. et al. *Hypertension* **75**, 796–805 (2020).
13. Ley, D. et al. *J. Pediatr.* **206**, 56–65.e8 (2019).
14. Stalh, A. et al. *Lancet* **394**, 1551–1559 (2019).
15. Härkin, P. et al. *J. Pediatr.* **177**, 72–77.e2 (2016).
16. Vohr, B. et al. *J. Pediatr.* **185**, 42–48.e1 (2017).
17. Liu, Y. et al. *J. Pediatr.* **200**, 91–97.e3 (2018).

## THIS IS AN EXCITING TIME FOR US TO REALLY MAKE A DIFFERENCE TO THEIR HEALTH.

getting specific issues that lead to disability later on: underdeveloped lungs, brain bleeds and altered eye development.

For instance, researchers hoping to protect babies' lungs gave a growth factor called IGF-1 – which the fetus usually gets from its mother during the first two trimesters of pregnancy – to premature babies in a phase II clinical trial reported<sup>13</sup> in 2016. Rates of a chronic lung condition that often affects premature babies halved, and babies were somewhat less likely to have a severe brain haemorrhage in their earliest months.

Another concern is visual impairment. Retina development halts prematurely when babies born early begin breathing oxygen. Later it restarts, but preterm babies might then make too much of a growth factor called VEGF, causing over-proliferation of blood vessels in the eye, a disorder known as retinopathy. In a phase III trial announced in 2018, researchers successfully treated 80% of these retinopathy cases with a VEGF-blocking drug called ranibizumab<sup>14</sup>, and in 2019 the drug was approved in the European Union for use in premature babies.

Some common drugs might also be of use: paracetamol (acetaminophen), for example, lowers levels of biomolecules called prostaglandins, and this seems to encourage a key fetal vein in the lungs to close, preventing fluid from entering the lungs<sup>15</sup>.

But among the most promising treatment programmes, some neonatologists say, are social interventions to help families after

# Books & arts



Many Somali refugees live in the crowded Dadaab camp for displaced people in Kenya, after fleeing droughts in their country.

LYNSEY ADDARIO / GETTY IMAGES REPORTAGE

## Migration is normal – accept it

Pull down the walls, argues a book on the movement of people, animals and plants. **By Emma Marris**

**W**hen the coronavirus pandemic led to stay-at-home orders around the globe, we all felt, at once, how central to human life movement is. Prevented from leaving our neighbourhoods, we became isolated, nervous, claustrophobic. Were we experiencing some *Zugunruhe*, the 'migratory restlessness' that seizes birds when it is time to take flight?

This near-universal disruption of how our species operates is another piece of evidence to add to the studies and anecdotes collected by journalist Sonia Shah to support her claim that migration is not aberrant, but is "an unexceptional ongoing reality". *The Next Great Migration* anticipates movements – human and non-human – in response to climate change as just the latest chapter in a story as old as life

itself. Such migrations, she posits, are natural, common and largely harmless.

Shah convincingly argues that politicians against immigration distort and misuse data to create unnecessary and cruel barriers. She tells gut-wrenching stories of struggling families on the move, and presents evidence that migrants are generally healthier and less apt to commit violent crimes than are the residents of the country they move to. And studies show that immigrants benefit host economies.

Turning to plants and animals, she takes biologists to task for abusing data to make introduced species look worse than they are. She charges that renowned ecologist Charles Elton "cherry-picked" case studies of the most disruptive introduced species (such as the sea lamprey (*Petromyzon marinus*), which originated in the Atlantic Ocean and devastated trout populations in the Great Lakes). She levels that later biologists exaggerated the economic impact of non-natives by including the costs of

removing them in calculations. Shah concludes that invasion biologists' predictions of "ecological Armageddon" have failed to transpire.

As a writer on ecology and conservation (for example, my 2011 book *Rambunctious Garden*), my expertise lies with the non-human migrations that Shah covers. I found it odd that she tends to group together many kinds of biological movement that specialists think of as quite different. Annual migrations – such as those of bar-headed geese (*Anser indicus*), which fly over the Himalayas on yearly trips from Mongolia to India and back again – as well as seed dispersal and animals roaming to find new territories and mates, aren't generally seen as comparable to human-mediated translocations across oceans, such as the deliberate introduction of Asian turtles in Hawaii to be bred for food. The former are considered natural, valuable processes; the latter are deemed unnatural, generally undesirable and potentially dangerous. Yet, Shah often ignores the human role.

Perhaps that's part of her point. The way we judge species' movements is a product of our culture, after all, and the borders we draw might be as arbitrary as those between our nation states. People are simply primates with iPhones; why should the species we carry be treated differently from those that move without us? If monkeys could arrive in the New World 30 million years after the Atlantic Ocean formed; if sweet potatoes could raft to Polynesia on their own; if seeds from Hawaii's koa tree could find purchase on the island of Réunion, half a world away, then even trans-oceanic range shifts are perfectly natural. How else could remote islands have developed any 'native' flora and fauna at all?

Most invasion biologists counter that philosophical discussions of 'naturalness' don't get us very far. Their concern is the effect on the travellers' new homes. The fact that many organisms move about widely of their own accord is not, in itself, evidence that introduced species can't be a problem. Most non-natives don't establish a population, or they have no unwanted effects – as Elton wrote, "There are enormously more invasions that never happen, or fail quite soon or even after a good many years." But a few do take hold. Some – mostly on islands or in lakes – threaten local inhabitants with extinction.

Such extinctions have surely occurred after long-distance range shifts that people had nothing to do with. The sudden arrival of rodents in South America 41 million years ago, for example, probably had far-reaching effects on the ecosystems there. Indeed, millions of years of planet-wide movements – including,



**Bar-headed geese, which migrate from Tibet and central Asia each winter, near Jammu, India.**

in the past few millennia, journeys of animals and plants as stowaways or as travelling companions of humans – have shaped the ecosystems we are trying to protect.

But the pace of radical range shifts in today's world is immeasurably higher. Before humans got to Hawaii, 30 new species made it there every million years. In the past 2 centuries, the islands have seen about 20 new species show up each year. And there is an ethical argument to be made that, when humans move species, we have a collective responsibility to mitigate the consequences. Cats and foxes that were introduced to Australia after European settlement in 1788 have helped to wipe out 22 native mammal species, such as the desert bandicoot (*Perameles eremicana*) and lesser stick-nest rat (*Leporillus apicalis*). Many introduced species will not cause significant problems. But if we wish to prevent extinctions, then a subset of these introduced species must be managed in some way. Shah does not parse these subtleties with the degree of nuance that specialists might want.

I do agree with her, however, that it is absolutely worth comparing our attitudes

towards human and non-human migrants, especially as climate change increases all such movements – our temporary lockdown notwithstanding. As Earth heats, trees climb mountains, butterflies flutter north and birds nest in new places. Humans, too, are relocating in a flow "from south to north along the gradient of our warming planet", Shah writes. She predicts that the biggest and most lasting human movements will be those sparked by persistent droughts. And climate migration won't be a straightforward, predictable human tide. Each story of migration is unique, she says, motivated by an interplay of idiosyncratic social, economic and personal factors.

The shifting climate means that we must allow – and assist – people, plants and animals who are forced to relocate to survive. Instead of hardening our borders and fighting to prevent change at all costs, we should help migrants to move in a safe and orderly way. Context will tell us when specific moves – such as animal predators to remote islands – are unwise. But we must face the inevitable: our social, political and ecological world is changing substantially. The altered communities that result won't just be different, they'll often be better adapted to thrive in our warming world. As Shah writes, "We can turn migration from a crisis into its opposite: the solution."

**The Next Great Migration: The Beauty and Terror of Life on the Move**  
Sonia Shah  
Bloomsbury (2020)

**Emma Marris** is an environmental writer who lives in Klamath Falls, Oregon.  
e-mail: e.marris@gmail.com

## Heather Houser Chronicling the infowhelm

**A deluge of contested data on crises is prompting writers and artists to consider the limits of knowledge.** Heather Houser, co-director of Planet Texas 2050, a climate-resilience project at the University of Texas at Austin, talks about her book *Infowhelm* (Columbia University Press, 2020).

### What do you mean by 'infowhelm'?

What I was capturing with that term is a deluge in which knowledge can be changing a lot, and then people in power can be poking holes in it. This creates a very difficult 'information situation' for us to process and understand.

### Why did you write this book?

I wanted to give an account of how art — literature and visual culture — is managing in the face of climate change and information overload. Probably the best-known example I talk about is Barbara Kingsolver's 2012 novel *Flight Behavior*; it's a narrative about climate change set in Appalachia in the United States. The main character uses scientific methods of data collection and observation and analysis. I'm interested in how art can entangle scientific data with aspects of lived experience, from emotion to doubt and uncertainty. I'm fascinated with the ways scientific data can conflict with one's social, economic and cultural position in the world.

### Can too much information harm knowledge?

Definitely. I begin each section with some poetry by Juliana Spahr, which really captures that sense of how too much information can be a hindrance and not a conduit to knowledge. Psychologists describe the numbing effects that it can cause, a sense of despair or apathy, or a sense of individual insignificance. They offer advice on moderating your intake of information, because otherwise, some of the emotional effects can obliterate what you do with it.

### What struck you about how scientists and artists present data?

The emotional impact of doing work every day on things such as extinction or climate change can be profoundly devastating



Studying climate change in regions such as Greenland can be emotionally devastating.

for scientists, and, in many cases, they have to detach that emotion from the official presentation of the research in journal articles or for conferences. I think the arts are a great way to allow all these other aspects of understanding and processing, as a person, to come into the narrative.

Even when it's recognized that there's some uncertainty behind the data, there can be a tendency for scientists to present them as more stable and objective and authoritative than some of the artists might do. I think what the artists bring to data is some deference, but also some awareness of limitations — all that the data can't communicate or help us understand.

### Scientists often display data pictorially. How does this relate to art?

I was really curious about data visualization, and what purposes it serves culturally, beyond scientific-research communication. Things like charts and graphs and different kinds of mapping strategies to display information that we're seeing in all the COVID-19 coverage today. I was interested in how artists adopt those as well. I write about Michael Crichton's 2004 novel *State of Fear*, which incorporates scientific charts and graphs.

### How have artists used visualization technologies, such as Google Earth?

A section of the book is on aerial imagery. I examine one environmental-watchdog group, and artists such as Laura Kurgan, who use satellite imagery. Kurgan doesn't change it very much but, through strategies of juxtaposition and different levels of zooming in, she has us think about what's hidden and what's shown, because of the technology's history and how it gets used.

### Did the artists' responses veer towards the apocalyptic?

I talk about 15–20 literary and visual artworks in this book, and despair is not their main tone. That's true even for something that might look as if it's heading towards the apocalyptic, such as Maya Lin's digital memorial, *What is Missing?*, begun in 2010, which is dedicated to extinction. It includes extinct species, but also extinct experiences, like the sounds of birds or starry nights. But because it is an open-source memorial that anyone can contribute to, it opens the way to imagining other futures.

### Interview by Gaia Vince

This interview has been edited for length and clarity.

FLORIAN LEDOUX

# John Horton Conway (1937–2020)

### Playful master of games who transformed mathematics.

John Horton Conway was one of the most versatile mathematicians of the past century, who made influential contributions to group theory, analysis, topology, number theory, geometry, algebra and combinatorial game theory. His deep yet accessible work, larger-than-life personality, quirky sense of humour and ability to talk about mathematics with any and all who would listen made him the centre of attention and a pop icon everywhere he went, among mathematicians and amateurs alike. His lectures about numbers, games, magic, knots, rainbows, tilings, free will and more captured the public's imagination.

Conway, who died at the age of 82 from complications related to COVID-19, was a lover of games of all kinds. He spent hours in the common rooms of the University of Cambridge, UK, and Princeton University in New Jersey playing backgammon, Go and other diversions, some of his own creation. Several of Conway's most celebrated contributions were made while he was thinking about games and their strategies. Perhaps his greatest discovery was a surprising correspondence between numbers and games that led him to a truly gigantic system, the surreal numbers, which stunned the mathematics community. It contained not only the positive and negative real numbers that we all know, but also new infinitely large numbers, infinitesimally small ones, and all sorts of new numbers in between.

Conway's work on surreal numbers emerged from the influential research project and book *Winning Ways for your Mathematical Plays* (1982), a compendium of information on the theory of games, written with Elwyn Berlekamp and Richard Guy. This fascination with games also led Conway to develop the Game of Life, a cellular automaton in which the pattern of live or dead cells in a two-dimensional grid evolves according to a set of rules for the 'birth' and 'death' of each cell, based on the status of its nearest neighbours. The simplicity and accessibility of this game was popularized in 1970 by *Scientific American* columnist Martin Gardner. By the mid-1970s, it was estimated that one-quarter of the world's computers were running Conway's Game of Life as their screensaver.

Conway, who was the John von Neumann professor of mathematics at Princeton University before his retirement in 2013, was born in Liverpool, UK, in 1937. His father made his living playing cards, and later worked as a



chemistry laboratory technician at a local high school attended by George Harrison and Paul McCartney. Conway, like his mother, was an avid reader. He showed early interests in mathematics; by the age of 11, he wanted to be a mathematician at Cambridge. He received his PhD from the University of Cambridge in 1964 under the adviseryship of Harold Davenport, was subsequently hired at Cambridge as a lecturer, and became professor in 1983. In 1987, he moved to Princeton.

**"His lectures about numbers, games, magic, knots, rainbows and more captured the public's imagination."**

Conway first attained fame in 1968 for determining all 8,315,553,613,086,720,000 symmetries of the Leech lattice – a remarkably regular arrangement of points in 24-dimensional space discovered by John Leech in 1967. This led to his discovery of the Conway simple groups, which were fundamental in the classification of finite simple groups – one of the capstone achievements of twentieth-century mathematics.

Conway had a primary role in researching and assembling the iconic symmetry book *ATLAS of Finite Groups* (1985). His deep knowledge of symmetries led him to propose, with his *ATLAS* co-author Simon Norton, the

Monstrous Moonshine conjectures. These, for the first time, seriously connected finite symmetry groups to analysis – and thus discrete maths to non-discrete maths. Today, the Moonshine conjectures play a key part in physics – including in the understanding of black holes in string theory – inspiring a wave of further such discoveries connecting algebra, analysis and beyond.

Conway's discovery of a new knot invariant – used to tell different knots apart – called the Conway polynomial became an important topic of research in topology. In geometry, he made key discoveries in the study of symmetries, sphere packings, lattices, polyhedra and tilings, including properties of quasi-periodic tilings as developed by Roger Penrose.

In algebra, Conway discovered another important system of numbers, the icosians, with his long-time collaborator Neil Sloane. In number theory, Conway showed that every whole number is the sum of at most 37 fifth powers. He also developed the 15-theorem (with his student William Schneeberger) and the 290-conjecture; these were vast generalizations of the four-squares theorem, proved by eighteenth-century mathematician Joseph-Louis Lagrange, which states that every positive whole number is the sum of four square numbers (for example, 21 is the sum of 16, 4, 1 and 0).

Conway was a memorable teacher and speaker, and the many tricks he performed to illustrate mathematical concepts included: stating immediately the day of the week for any date in history, twirling a hanger with a penny balanced on its inside edge, contorting his tongue into a variety of shapes, balancing objects on his chin, and delivering entire lectures in which every word he said had only one syllable.

He loved to talk about mathematics and games, as well as history, etymology and philosophy. His contributions to culture, through his work and outreach, will have a lasting impact. For the remarkable profundity of his mathematical discoveries – and the playful and generous way in which he shared these with others – he will be sorely missed.

**Manjul Bhargava** is the R. Brandon Fradd professor of mathematics at Princeton University, New Jersey, USA. He was a first-year graduate advisee, and later colleague, of John H. Conway at Princeton.  
e-mail: bhargava@math.princeton.edu

# Comment



SEONGJOOON CHO/BLOOMBERG VIA GETTY

Passengers on an underground train in Seoul. South Korea used contact tracing to great effect early in the pandemic.

## Ethical guidelines for COVID-19 tracing apps

Jessica Morley, Josh Cowls, Mariarosaria Taddeo & Luciano Floridi

**Protect privacy, equality and fairness in digital contact tracing with these key questions.**

**T**echnologies to rapidly alert people when they have been in contact with someone carrying the coronavirus SARS-CoV-2 are part of a strategy to bring the pandemic under control. Currently, at least 47 contact-tracing apps are available globally ([see go.nature.com/2zc1qhk](http://go.nature.com/2zc1qhk)). They are already in use in Australia, South Korea and Singapore, for instance. And many other governments are testing or considering them.

Herewe set out 16 questions to assess whether – and to what extent – a contact-tracing app is ethically justifiable. These questions could assist governments, public-health agencies and providers to develop ethical apps – they have already informed developments in France,

Italy and the United Kingdom. They will also help watchdogs and others to scrutinize such technologies.

What do COVID-19 contact-tracing apps do? Running on a mobile phone, they inform people that they have spent time near someone with the virus. The contacts should then respond according to local rules, for example by isolating themselves. Prompt alerts are key because the incubation time of the virus is up to two weeks<sup>1–4</sup>.

These digital interventions come at a price. Collecting sensitive personal data potentially threatens privacy, equality and fairness. Even if COVID-19 apps are temporary, rapidly rolling out tracing technologies runs the risk of creating permanent, vulnerable records of people's health, movements and social interactions, over which they have little control.

More ethical oversight is essential. So far, such concerns have focused on rights to privacy ([see go.nature.com/3e7jntx](http://go.nature.com/3e7jntx)). Some governments have pledged to protect data privacy ([see go.nature.com/3grwfe8](http://go.nature.com/3grwfe8)). Apple and Google are developing a common interface to support apps that do not require central data storage ([see \*Nature\* <http://doi.org/dwc6>](http://Nature http://doi.org/dwc6)).

2020). However, other ethical and social considerations must not be cast aside in the rush to quell the pandemic.

For instance, contact-tracing apps should be available and accessible to anyone, irrespective of the technology needed or their level of digital literacy. Yet many apps work only with certain phones. Australia, for example, has no plans to make its app work with phones that use software older than Apple's iOS 10 or Android 6.0. In the United Kingdom, around one-fifth of adults do not use a smartphone, and so might be excluded from a digital contact-tracing programme.

Rolling out an app without considering its wide ethical and social implications can be dangerous, costly and useless. For example, Bluetooth signals that show the proximity of two individuals' mobile phones are not a certain indicator of infection risk – two people might be in the same space but physically separated, for example, by a wall. A high level of false positives from such an app (for instance, as a result of self-reporting) could lead to unjustified panic. And minimal protections against false negatives (people not using the app to report that they are unwell) could spur a false sense of safety in others and increase the risk of infection.

The public might reject apps that breach principles of privacy, equality and fairness. This would frustrate the efforts and waste the resources being invested in developing and deploying such technology. Lack of consideration of ethics could erode trust in the government and public-health services – as happened last month, when the Norwegian Data Protection Authority accused the Norwegian Institute of Public Health of failing to carry out a proper risk assessment of its contact-tracing app, Smittestopp.

### Many approaches

Temporarily restricting some fundamental rights and freedoms might be ethically justifiable in the context of hastening the end of the pandemic. Quarantining individuals, for example, helps to prevent the spread of the disease. Arguably, it might be unethical not to use digital tracing apps when necessary. Nevertheless, much depends on the effectiveness of the app, the goal pursued, the type of system and the context in which it will be deployed.

Countries and regions are taking different approaches. China's Alipay Health Code app assigns a digital QR code to each user, which is colour-coded red, amber or green to indicate that person's quarantine status and thus their ability to move around. People quarantined in Hong Kong must wear an electronic bracelet that shares their location with local authorities through an app. Poland requires citizens to self-isolate for 14 days after returning from overseas, and to

# Is this contact-tracing app ethically justifiable?

**Those responsible for contact-tracing apps should answer the following.**

## Principles: is this the right app to develop?

### 1. Is it necessary?

- Yes, it must be developed to save lives (+).
- No, there are better solutions (-).

### 2. Is it proportionate?

- Yes, the gravity of the situation justifies the potential negative impact (+).
- No, the potential negative impact is disproportionate to the situation (-).

### 3. Is it sufficiently effective, timely, popular and accurate?

- Yes, evidence shows that it will work, is timely, will be adopted by enough people and yields accurate data and insights (+).
- No, it does not work well, is available too late or too early, will not be used widely, and is likely to collect data that have false positives and/or false negatives (-).

### 4. Is it temporary?

- Yes, there is an explicit and reasonable date on which it will cease (+).
- No, it has no defined end date (-).

## Requirements: is this app being developed in the right way?

### 5. Is it voluntary?

- Yes, it is optional to download and install (+).
- No, it is mandatory and people can be penalized for non-compliance (-).

### 6. Does it require consent?

- Yes, people have complete choice over what data are shared and when, and can change this at any time (+).
- No, default settings are to share everything all the time, and this cannot be altered (-).

### 7. Are the data kept private and users' anonymity preserved?

- Yes, data are anonymous and held only on the user's phone. Others who have been in contact are notified only that there is a risk of contagion, not from whom or where. Methods such as differential privacy are used to ensure this. Cyber-resilience is high (+).
- No, data are (re)identifiable owing to the level of data collected, and stored centrally.

Locations of contacts are also available. Cyber-resilience is low (-).

## 8. Can users erase the data?

- Yes, they can do so at will; all data are deleted at the end point (+).
- No, there is no provision for data deletion, nor a guarantee that it can ever be deleted (-).

## 9. Is the purpose of data collection defined?

- Yes, explicitly; for example, to alert users that they have encountered a potentially infected person (+).
- No, the purposes of data collection are not explicitly defined (-).

## 10. Is the purpose limited?

- Yes, it is used for tracing and tracking of COVID-19 only (+).
- No, it can be regularly updated to add extra features that extend its functionality (-).

## 11. Is it used only for prevention?

- Yes, it is used only to enable people voluntarily to limit spread (+).
- No, it is also used as a passport to enable people to claim benefits or return to work (-).

## 12. Is it used for compliance?

- No, it is not used to enforce behaviour (+).
- Yes, non-compliance can result in punishment such as a fine or jail time (-).

## 13. Is it open-source?

- Yes, the code is publicly available for inspection, sharing and collaborative improvement (+).
- No, the source code is proprietary, and no information about it is provided (-).

## 14. Is it equally available?

- Yes, it is free and distributed to anyone (+).
- No, it is arbitrarily given only to some (-).

## 15. Is it equally accessible?

- Yes, it is user-friendly, even for naive users, and works on the widest possible range of mobile phones (+).
- No, it can be used only by those with specific devices and with sufficient digital education (-).

## 16. Is there a decommissioning process?

- Yes, there is a process for shutting it down (+).
- No, there are no policies in place (-).

send geotagged ‘selfies’ to the police to prove they are at home. Singapore’s TraceTogether app has been downloaded by about 25% of its population, much less than the 60% needed. This has led the country to introduce its SafeEntry system, which requires users to check in to public places using their national identity card or by scanning a QR code with their phone.

Apps differ in how they collect and store data. For example, they might rely on systems that are centralized, as in Australia and Singapore, or decentralized, as in Germany and Italy (see also *Nature* <http://doi.org/dwc6>; 2020). Centralized apps send pseudonymized data collected by a user’s phone to a central database controlled by, for example, a national health agency, where contacts are matched. Decentralized approaches instead match contacts on the user’s device (see go.nature.com/3e7jntx). Use of an app can be voluntary, as the European Commission recommended in April (see go.nature.com/2x2hrat), or not. India’s app, for instance, is mandatory for citizens living in virus-containment zones and for all government and private-sector employees. Apps in Argentina and the United Kingdom ask users to self-report their symptoms, whereas the Norwegian app relies on the user having a formal diagnostic test.

More coordination is needed. Some supranational efforts to harmonize the apps are under way. The World Health Organization, for example, is developing a symptom-checking app that might also enable contact tracing in under-resourced countries. The European Data Protection Supervisor has called for a Europe-wide contact-tracing app<sup>5</sup>. The European Commission has outlined requirements for digital tracing solutions deployed in the European Union, including compliance with EU data protection and privacy rules<sup>1</sup>.

Countries and regions should consider a broader set of ethical concerns, including equality and fairness. Government agencies and developers working under pressure might find it hard to make these judgement calls quickly. In other contexts, such as bioethics, ethical review boards typically have much more time to deliberate. Expert groups might be set up to advise, as France, Italy and the United Kingdom have done. (L.F. is a member of the UK National Health Service COVID-19 App Data Ethics Advisory Board; see go.nature.com/3cxyzrw).

## Four principles

To be ethical, a contact-tracing app must abide by four principles: it must be necessary, proportional, scientifically valid and time-bound. These principles are derived from the European Convention on Human Rights, the International Covenant on Civil and Political Rights (ICCPR) and the United

Nations Siracusa Principles, which specify the provisions in the ICCPR that limit how it can be applied.

However, there are many ways in which an app can meet these principles. To address this gap, we have synthesized 16 questions that designers, deployers and evaluators should answer (see ‘Is this contact-tracing app ethically justifiable?’). For each, we give examples of how an app might be designed and used in a more (+) or less (–) ethically justifiable way. These questions apply to apps that have been released, as well as for those in development<sup>6</sup>.

In theory, an ethical app should satisfy all 16 factors. The questions themselves might not be controversial, but the answers are likely to generate disagreement about whether and how much an app satisfies a

## “Governments might not have a second chance to get an intervention right.”

factor, and which ethical factors should be a priority.

In practice, there will be trade-offs. These will depend on the laws, values, attitudes and norms in different regions, as well as on changes over time in the spread and scale of the virus and the available technology. For example, it might be more ethically justifiable to deploy an app that does not fully meet the stipulation that it should “work on the widest possible range of mobile phones” in a country with high smartphone penetration, such as South Korea – where more than 95% of people owned a smartphone in 2018. But it might be less justifiable in Japan, where 66% of the population did.

Similarly, what was ethically justifiable in one place yesterday might not be so tomorrow. For example, Germany shifted from a centralized to a decentralized app after some 300 experts signed an open letter strongly criticizing the centralized approach. The same happened in Italy after Apple and Google announced their plan to support decentralized apps. Singapore could follow suit. Its centralized TraceTogether app was developed before the Apple–Google interface was available, and developers are now aiming to make it compatible.

An app’s implementation strategy and impact must also be considered. Something that looked good on paper can turn out to be ineffective in practice. This was the case with the Australian COVIDsafe app. Concerns about third-party access to user data and low compatibility with phones running old operating systems have led to a low level of adoption. More than a month since deployment, the minimum threshold of 40% has not

been met. This is making the app irrelevant for managing the pandemic in Australia.

If an app fails, it becomes unnecessary, and thus unethical. Apps that are no longer beneficial should be improved or decommissioned. A review and exit strategy must be in place to establish when and how fast this should happen. These assessments should be conducted by an independent body, such as a regulator or an ethics advisory board, and not by the designers or the government itself. Circumstances and attitudes are changing quickly, so the questions in our framework must be asked anew at regular intervals.

## One chance

Governments might not have a second chance to get an intervention right – failure now could breach public trust for the foreseeable future. Governments, developers and deployers must ensure that COVID-19 contact-tracing apps satisfactorily address the ethical questions we set out. Apps that do not should not be deployed; alternatives should be considered.

Simply rolling out a tracing app without ethical consideration is not acceptable. Even in a crisis, a ‘try-everything’ approach is dangerous when it ignores the real costs, including serious and long-lasting harms to fundamental rights and freedoms, and the opportunity costs of not devoting resources to something else.

## The authors

**Jessica Morley** is a graduate researcher at the Oxford Internet Institute, University of Oxford, UK. **Josh Cowls** is a doctoral researcher at the Oxford Internet Institute, University of Oxford, UK, and at The Alan Turing Institute, London, UK. **Mariarosaria Taddeo** is a senior research fellow at the Oxford Internet Institute, University of Oxford, UK, and a Turing Fellow/DSTL Ethics Fellow at The Alan Turing Institute, London, UK. **Luciano Floridi** is professor of philosophy and ethics of information at the Oxford Internet Institute, University of Oxford, UK, and a Turing Fellow and chair of the Data Ethics Group at The Alan Turing Institute, London, UK.

e-mail: pa.floridi@oii.ox.ac.uk

- European Commission. *Commission Recommendation (EU) 2020/518 of 8 April 2020* (EC, 2020); available at <https://go.nature.com/2jkmmpt>
- European Centre for Disease Prevention and Control. *Novel Coronavirus Disease 2019 (COVID-19) Pandemic: Increased Transmission in the EU/EEA and the UK – Sixth Update* (ECDC, 2020).
- Ferretti, L. et al. *Science* **368**, eabb6936 (2020).
- Keeling, M. J., Hollingsworth, T. D. & Read, J. M. Preprint at medRxiv <https://doi.org/10.1101/2020.02.14.20023036> (2020).
- Wiewiórowski, W. *EU Digital Solidarity: A Call for a Pan-European Approach Against the Pandemic* (European Data Protection Supervisor, 2020).
- Floridi, L. *Phil. Trans. R. Soc. A* **376**, 20180081 (2018).

# Correspondence

## Fight pseudoscience responsibly

Battle cries to take up “cudgels” against pseudoscience sparked by COVID-19 (T. Caulfield *Nature* <http://doi.org/ggtbj>; 2020) could backfire. Although Caulfield makes important points about the need to counteract misinformation, war metaphors and hostility are more likely, in our experience, to antagonize perpetrators of misinformation – and so exacerbate the problem. We should instead be pre-emptively disseminating factual evidence so that people become more resistant to false information.

As members of Critica, a non-profit organization that corrects scientific and medical misinformation ([www.criticascience.org](http://criticascience.org)), we contend that the problem does not stem from just a lack of knowledge. Many view COVID-19 as a political rather than a scientific issue, for example. And pseudoscience perpetrators are wary of experts – marginalizing and ignoring them over vaccination, for instance. This must not happen if and when a vaccine against COVID-19 emerges.

More-effective communication by scientists is the key. Although we do not yet have the luxury of making recommendations based on settled science, enough evidence exists to guide our attempts at communication. Respectful online discussion is more likely than ridicule to engage the curious and convince the unconvinced.

**David Scales** Weill Cornell Medicine, New York, USA.  
das9289@med.cornell.edu

**Sara Gorman, Jack Gorman**  
Critica, New York, USA.

## Heat and COVID-19 could be twin killers

People who are especially vulnerable to COVID-19 also tend to be more affected by extreme weather events. Examples include elderly people and those living in densely populated cities. As the pandemic continues, scientists should help politicians to develop plans and policies that protect those most susceptible to the cascade of socio-economic risks that could arise when these hazards combine.

We need to find out how human behaviour is changing in response to COVID-19, how this affects demands for food, energy, the Internet and transportation, and whether these altered demands are likely to increase or decrease people's susceptibility to the effects of extreme weather. Heatwaves could make COVID-19 shelter-in-place policies dangerous, for example, if elderly people or those with low incomes do not have air conditioning.

Cities and scientists must work together to protect urban residents. They need to improve city infrastructure, expand data sharing and create open communication channels between policymakers, planners and researchers. They should focus first on those most in need.

**Timon McPhearson, Ahmed Mustafa, Luis Ortiz** The New School, New York, USA.  
timon.mcphearson@newschool.edu

## Opening up journals' editorial practices

We have launched a Platform for Responsible Editorial Policies ([www.responsiblejournals.org](http://www.responsiblejournals.org)). This facilitates transparent review and research into peer-review procedures, thereby contributing to open science and optimal journal management (see P. Wouters *et al. Nature* **569**, 621–623 (2019) and F. Squazzoni *et al. Nature* **578**, 512–514; 2020).

Funded by the Netherlands Organisation for Health Research and Development, and created with the Leiden Centre for Science and Technology Studies, the platform currently hosts a database of 387 journals. It evaluates these journals' peer-review procedures according to 12 criteria, including: the level of anonymity afforded to authors and reviewers; the use of digital tools such as plagiarism scanners; and the timing of peer review in the research and publication process (see S. P. J. M. Horbach and W. Halffman *Scientometrics* **118**, 339–373; 2019). The platform displays the procedures used by each journal, along with aggregate statistics on the various editorial practices.

We invite journal editors to provide such information for inclusion in the database. In response, we offer tailored suggestions for improvement, information on alternative review procedures and suggestions on fair use of journal metrics for evaluation. The platform makes data on peer review freely available for further research.

**Serge Horbach, Wytske Hepkema, Willem Halffman** Radboud University, Nijmegen, the Netherlands.  
s.horbach@science.ru.nl

## Lockdown classes with Pam

As a research student working from home during the COVID-19 lockdown, I've discovered a productive new pastime – remotely teaching immunology basics to laboratory members.

My research in Terry Dermody's lab aims to define innate immune responses to reovirus infection in the intestine, a condition that could be linked to coeliac disease (R. Bouziat *et al. Science* **356**, 44–50; 2017). When a research assistant joined the project who was not familiar with the immunology involved, I set up a weekly crash course to help her out. We covered topics ranging from interferon signalling to oral tolerance and virus–host interactions.

Then COVID-19 hit. The lab was shut down. As our sessions resumed online, other lab members started to join these weekly 'Immunology with Pam' classes. I propose topics for discussion, research them and create PowerPoint presentations. In a few weeks, we have covered general concepts, innate immune responses and pathogen-recognition receptors. To my surprise, I felt confident – despite my own initially limited grounding in the subject.

The rewards have included a newfound love of teaching and engaging with my otherwise-preoccupied fellow researchers.

**Pamela Briggleb** University of Pittsburgh School of Medicine, Pennsylvania, USA.  
pab149@pitt.edu

# News & views

## Origins of life

# A plausible route to the first genetic alphabet

Kristian Le Vay & Hannes Mutschler

Understanding the prebiotic origins of the nucleic acids is a long-standing challenge. The latest experiments support the idea that the first nucleic acid encoded information using a mixed 'alphabet' of RNA and DNA subunits. See p.60

The genetic polymers RNA and DNA are central to information storage in all biological systems, and as such form the core of most hypotheses about the origin of life. The most prominent of these theories is the 'RNA world' hypothesis, which posits that RNA was once both the central information-carrier and the catalyst for biochemical reactions on Earth before the emergence of life<sup>1</sup>. However, studies in the past few years (see ref. 2, for example) have suggested that the first genetic systems might have been based on nucleic-acid molecules that contain both RNA and DNA nucleotides, which then gradually self-separated into today's RNA and DNA. On page 60, Xu *et al.*<sup>3</sup> offer fascinating experimental support for a mixed RNA–DNA world.

Primordial geochemical processes are thought to have led to the formation of the building blocks of nucleic acids – nucleotides and nucleosides (nucleotides that lack a phosphate group). Under suitable conditions, these building blocks polymerized and the resulting strands eventually replicated, without assistance from modern protein enzymes.

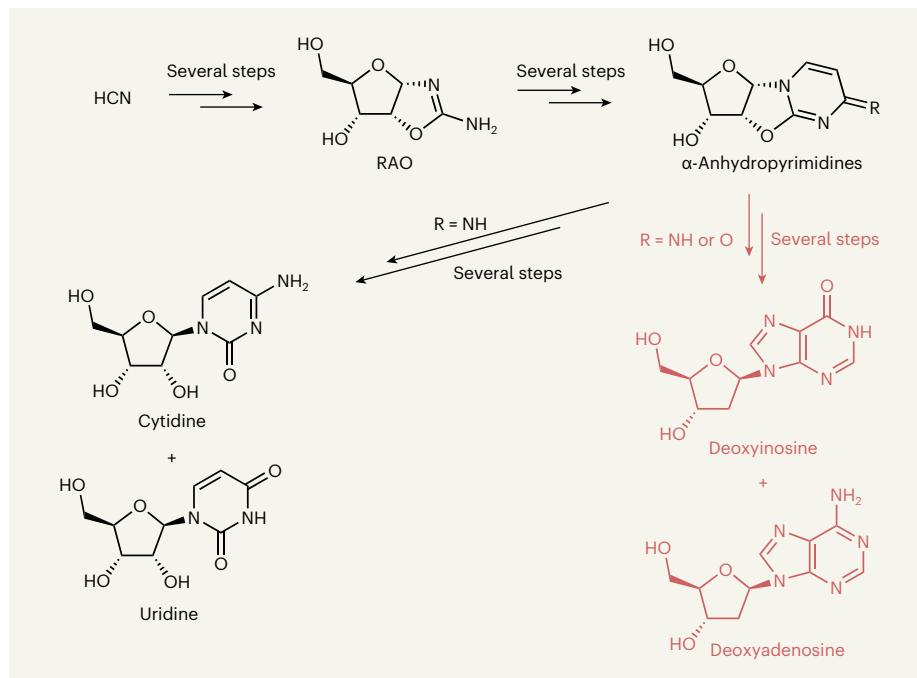
Workers from the same research group as Xu *et al.* had previously identified<sup>4</sup> a network of reactions promoted by ultraviolet light that resulted in the synthesis of two of the standard nucleosides found in RNA: uridine (U) and cytidine (C), which are collectively known as pyrimidines (Fig. 1). These reactions started from hydrogen cyanide (HCN) and derivatives thereof, simple molecules thought to have been readily available on early Earth. Further studies and development of this reaction network raised the intriguing possibility that protein and lipid precursors could have arisen simultaneously alongside nucleosides<sup>5</sup> – thereby providing three of the main types of

molecule needed to make cells. However, a complementary route for the formation of the other two standard RNA nucleosides (adenosine and guanosine, known as the purines) using the same HCN-based chemistry has remained elusive.

In the present work, Xu *et al.* revisited compounds produced as intermediates

in the previously established reaction network<sup>4</sup> that synthesizes U and C. They identified a pathway in which a key intermediate of pyrimidine-nucleoside synthesis, ribo-aminoazoline (RAO; Fig. 1), can also be converted into two purine DNA nucleosides, deoxyadenosine (dA) and deoxyinosine (dI), which is not one of the standard nucleosides found in modern DNA. Crucially, these DNA nucleosides can form base pairs with U and C. The four nucleosides – U, C, dA and dI – therefore constitute a complete 'alphabet' that could have encoded genetic information in nucleic acids in a prebiotic RNA–DNA world.

Importantly, the synthesis of dA and dI can occur in parallel with that of U and C, producing mixtures of the four products in yields and ratios suitable for the construction of a genetic system. This mutual compatibility of the two synthetic pathways increases the plausibility of the reaction network as a prebiotic system – if the two syntheses were incompatible, then geological scenarios would need to be contrived to explain how they could have been separated into different pools to enable the chemistry to occur, and then



**Figure 1 | A reaction network that produces both DNA and RNA subunits.** It was known that a network of chemical reactions produces the RNA subunits cytidine (C) and uridine (U), under conditions that could have occurred on prebiotic Earth<sup>4</sup>. The network starts from hydrogen cyanide (HCN) and proceeds through an intermediate called ribo-aminoazoline (RAO). Xu *et al.*<sup>3</sup> now report that compounds known as  $\alpha$ -anhydropurimidines produced in the pathway to C and U can also be converted in parallel into the DNA subunits deoxyinosine (dI) and deoxyadenosine (dA). These subunits can form base pairs with C and U. The four subunits – C, U, dI and dA – therefore constitute a complete genetic 'alphabet' that might have been used to encode biological information on early Earth.

combined to enable the formation of hybrid RNA–DNA molecules. Notably, under certain reaction conditions, U and C can survive only in the presence of the thioanhydopurine compounds that act as direct precursors of dA and dI.

Many organic molecules can be produced as left- and right-handed versions, known as enantiomers, which are mirror images of each other. However, modern nucleotides and their building blocks all take the same enantiomeric form. One of the main difficulties in origins-of-life research is to explain how single enantiomers could have been generated from simple precursor molecules that have no handedness and which could have formed on prebiotic Earth. Xu and colleagues' purine synthesis is attractive in this respect, because it is highly selective for the enantiomers and other isomers of nucleosides observed in modern biology.

Alternative routes have been reported for the combined prebiotic synthesis of pyrimidine and purine nucleosides and nucleotides<sup>6,7</sup>. These routes require chemically and enantiomerically pure sugars to be used as starting materials, which poses the problem that other, often unknown, prebiotic processes would have been necessary to provide those starting materials<sup>8</sup>. By contrast, the enantioselectivity reported by Xu *et al.* derives from RAO, which can crystallize as a single enantiomer from reactions in which the starting materials are nearly racemic<sup>9</sup> (that is, the starting materials consist of an almost equal mixture of enantiomers).

Nucleoside synthesis can also lead to products in which the nucleoside's base is attached to the sugar in the wrong orientation. In Xu and co-workers' synthetic pathway, a UV-induced chemical reduction occurs that leads to the strikingly selective destruction of these unwanted by-products, ultimately producing only the biologically relevant isomers of the purines. Given that early Earth was highly irradiated by UV, the remarkable selectivity of this reaction suggests a possible mechanism by which the total pool of potential nucleic-acid isomers was reduced to the subset of isomers observed today in nature.

Xu and colleagues' work supports a vision of early molecular evolution somewhat removed from the conventional 'pure' RNA-world hypothesis, and perhaps offers a more plausible route to the origin of life from mixed and complex chemical environments. Given the lack of 'chemical fossils', and the uncertainty over the exact conditions and chemistry that occurred on early Earth, it is impossible to say which chemical pathways actually took place. Instead, we must ensure that proposed systems conform as closely as possible to our understanding of what could realistically have happened on prebiotic Earth – not just the chemistry, but also the overall complexity of

the reaction networks and their compatibility with other processes.

In the current work, the authors show that the four nucleosides can indeed be produced through processes that could reasonably be expected to have occurred on early Earth (such as hydrolysis, drying and UV irradiation), and provide plausible synthetic pathways that could supply the reactions with their required starting materials. However, as for all prebiotic syntheses, it remains hard to envisage the actual microenvironment that could have supported the many specific chemical transformations required to produce the building blocks of life in quantity.

Nevertheless, Xu and colleagues' work impressively demonstrates how a complete genetic alphabet might have arisen. Regardless of whether we think that life developed from RNA alone, or from more-complex mixtures of nucleic acids, systems-level thinking to find mutually compatible prebiotic chemical

pathways will be crucial for developing truly plausible models of the first stages of life's emergence.

**Kristian Le Vay** and **Hannes Mutschler**

are in the Biomimetic Systems research group, Max Planck Institute of Biochemistry, Martinsried 82152, Germany.  
e-mail: mutschler@biochem.mpg.de

1. Joyce, G. F. & Szostak, J. W. *Cold Spring Harb. Perspect. Biol.* **10**, a034801 (2018).
2. Gavette, J. V., Stoop, M., Hud, N. V. & Krishnamurthy, R. *Angew. Chem. Int. Edn* **55**, 13204–13209 (2016).
3. Xu, J. *et al.* *Nature* **582**, 60–66 (2020).
4. Powner, M. W., Gerland, B. & Sutherland, J. D. *Nature* **459**, 239–242 (2009).
5. Patel, B. H., Percivalle, C., Ritson, D. J., Duffy, C. D. & Sutherland, J. D. *Nature Chem.* **7**, 301–307 (2015).
6. Teichert, J. S., Kruse, F. M. & Trapp, O. *Angew. Chem. Int. Edn* **58**, 9944–9947 (2019).
7. Becker, S. *et al.* *Science* **366**, 76–82 (2019).
8. Yadav, M., Kumar, R. & Krishnamurthy, R. *Chem. Rev.* <https://doi.org/10.1021/acs.chemrev.9b00546> (2020).
9. Hein, J. E., Tse, E. & Blackmond, D. G. *Nature Chem.* **3**, 704–706 (2011).

### Neurodegeneration

## Gut microbes tune inflammation and lifespan

**Ping Fang & Elaine Y. Hsiao**

There is growing evidence that gut microbes can influence disease. Analysis of a mouse model of the neurodegenerative condition amyotrophic lateral sclerosis offers insight into how gut bacteria might contribute to this illness. See p.89

Animals have co-evolved with diverse communities of microorganisms that are integral to the development and activity of their immune and nervous systems<sup>1</sup>. Alterations in the composition and function of the community of gut microorganisms (termed the microbiota) are increasingly being implicated in neurological disorders that involve neuroinflammation, including multiple sclerosis<sup>2</sup>, autism spectrum disorder<sup>3</sup> and Parkinson's disease<sup>4</sup>. Studies are also emerging that link the gut microbiota to amyotrophic lateral sclerosis (ALS), a neurodegenerative disorder characterized by the progressive loss of motor neurons crucial for movement, speech and cognition. This devastating disease is usually fatal within a few years of diagnosis. On page 89, Burberry *et al.*<sup>5</sup> fill some gaps in our knowledge of how gut microbes might contribute to ALS, from studies of the condition in a mouse model. Their findings might help to shed light on how a gene linked to ALS called *C9orf72* affects this disease.

Initial studies<sup>6,7</sup> have shown that the gut microbiota of people who have ALS differ

from those of unaffected individuals. A study of a mouse model of the disease, based on an ALS-associated mutation in the *Sod1* gene<sup>6</sup>, has provided strong evidence that alterations in the microbiota can exacerbate neurodegeneration and drive early mortality. That study also identified microbes and microbial molecules that promote improved motor function and longer lifespan in the mice. It showed that the particular positive or negative effects observed might depend on differences in the microbes encountered in the animals' housing facility (termed a vivarium). Mouse models of inflammatory diseases have also revealed that the animals' environment has such an effect<sup>8</sup>.

Burberry *et al.* used a mouse model of ALS (Fig. 1) in which the animals have a mutant version of the gene *C9orf72*, resulting in a deficiency in the encoded *C9orf72* protein (these mice also model a neurodegenerative condition called frontotemporal dementia). The authors observed that if the animals were reared in the Harvard University animal facility, they had a shorter lifespan,

## Reproducibility

# Pipeline choices alter neuroimaging findings

Martin Lindquist

Seventy laboratories that analysed the same neuroimaging data each produced different results. This finding highlights the potential consequences of a lack of standardized pipelines for processing complex data. See p.84

For most types of big data, from genome sequences to medical images, there is no single 'best' way to process the data. This issue is exemplified by the substantial differences in how individual laboratories preprocess and analyse data from functional magnetic resonance imaging (fMRI) experiments, which generate information about brain activity. Indeed, a survey of fMRI studies found that nearly every study used a different analysis pipeline<sup>1</sup>. Botvinik-Nezer *et al.*<sup>2</sup> provide further evidence of this variability on page 84, highlighting how analytical choices made by individual researchers can greatly influence the findings gleaned from an fMRI data set. The work is bound to spark lively discussion.

Functional MRI experiments produce a series of images of the brain at work. These images go through several stages of processing and analysis to determine which brain regions show significant activity. However, the choice of pipeline can alter the outcome of an fMRI study. In 2012, for instance, 6,912 unique processing and analysis pipelines were applied to the same fMRI data set<sup>3</sup>. Many of the pipelines indicated that the same

general brain regions were active, but the locations at which neural activity was deemed to be highest varied widely depending on the pipeline used.

Botvinik-Nezer *et al.* explored this phenomenon further. The authors gave 70 independent research teams the same fMRI data set, generated from 108 people performing decision-related tasks. The

## "The fact that each team's results were so pipeline-dependent is highly problematic."

teams were asked to use the data set to test nine hypotheses, each of which posited that activity in a specific brain region related to a particular feature of the tasks. This allowed Botvinik-Nezer *et al.* to evaluate the impact of analytical flexibility on fMRI results 'in the wild' (rather than performing the analyses themselves, as was done in the 2012 study). Notably, no two teams chose identical

workflows to analyse the data, resulting in substantial variation in the results (Fig. 1).

It is standard in neuroimaging to test types of hypothesis such as the nine put forward by Botvinik-Nezer *et al.* by generating a statistical map. The map comprises a compendium of statistical tests performed on different parts (voxels) of the image. The results of these tests are subjected to a process called thresholding to set the level, for each voxel, at which activity picked up by the experiment is deemed to indicate real neuronal activity rather than noise. The authors found considerable differences between each research team's results even when the underlying statistical maps they had used were highly correlated.

The strongest factor in explaining the differences between each team's results was the spatial smoothness of the data being analysed. Spatial smoothing is a preprocessing step in which the activity of each voxel is averaged with that of its neighbours – a process designed to reduce noise. Higher estimated smoothness was associated with a greater likelihood of reaching a significant outcome for each hypothesis.

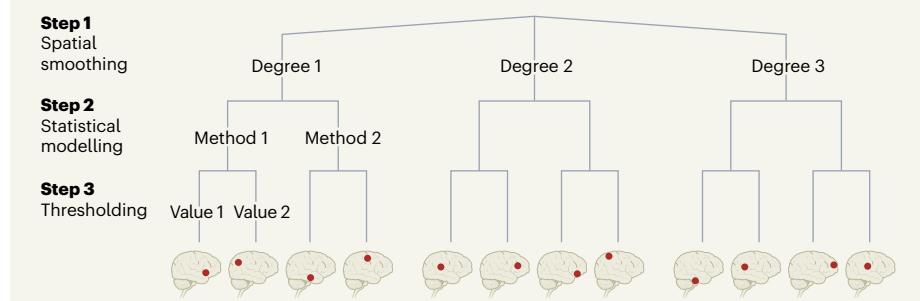
The fact that each team's results were so pipeline-dependent is highly problematic, particularly because the exact configuration of analytical pipelines is often poorly described in research articles. Moreover, sensitivity analyses – which assess how different pipeline choices might affect an experiment's outcome – are rarely performed in neuroimaging. However, Botvinik-Nezer and colleagues offer several reasonable suggestions for addressing the concerns that their work will raise.

The first is to share unthresholded activity maps, because this will allow image-based meta-analysis. The authors found that such an analysis, which aggregated information across teams, yielded consensus results, no doubt aided by the fact that the spatial patterns in the activity maps were highly correlated across groups.

Second is a call to publicly share both data and code, making it easier for others to attempt to reproduce a paper's findings. In this regard, the authors model good behaviour by making all their data and processing pipelines publicly available.

Third is the use of pre-registration – in which a hypothesis and analysis plan is made public before the experiment is performed. It is unfortunately common for researchers to explore various pipelines to find the version that yields the 'best' results, ultimately reporting only that pipeline and ignoring the others. This practice can lead to errors and make it difficult to replicate findings<sup>4</sup>. Pre-registration would make it easy to detect cases in which researchers had explored various pipelines.

Fourth is to analyse all data through



**Figure 1 | Implications of choosing a neuroimaging pipeline.** Botvinik-Nezer *et al.*<sup>2</sup> report that researchers process neuroimaging data using a wide variety of pipelines, which can produce varying results. In this simplified example, the pipeline has three steps: spatial smoothing of the images to reduce noise, which in this example is done to three different degrees; statistical modelling, which in this example can be performed in one of two ways; and 'thresholding' of statistical tests associated with these models to determine the level at which neuronal activity in each brain region is deemed to be significant, which in this example is set to two different values. Making different choices for each step leads to a different end point – the red dots represent how activation moves throughout the brain depending on which pipeline is used.

multiple pipelines and use the results to obtain consensus findings. This could be achieved by implementing the type of meta-analysis used by Botvinik-Nezer and co-workers.

Another goal of the study was to evaluate how accurately researchers could predict the number of teams that would report significant results for each hypothesis. To study this, the authors ran separate ‘prediction markets’, one for the analysis teams and one for researchers who did not participate in the analysis. In them, researchers attempted to predict the outcomes of the scientific analyses and received monetary payouts on the basis of how well they predicted performance. Participants – even researchers who had direct knowledge of the data set – consistently overestimated the likelihood of significant findings. Botvinik-Nezer *et al.* do not explicitly explore how the analysts’ prior beliefs affected their findings and pipeline choices. For example, if a research finding does not initially coincide with expectations, will groups seek to alter pipelines until expectations and results align? The prevalence of pipeline exploration implies that this is likely.

What other improvements could be made for the future? One approach is to use pipeline-optimization tools<sup>5,6</sup> to reduce analysis flexibility. These tools automatically identify pipelines that maximize reproducibility, which can reduce the risk of excessive pipeline exploration and selective reporting. In addition, increased use of sensitivity analyses to evaluate the effects of pipeline decisions would provide a better understanding of the link between analysis choices and research findings. The fact that activity maps were highly correlated across groups implies that multivariate statistical approaches, which identify spatial patterns in the data, might provide more-consistent results across pipelines than would a series of tests performed at individual voxels.

Ultimately, neuroimaging results should be carefully verified using independent data sets to demonstrate generalizability across samples, research contexts and populations. A positive example of this already being done comes from an approach for developing predictive models on the basis of brain activation, which can be shared, tested in multiple contexts and used in applied settings<sup>7</sup>.

It seems unlikely that the fMRI field will ever coalesce on a standard workflow that is applicable to all types of study, because studies tend to be too varied for one pipeline to always be appropriate. But Botvinik-Nezer *et al.* conclude their paper by calling for an increased awareness of the situation, and a drive to improve the quality of method reporting. This is wise and prudent advice that researchers in any field analysing high-dimensional data would be well advised to heed.

**Martin Lindquist** is in the Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21205, USA.  
e-mail: mlindqui@jhsph.edu

3. Carp, J. *Front. Neurosci.* **6**, 149 (2012).
4. Simmons, J. P., Nelson, L. D. & Simonsohn, U. *Psychol. Sci.* **22**, 1359–1366 (2011).
5. Strother, S. C. *et al.* *NeuroImage* **15**, 747–771 (2002).
6. Churchill, N. W. *et al.* *Hum. Brain Mapp.* **33**, 609–627 (2012).
7. Woo, C. W., Chang, L. J., Lindquist, M. A. & Wager, T. D. *Nature Neurosci.* **20**, 365 (2017).

This article was published online on 20 May 2020.

## Optical physics

# Light turned into exotic Laughlin matter

Laura Corman

A Laughlin state is a phase of matter that has remarkable features, such as excitations that behave as a fraction of a particle. The long-sought creation of a photonic Laughlin state is a milestone for the field of quantum simulation. See p.41

Solving a difficult physics problem can be surprisingly similar to assembling an interlocking mechanical puzzle. In both cases, the particles or pieces look alike, but can be arranged into a beautiful structure that relies on the precise position of each component (Fig. 1). In 1983, the physicist Robert Laughlin made a puzzle-solving breakthrough by explaining the structure formed by interacting electrons in a device known as a Hall bar<sup>1</sup>. Although the strange behaviour of these electrons still fascinates physicists, it is not possible to simulate such a system or accurately measure the particles’ ultrashort

time and length scales. On page 41, Clark *et al.*<sup>2</sup> report the creation of a non-electronic Laughlin state made of composite matter–light particles called polaritons, which are easier to track and manipulate than are electrons.

To picture a Laughlin state, consider a Hall bar, in which such states are usually observed (Fig. 2a). In these devices, electrons that are free to move in a two-dimensional plane are subjected to a strong magnetic field perpendicular to the plane. In classical physics, an electron at any position will start moving along a circular trajectory known as a cyclotron orbit, the radius of which depends on the particle’s kinetic



**Figure 1 | An interlocking mechanical puzzle.** Certain many-body phases of matter, such as Laughlin states, depend on the precise position of all the particles in the system, just like the pieces in an interlocking puzzle.

multiple pipelines and use the results to obtain consensus findings. This could be achieved by implementing the type of meta-analysis used by Botvinik-Nezer and co-workers.

Another goal of the study was to evaluate how accurately researchers could predict the number of teams that would report significant results for each hypothesis. To study this, the authors ran separate ‘prediction markets’, one for the analysis teams and one for researchers who did not participate in the analysis. In them, researchers attempted to predict the outcomes of the scientific analyses and received monetary payouts on the basis of how well they predicted performance. Participants – even researchers who had direct knowledge of the data set – consistently overestimated the likelihood of significant findings. Botvinik-Nezer *et al.* do not explicitly explore how the analysts’ prior beliefs affected their findings and pipeline choices. For example, if a research finding does not initially coincide with expectations, will groups seek to alter pipelines until expectations and results align? The prevalence of pipeline exploration implies that this is likely.

What other improvements could be made for the future? One approach is to use pipeline-optimization tools<sup>5,6</sup> to reduce analysis flexibility. These tools automatically identify pipelines that maximize reproducibility, which can reduce the risk of excessive pipeline exploration and selective reporting. In addition, increased use of sensitivity analyses to evaluate the effects of pipeline decisions would provide a better understanding of the link between analysis choices and research findings. The fact that activity maps were highly correlated across groups implies that multivariate statistical approaches, which identify spatial patterns in the data, might provide more-consistent results across pipelines than would a series of tests performed at individual voxels.

Ultimately, neuroimaging results should be carefully verified using independent data sets to demonstrate generalizability across samples, research contexts and populations. A positive example of this already being done comes from an approach for developing predictive models on the basis of brain activation, which can be shared, tested in multiple contexts and used in applied settings<sup>7</sup>.

It seems unlikely that the fMRI field will ever coalesce on a standard workflow that is applicable to all types of study, because studies tend to be too varied for one pipeline to always be appropriate. But Botvinik-Nezer *et al.* conclude their paper by calling for an increased awareness of the situation, and a drive to improve the quality of method reporting. This is wise and prudent advice that researchers in any field analysing high-dimensional data would be well advised to heed.

**Martin Lindquist** is in the Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21205, USA.  
e-mail: mlindqui@jhsph.edu

3. Carp, J. *Front. Neurosci.* **6**, 149 (2012).
4. Simmons, J. P., Nelson, L. D. & Simonsohn, U. *Psychol. Sci.* **22**, 1359–1366 (2011).
5. Strother, S. C. *et al.* *NeuroImage* **15**, 747–771 (2002).
6. Churchill, N. W. *et al.* *Hum. Brain Mapp.* **33**, 609–627 (2012).
7. Woo, C. W., Chang, L. J., Lindquist, M. A. & Wager, T. D. *Nature Neurosci.* **20**, 365 (2017).

This article was published online on 20 May 2020.

## Optical physics

# Light turned into exotic Laughlin matter

Laura Corman

A Laughlin state is a phase of matter that has remarkable features, such as excitations that behave as a fraction of a particle. The long-sought creation of a photonic Laughlin state is a milestone for the field of quantum simulation. See p.41

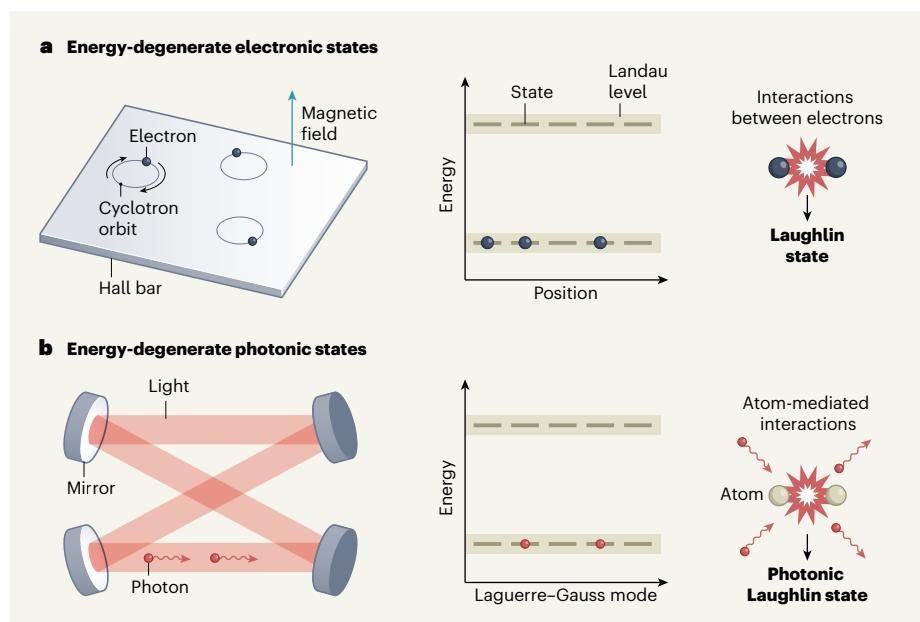
Solving a difficult physics problem can be surprisingly similar to assembling an interlocking mechanical puzzle. In both cases, the particles or pieces look alike, but can be arranged into a beautiful structure that relies on the precise position of each component (Fig. 1). In 1983, the physicist Robert Laughlin made a puzzle-solving breakthrough by explaining the structure formed by interacting electrons in a device known as a Hall bar<sup>1</sup>. Although the strange behaviour of these electrons still fascinates physicists, it is not possible to simulate such a system or accurately measure the particles’ ultrashort

time and length scales. On page 41, Clark *et al.*<sup>2</sup> report the creation of a non-electronic Laughlin state made of composite matter–light particles called polaritons, which are easier to track and manipulate than are electrons.

To picture a Laughlin state, consider a Hall bar, in which such states are usually observed (Fig. 2a). In these devices, electrons that are free to move in a two-dimensional plane are subjected to a strong magnetic field perpendicular to the plane. In classical physics, an electron at any position will start moving along a circular trajectory known as a cyclotron orbit, the radius of which depends on the particle’s kinetic



**Figure 1 | An interlocking mechanical puzzle.** Certain many-body phases of matter, such as Laughlin states, depend on the precise position of all the particles in the system, just like the pieces in an interlocking puzzle.



**Figure 2 | How to make Laughlin matter.** **a**, An electronic Laughlin state is usually produced in a device called a Hall bar. There, electrons are subjected to a strong perpendicular magnetic field and move along circular trajectories known as cyclotron orbits. This set-up leads to large sets of equal-energy (energy-degenerate) electronic states, collectively called Landau levels. When repulsive interactions are added, the electrons enter a Laughlin state. **b**, Clark *et al.*<sup>2</sup> created a photonic Laughlin state by passing light between a system of four mirrors. This arrangement caused modes of light known as Laguerre–Gauss modes to have the same energy as each other, forming analogues of Landau levels. Finally, the authors added atom-mediated interactions between the photons.

energy. In quantum mechanics, the electron's position will still be free, but its orbital radius – and, therefore, its kinetic energy – can be increased or decreased only in discrete steps. This feature leads to large sets of equal-energy (energy-degenerate) states called Landau levels. Non-interacting electrons added to the lowest-energy Landau level can be distributed between the level's energy-degenerate states in many different ways.

Adding repulsive interactions between the electrons constrains the particles' distribution over the states of the lowest Landau level, favouring configurations in which any two electrons have zero probability of being at the same spot. The states described by Laughlin have exactly this property and explain the main features of the fractional quantum Hall effect, whereby electrons in a strong magnetic field act together to behave like particles that have fractional electric charge. This work earned Laughlin a share of the 1998 Nobel Prize in Physics. Laughlin states are truly many-body states that cannot be described by typical approximations, such as the mean-field approximation. Instead, the state of each particle depends on the precise state of all the others, just as in an interlocking puzzle.

To obtain a Laughlin state composed of polaritons, Clark and co-workers needed to add two elements to a photonic system. The first was an equivalent of the magnetic field that would cause photonic states to become energy degenerate. The second was a way for

photons to interact with one other, which would force the system to form a Laughlin state, rather than any other combination of degenerate lowest-energy states. In a pioneering set of experiments, the authors' research group had previously demonstrated each element independently<sup>3,4</sup>. In the current work, these elements were combined.

Clark and colleagues used a photonic system called an optical cavity that consisted of four mirrors (Fig. 2b). This set-up caused different

## “This work is a milestone for the field of quantum simulation.”

Laguerre–Gauss modes (modes of light that have doughnut-shaped intensity profiles) to have the same energy as each other, forming analogues of Landau levels. Light propagating in the cavity interacted with a cloud of rubidium atoms to promote these atoms to highly excited (Rydberg) states, which are relatively large and interact strongly with one another. The interacting cavity photons and atoms formed the polaritons that the researchers investigated. These polaritons inherited both the energy degeneracy of their photonic part and the interaction of their atomic part, and could, therefore, enter a Laughlin state.

The authors observed the smallest possible Laughlin state – one made of two polaritons.

They demonstrated the properties of this quantum state by carefully studying the light emitted from the cavity. Because the time-scales in this system are much longer than those in electronic systems, Clark *et al.* could measure the number of photons in each light mode, as well as the photons' temporal and spatial correlations, proving the formation of this long-sought state.

This work is a milestone for the field of quantum simulation<sup>5,6</sup>. Physicists in this field build experiments to simulate phenomena occurring in systems that are difficult to manipulate and impossible to model computationally. The simulation system is completely different from that of the original phenomenon, but has the same number of control parameters, and can be used to try to reproduce and understand the underlying principles of the phenomenon.

Clark and colleagues' results further validate the quantum-simulation approach to studying strongly correlated systems. This approach is key to finding materials that have exotic properties and excitations that could revolutionize electronics or quantum computing. The work was implemented on a platform pioneered by the authors' research group, clearing hurdles at which others had stumbled. For instance, the experiment subtly managed to enforce energy degeneracy, which is difficult to attain when the system is not intrinsically symmetric<sup>7</sup>, as was the case here.

The main achievement of this work was to observe the smallest possible Laughlin state made of polaritons, and to carefully characterize it. Future studies could produce larger states that are computationally intractable and probe the excitations of this system. Such advances would test the ability of the platform to accommodate more than two polaritons and to build bigger puzzles of matter and light.

**Laura Corman** is at X-Rite Europe GmbH, 8150 Regensdorf, Switzerland.  
e-mail: lauracorman@xrite.com

1. Laughlin, R. B. *Phys. Rev. Lett.* **50**, 1395–1398 (1983).
2. Clark, L. W., Schine, N., Baum, C., Jia, N. & Simon, J. *Nature* **582**, 41–45 (2020).
3. Schine, N., Ryou, A., Gromov, A., Sommer, A. & Simon, J. *Nature* **534**, 671–675 (2016).
4. Jia, N. *et al.* *Nature Phys.* **14**, 550–554 (2018).
5. Cirac, J. I. & Zoller, P. *Nature Phys.* **8**, 264–266 (2012).
6. Carusotto, I. & Ciuti, C. *Rev. Mod. Phys.* **85**, 299–366 (2013).
7. Clark, L. W. *et al.* *Nature* **571**, 532–536 (2019).

## Parkinson's disease

# Extra help for neuron grafts

Lorenz Studer & Viviane Tabar

Grafts of stem-cell-derived precursors of dopamine neurons could be used to treat Parkinson's disease, but this approach has limitations. Injecting a growth factor three weeks after transplantation can overcome some of these limits.

The idea of repairing the brain by replacing the neurons that die in Parkinson's disease has been a long-standing dream for researchers in the field. Over the past few decades, many cell types have been put forward as candidates to replace dying neurons<sup>1</sup>, including cells derived from the midbrain of human fetuses. But the focus has now turned to pluripotent stem cells, which can give rise to almost any cell type. Transplantation of specific neuron progenitors derived from these stem cells is on the verge of being tested in the clinic<sup>2</sup>, but limitations remain, including low rates of survival and suboptimal function of the transplanted neurons<sup>3</sup>. Writing in *Cell Stem Cell*, Gantner *et al.*<sup>4</sup> address some of these limitations.

Parkinson's disease is caused by the degeneration of a population of neurons that reside in the brain's substantia nigra and project to the striatum, where they release the neurotransmitter molecule dopamine. A plethora of growth factors are involved in the development of these dopaminergic neurons in the immature brain, but these factors are either lacking or expressed at much lower levels in the adult brain. The introduction of growth factors into the adult brain, therefore, might be a way

to improve the success of cell-engraftment approaches.

Gantner *et al.* set out to investigate the possibility of using one such factor, glial-cell-line-derived neurotrophic factor (GDNF). This type of approach has a long history. Several clinical studies have analysed whether the administration of GDNF to the brain could improve the outlook for people with Parkinson's disease (although these attempts have met with limited success<sup>5</sup>). And work has indicated that GDNF can enhance the survival and function of grafts of fetal dopaminergic neurons<sup>6</sup>.

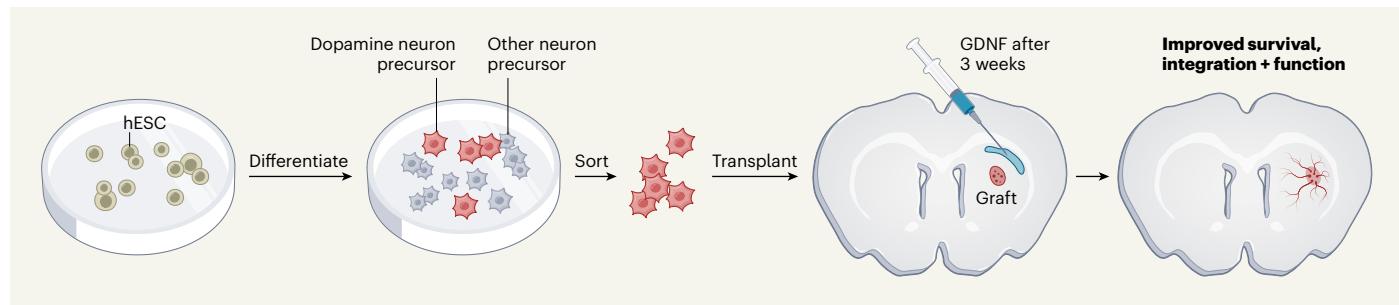
Gantner and colleagues used a type of pluripotent stem cell called human embryonic stem cells. They engineered the cells to fluoresce when they expressed either a protein specific to midbrain dopaminergic neurons, called PITX3, or a protein specific to dopaminergic-neuron progenitors, called LMX1A. This approach allowed the researchers to easily quantify the number of dopaminergic neurons and precursors both *in vitro* and in brain grafts. The group used an established protocol to convert the stem cells to dopaminergic-neuron progenitors *in vitro*, before transplanting them into the brains of mice

or rats lacking dopaminergic neurons in the midbrain (Fig. 1). They then assessed how GDNF (injected into the striatum in a viral vector) affected the survival of the grafted dopaminergic neurons and their integration into the rodent brain.

The authors delivered GDNF one week before or three weeks after transplantation – time points selected to capture the effects of GDNF on the initial survival of the transplanted cells or on later stages of neuron maturation and integration. They found that GDNF delivery increased neuron survival at both time points, although more so if it was delivered before the transplant. However, several lines of evidence indicated that only late GDNF delivery improved functional integration into the brain – this timing resulted in increased density of neuronal fibres in the striatum, increased synaptic connections between grafted neurons and host neurons, and improvements in motor coordination.

The work, therefore, demonstrates that it is crucial to understand the biology of GDNF action at each stage of dopaminergic-neuron development if this type of treatment is to be beneficial. It also suggests that GDNF treatment could be used to improve the next wave of stem-cell-based neuron transplants. It is likely that the optimal timing of GDNF delivery in humans would be comparable, because the timing of survival and maturation of transplanted human neurons is remarkably similar when grafted into a mouse, rat or primate brain<sup>7</sup>.

In addition to problems with neuronal integration, current approaches to stem-cell differentiation result in a mixed cell population from which only a relatively low percentage of grafted cells become dopaminergic neurons. For example, six months after Gantner *et al.* transplanted unsorted dopaminergic-neuron precursors into the rodents, less than 1% of cells in the graft were dopaminergic neurons – even in the presence of GDNF. The authors took advantage of the fact that their



**Figure 1 | Improving grafts to treat Parkinson's disease.** There is a focus on stem-cell-based transplants to treat Parkinson's disease, but problems exist with survival and integration of transplanted neurons in the brain. Gantner *et al.*<sup>4</sup> have outlined a protocol that could help to address these issues. The authors used human embryonic stem cells (hESCs) that were engineered to fluoresce (not shown) when they expressed a protein specific to precursors of dopamine-releasing neurons. The group differentiated the hESCs into neuronal

precursors, then sorted the fluorescing cells from the rest of the population. They transplanted these cells into the brains of rodents that lacked neurons in the striatum – the region of the brain affected in Parkinson's disease. After three weeks, the researchers injected the growth factor GDNF into the brain. Together, cell sorting and GDNF treatment improved neuronal survival and functional integration into the brain, reducing motor symptoms of Parkinson's disease in the animals.

## News & views

cells carried fluorescent reporters to address this point, by purifying the cells before transplantation and engrafting only those that expressed LMX1A. This approach greatly improved the percentage of dopaminergic neurons in the graft, to 10% of total cells after six months.

Gantner and colleagues' experiments point to ways to improve the efficiency of transplant approaches to treat Parkinson's disease. However, they also highlight further issues that need to be overcome. For instance, in unsorted grafts, GDNF treatment increased overall cell numbers (and therefore the number of cells of unwanted types), highlighting that improvements in cell survival alone are not enough. The fluorescence-based sorting approach used by the authors would present considerable manufacturing and scalability challenges in the clinic. However, drug-based or magnetic sorting strategies might be appropriate alternatives. In addition, it will be important to demonstrate the long-term safety of combining grafting and sustained GDNF expression, ideally in primates, before testing the authors' approach in people.

Another major question going forward – both for approaches to treat Parkinson's disease and for regenerative therapies more broadly – is whether generating the 'authentic' cell that is lost in disease is sufficient, or

whether genetically engineering cells could enhance the benefit of transplants. The cell-lineage reporters used by Gantner and colleagues are just one example of the second approach. Another prominent idea is the development of genetically engineered universal donor cells<sup>8,9</sup>. Currently, many cell-based therapies use donor cells, in which case patients must receive immunosuppressive treatments to prevent graft rejection. A universal donor cell engineered to lack surface antigen proteins that can be recognized by the immune system would prevent this issue.

For Parkinson's disease specifically, cells engineered to express lower-than-normal levels of the disease-associated protein α-synuclein are of interest, because this approach might prevent disease spreading into the graft<sup>10</sup>. Finally, methods to genetically control graft activity or connectivity with host cells could help to fine-tune graft function according to the individual needs of the patient, whose disease condition might be changing over time<sup>11</sup>.

The increasingly routine nature of genetic engineering in pluripotent stem cells offers an exciting opportunity for regenerative medicine – being able to generate any cell type expressing any therapeutic gene product on demand. However, to take advantage of such technology, it is paramount to understand the

biology of regeneration. Taking another look at molecules such as GDNF is a timely advance that could give neuronal grafts some extra help, optimizing therapeutic success.

**Lorenz Studer** is in the Developmental Biology Program and the Center for Stem Cell Biology, the Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. **Viviane Tabar** is in the Department of Neurosurgery, the Cancer Biology & Genetics Program and the Center for Stem Cell Biology, the Memorial Sloan Kettering Cancer Center. e-mails: studerl@mskcc.org; tabarv@mskcc.org

1. Barker, R. A., Drouin-Ouellet, J. & Parmar, M. *Nature Rev. Neurosci.* **11**, 492–503 (2010).
2. Barker, R. A., Parmar, M., Studer, L. & Takahashi, J. *Cell Stem Cell* **21**, 569–573 (2017).
3. Tabar, V. & Studer, L. *Nature Rev. Genet.* **15**, 82–92 (2014).
4. Gantner, C. W. et al. *Cell Stem Cell* **26**, 511–526 (2020).
5. Kirik, D., Cederfjäll, E., Halliday, G. & Petersén, A. *Neurobiol. Dis.* **97**, 179–188 (2017).
6. Rosenblad, C., Martínez-Serrano, A. & Björklund, A. *Neuroscience* **75**, 979–985 (1996).
7. Kikuchi, T. et al. *Nature* **548**, 592–596 (2017).
8. Deuse, T. et al. *Nature Biotechnol.* **37**, 252–258 (2019).
9. Gornalusse, G. G. et al. *Nature Biotechnol.* **35**, 765–772 (2017).
10. Luk, K. C. et al. *Science* **338**, 949–953 (2012).
11. Steinbeck, J. A. & Studer, L. *Neuron* **86**, 187–206 (2015).

The authors declares competing financial interests: see go.nature.com/3ehtfkc for details.

This article was published online on 26 May 2020.

nature research

## Publishing high-quality Research & Reviews in all areas of neuroscience.

Discover our portfolio of leading journals which cover all areas of neuroscience, including Research & Reviews, News, Commentaries & Historical perspectives.

**Nature Neuroscience:** [nature.com/neuro](http://nature.com/neuro)

**Nature Reviews Neuroscience:** [nature.com/nrn](http://nature.com/nrn)

**Nature Human Behaviour:** [nature.com/nathumbehav](http://nature.com/nathumbehav)



Nature

@NatureNeuro

@NatRevNeurosci

@NatureHumBehav

A86030

Part of SPRINGER NATURE

## Article

# Observation of Laughlin states made of light

<https://doi.org/10.1038/s41586-020-2318-5>

Logan W. Clark<sup>1,2,3</sup>, Nathan Schine<sup>1,2,3</sup>, Claire Baum<sup>1,2</sup>, Ningyuan Jia<sup>1,2</sup> & Jonathan Simon<sup>1,2</sup>✉

Received: 12 July 2019

Accepted: 31 March 2020

Published online: 3 June 2020

 Check for updates

Much of the richness in nature emerges because simple constituents form an endless variety of ordered states<sup>1</sup>. Whereas many such states are fully characterized by symmetries<sup>2</sup>, interacting quantum systems can exhibit topological order and are instead characterized by intricate patterns of entanglement<sup>3,4</sup>. A paradigmatic example of topological order is the Laughlin state<sup>5</sup>, which minimizes the interaction energy of charged particles in a magnetic field and underlies the fractional quantum Hall effect<sup>6</sup>. Efforts have been made to enhance our understanding of topological order by forming Laughlin states in synthetic systems of ultracold atoms<sup>7,8</sup> or photons<sup>9–11</sup>. Nonetheless, electron gases remain the only systems in which such topological states have been definitively observed<sup>6,12–14</sup>. Here we create Laughlin-ordered photon pairs using a gas of strongly interacting, lowest-Landau-level polaritons as a photon collider. Initially uncorrelated photons enter a cavity and hybridize with atomic Rydberg excitations to form polaritons<sup>15–17</sup>, quasiparticles that here behave like electrons in the lowest Landau level owing to a synthetic magnetic field created by Floquet engineering<sup>18</sup> a twisted cavity<sup>11,19</sup> and by Rydberg-mediated interactions between them<sup>16,17,20,21</sup>. Polariton pairs collide and self-organize to avoid each other while conserving angular momentum. Our finite-lifetime polaritons only weakly prefer such organization. Therefore, we harness the unique tunability of Floquet polaritons to distil high-fidelity Laughlin states of photons outside the cavity. Particle-resolved measurements show that these photons avoid each other and exhibit angular momentum correlations, the hallmarks of Laughlin physics. This work provides broad prospects for the study of topological quantum light<sup>22</sup>.

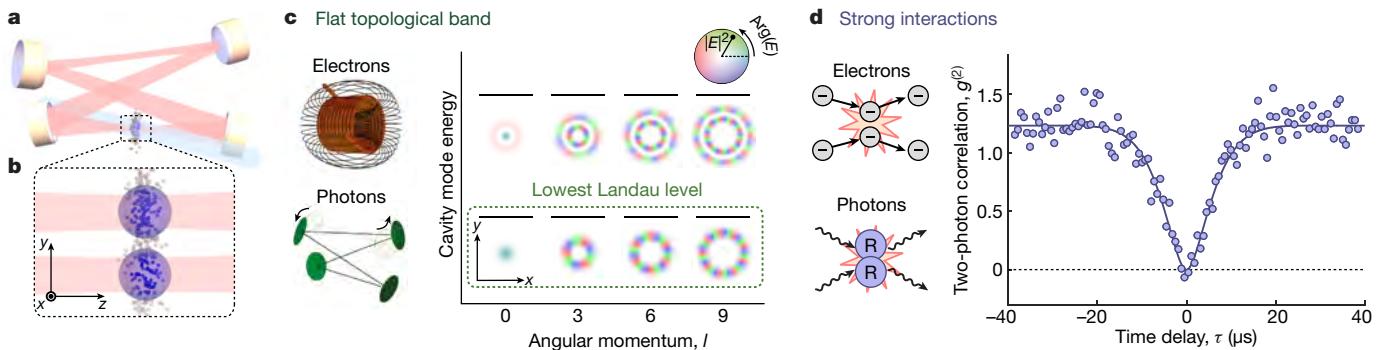
The simplest method for realizing topological order is to place strongly interacting particles in an effective magnetic field. The magnetic field quenches the kinetic energy of the particles so that they order themselves solely to minimize their interaction energy, resulting in intricate patterns of long-range entanglement<sup>3</sup>. These states exhibit fascinating properties largely unseen in other forms of matter; in addition to robust quantized edge transport, which also appears without interactions<sup>23</sup>, topologically ordered phases host excitations with fractional charge and anyonic statistics<sup>24</sup>. More exotic phases even host non-Abelian anyons, promising constituents for fault-tolerant quantum computers owing to their insensitivity to local perturbations<sup>4</sup>.

The scarcity of physical platforms hosting topological order has spurred great interest in explaining its exotic properties using the wide tunability, particle-resolved control and versatile detection capabilities afforded by synthetic quantum matter<sup>7,8,10,11</sup>. The typical constituents of synthetic matter are atoms and photons, neither of which experience a Lorentz force in ordinary magnetic fields, as they are charge neutral. Accordingly, the key challenge is to implement a synthetic magnetic field that creates an effective Lorentz force and is compatible with strong interparticle interactions. A classic approach employed the Coriolis force in rotating ultracold atomic gases<sup>25</sup>, and such systems approached the few-body fractional quantum Hall regime<sup>26</sup>. More recent efforts with ultracold atoms have focused on Floquet engineering of synthetic magnetic fields<sup>27</sup> combined with

strong atomic interactions owing to tight optical lattice confinement<sup>28</sup>. Photonic systems have also shown a variety of synthetic magnetic fields<sup>11</sup> compatible with strong interactions via coupling to superconducting qubits in the microwave domain<sup>29,30</sup> and cold atoms<sup>16,17,20,21</sup> or quantum dots<sup>31</sup> in the optical domain. However, synthetic magnetic fields and strong interactions have yet to be effectively combined and scaled, so the formation of topologically ordered synthetic matter has not yet been achieved.

Here we assemble optical photon pairs into a Laughlin state. To achieve this, we construct a polariton system analogous to an electronic fractional quantum Hall fluid by combining two key components: a synthetic magnetic field for light induced by a twisted optical cavity<sup>19</sup> and strong polariton interactions mediated by Rydberg atoms<sup>17</sup>. We inject photons with angular momentum into this system and observe that the resulting intracavity polaritons do indeed behave like fractional quantum Hall electrons, undergoing mode-changing, angular-momentum-conserving collisions to minimize their interactions with one another. Given the finite lifetime of our polaritons, this leads towards but not to an intracavity Laughlin state. We then distil a Laughlin state of photons outside the cavity by manipulating the optical leakage rate of each single-polariton state and compensating for a relative phase between photons and polaritons (Supplementary Information section A1). The distilled photon pairs exhibit angular momentum correlations and spatial avoidance characteristic of a Laughlin state.

<sup>1</sup>James Franck Institute, University of Chicago, Chicago, IL, USA. <sup>2</sup>Department of Physics, University of Chicago, Chicago, IL, USA. <sup>3</sup>These authors contributed equally: Logan W. Clark, Nathan Schine. ✉e-mail: simonjon@uchicago.edu



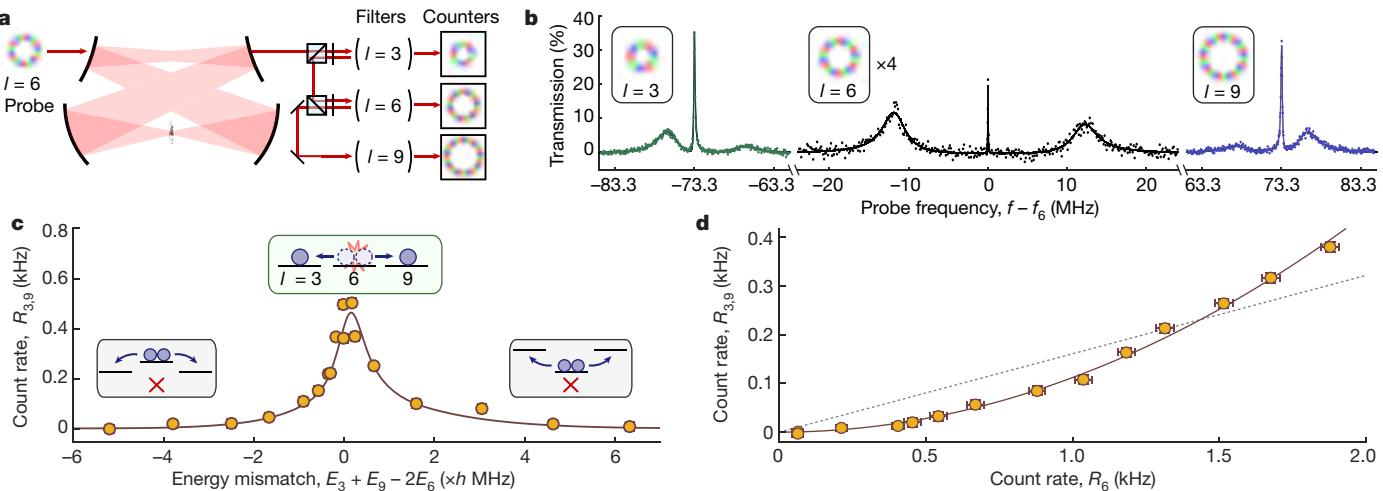
**Fig. 1 | Components for Laughlin states of flight.** **a**, Our experiment couples optical photons (red) with a gas of  $^{87}\text{Rb}$  atoms at the waist of a twisted, four-mirror cavity. **b**, This coupling turns each photon entering the cavity into a polariton, a quasiparticle combining the photon with a collective Rydberg excitation of the atomic gas. Polaritons move around in the transverse modes available to their photonic component, and a pair of polaritons (depicted) strongly interacts because of their Rydberg components (blue spheres). Two key components enable this system to explore topological order. **c**, First, we form a flat topological band of single-photon states using a twisted optical cavity, which hosts a set of degenerate photonic modes that are equivalent to

the states in the lowest Landau level available to electrons in a strong magnetic field. For each mode, the complex electric field  $E$  in the plane perpendicular to photon propagation is shown. Illustrations at the left depict the origins of the fields in each system. **d**, Second, the strong polaritonic interactions mediated by Rydberg atoms ( $R$ ) are analogous to the Coulomb interactions between electrons in a traditional fractional quantum Hall system. Polaritons confined to a single cavity mode reveal their strong interactions via transport blockade, wherein a single polariton present in the cavity prevents a second photon from entering. Blockade results in antibunched correlations of photons exiting the cavity, shown here for  $l=0$ .

A typical experiment begins by loading a pancake of  $6(1) \times 10^3$  laser-cooled  $^{87}\text{Rb}$  atoms at the waist of an optical cavity (Fig. 1a, Methods section ‘Experimental setup’; numbers in parentheses indicate standard error). We then continuously shine a weak probe laser beam on the cavity for 100 ms. Initially uncorrelated photons from the beam entering the cavity are strongly coupled to a resonant atomic transition. With an additional Rydberg electromagnetically induced transparency field, they become polaritons, quasiparticles consisting of a superposition between a photon and a collective Rydberg excitation of the atomic

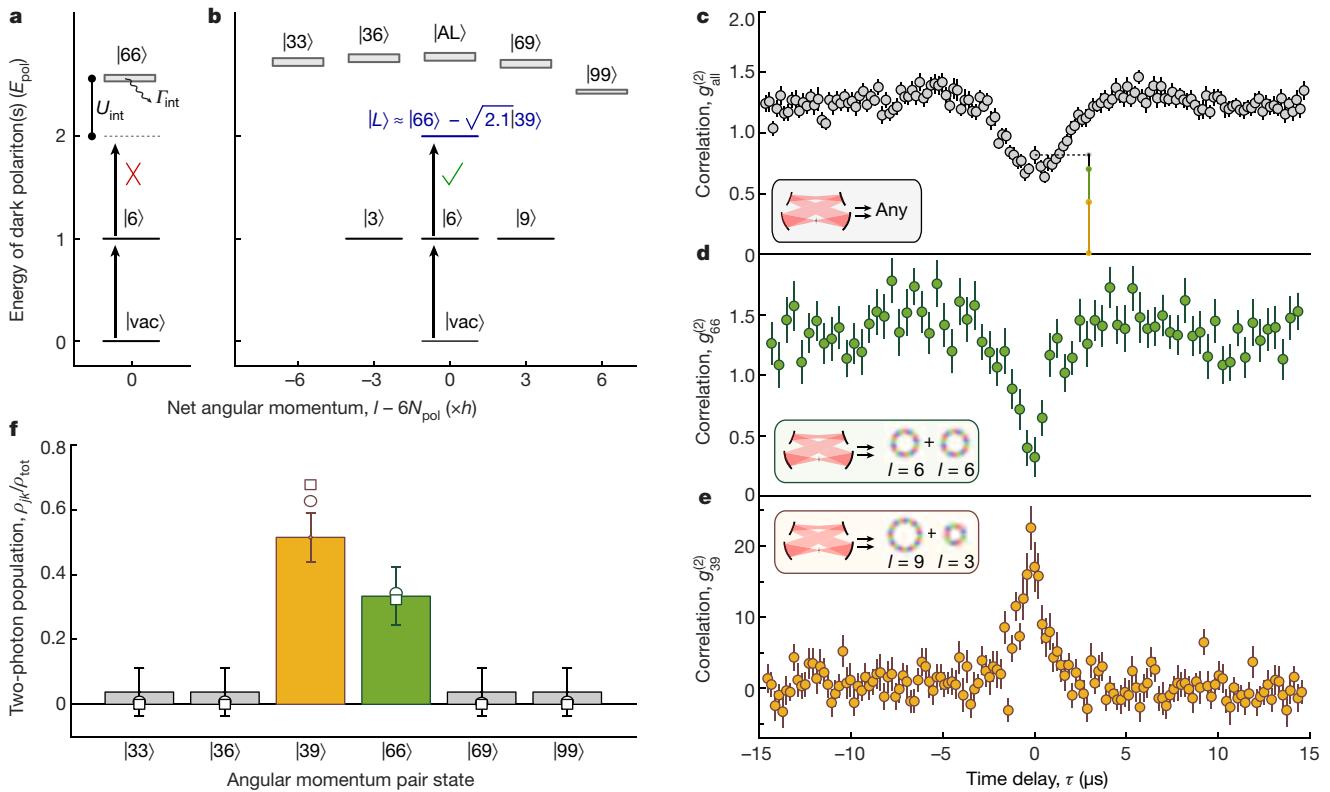
gas (Fig. 1b, Methods section ‘Making polaritons with cavity Rydberg electromagnetically induced transparency’)<sup>16,17</sup>.

Polaritons inherit key properties from both their photonic and Rydberg components. The motion of individual polaritons is determined by the cavity modes accessible to their photonic part. The large energy spacing between longitudinal cavity manifolds restricts the polaritons to a single manifold, confining them to undergo two-dimensional motion among their transverse modes. We utilize a twisted optical cavity and Floquet engineering to form a set of



**Fig. 2 | Collisions between polaritons in the lowest Landau level.** **a**, To test for mode-changing collisions between polaritons, we inject photons with orbital angular momentum  $l=6$  and then count the photons emerging from the cavity in other angular momentum modes (Supplementary Information section A2). **b**, We perform spectroscopy of the single-polariton eigenstates by weakly injecting  $l=3$ ,  $l=6$  and  $l=9$  photons and scanning their energies. We observe Floquet Rydberg polariton resonances<sup>18</sup> (three narrow peaks) for each of the three accessible lowest-Landau-level eigenstates, separated by roughly 73 MHz. The  $l=3$  and  $l=9$  polaritons are more photon-like (broader and taller) than the  $l=6$  polariton (see text). We ignore the broad ‘bright’ polaritons<sup>32</sup>, as they do not interact<sup>17</sup>, instead probing collisions of photons injected on the Rydberg polariton resonances. **c**, When  $l=6$  photons are injected into the

cavity at a higher rate, the total rate  $R_{3,9}$  at which photons emerge with angular momentum  $l=3$  or  $l=9$  peaks when the collisions conserve energy (inset);  $E_i$  is the angular momentum  $l$  polariton’s energy. Atomistic numerics (solid curve) capture key features of the data, including the full-width at half-maximum of 1.1(1) MHz and slight energy mismatch 0.15(5) h MHz<sup>-1</sup> of the peak due to the finite Rydberg interaction range (Supplementary Information section B5). **d**, At zero energy mismatch, varying the photon injection rate reveals that  $R_{3,9}$  grows quadratically (solid curve) rather than linearly (dotted line) with the rate  $R_6$  of injected  $l=6$  photons, consistent with  $l=3$  and  $l=9$  polaritons produced in two-body collisions between  $l=6$  polaritons. The probe power in **c** is the same as the  $R_6 = 1.32(3)$  kHz datapoint in **d**; observed rates are incompatible due to mild  $\nabla E$  broadening of **d**. Error bars indicate the standard error of the mean.



**Fig. 3 | Laughlin state characterization in angular momentum space.** **a, b**, Schematic many-body energy spectrum for states containing polariton numbers  $N_{\text{pol}} \leq 2$  with each particle able to explore only a single mode (**a**) or three modes of the lowest Landau level (**b**). In a single mode (here  $l = 6$ ), the polaritons cannot avoid each other, so the two-polariton state  $|66\rangle$  has its energy shifted ( $U_{\text{int}}$ ) and rapidly decays ( $\Gamma_{\text{int}}$ ) due to interactions; these effects induce blockade. With three modes accessible, most two-polariton states are still shifted by interactions. However, a Laughlin state  $|L\rangle$  arises in which the interference between  $|66\rangle$  and  $|39\rangle$  reduces the interaction energy by keeping the particles apart. This pair state is longer-lived and thus preferentially excited by the  $l = 6$  probe laser (black arrows). In practice, we study and measure correlations in the related but inequivalent optical quantum field emitted from

the cavity. **c**, When all emitted photons are included regardless of angular momentum, photons with access to all three modes exhibit weak blockade. **d**, Correlations between photons with  $l = 6$  exhibit stronger blockade, with remnant non-zero correlation  $g_{66}(0) = 0.32(3)$  at zero delay. **e**, Photons with angular momenta  $l = 3$  and  $l = 9$  exhibit a large positive cross-correlation because they are produced together in collisions. **f**, Relative populations in the two-photon manifold determined from coincidence events (bars) are comparable to the Laughlin state (squares) and an atomistic numerical model<sup>18</sup> (circles). Vertical bars in **c** indicate the portion of  $g_{\text{all}}^{(2)}(0)$  accounted for by pairs in  $|39\rangle$  (orange),  $|66\rangle$  (green) and all other states (black), as detailed in Supplementary Information section A3. All error bars indicate standard error.

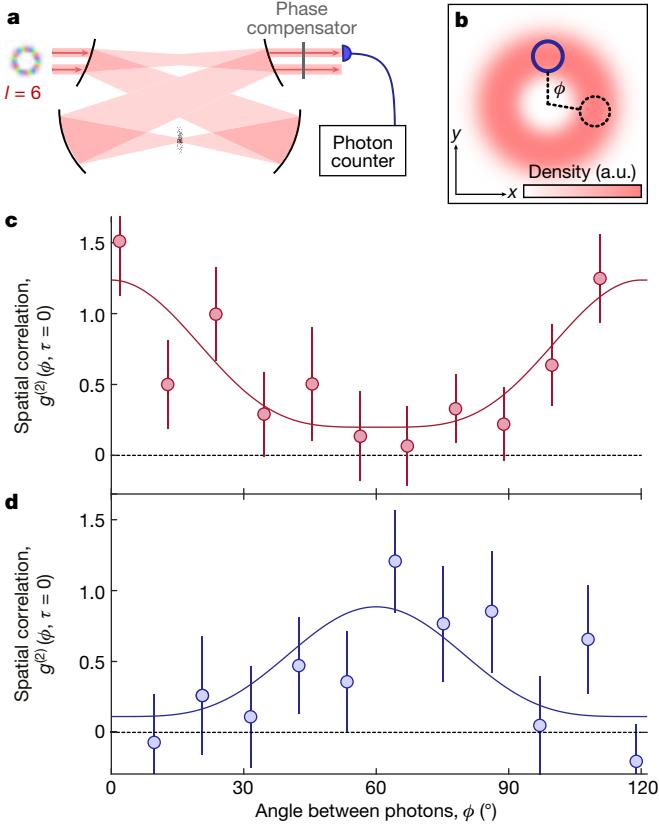
degenerate orbital angular momentum modes equivalent to the lowest Landau level accessible to electrons in a magnetic field<sup>18,19</sup> (Fig. 1c).

To form ordered states, polaritons must also interact with one another (Fig. 1d). Polaritons inherit the strong interactions of their Rydberg components, causing them to avoid each other. When confined to a single transverse mode, one polariton can avoid another only by blockading the cavity, preventing the second polariton from entering, akin to a zero-dimensional electronic quantum dot<sup>17</sup>. This phenomenon can be characterized via the two-photon correlation function  $g^{(2)}(\tau)$  quantifying the likelihood that two photons emerge from the cavity separated by a time  $\tau$  compared with uncorrelated photons (Supplementary Information section A3). Blockade manifests strikingly through antibunching; the  $\tau = 0$  correlation  $g^{(2)}(0)$  falls to zero because there are never two polaritons present in the cavity simultaneously.

When polaritons have access to multiple transverse modes, as in the lowest Landau level, new physics emerges. It becomes possible for two polaritons to enter the cavity simultaneously while avoiding one another: interactions need not lead to blockade, but can instead drive collisions between polaritons, causing them to move among the states of the Landau level and thereby reduce their interaction energy. To test for these collisions, we provide the polaritons with access to exactly three states in the lowest Landau level—those with orbital angular momentum  $l/\hbar = 3, l/\hbar = 6$  and  $l/\hbar = 9$ , where  $\hbar$  is the

reduced Planck constant—using our Floquet scheme<sup>18</sup> (see Methods section ‘Forming the Landau level with Floquet polaritons’). Spectroscopy reveals the narrow Rydberg-polariton eigenstates that we employ for this work, alongside wider, non-interacting ‘bright’ polaritons<sup>32</sup> that we can ignore (Fig. 2b). By using a digital micromirror device for wavefront shaping, we inject photons with angular momentum  $l = 6$ . Photons entering the cavity become polaritons and can then collide and begin to order. When polaritons emerge from the cavity as photons through the output mirror, we count how many photons come from each angular momentum mode (Fig. 2a).

Despite the exotic nature of polaritonic quasiparticles, we find that they undergo collisions much like ordinary particles<sup>22</sup>: collisions between polaritons conserve total energy, as well as angular momentum owing to the rotational symmetry of the Landau level and interactions. Accordingly, the only collision process that conserves angular momentum converts two input polaritons with  $l = 6$  into one output polariton with  $l = 3$  and another with  $l = 9$ . Similarly, as we tune the relative energies between the angular momentum states, we detect only photons with  $l = 3$  or  $l = 9$  when the aforementioned collision process conserves energy (Fig. 2c, Supplementary Information section B6). Moreover, increasing the probe beam power increases the likelihood of two polaritons being present simultaneously and thereby makes collisions more likely, a feature analogous to nonlinear reaction rates in chemistry. Indeed, we observe that the rate at which collision



**Fig. 4 | Spatial correlations of a photonic Laughlin puddle.** **a**, A single-mode fibre (purple) admits only photons at its location, enabling us to filter photons by their real-space position. A mode-dependent phase compensator counteracts the phase difference between polaritons and photons imprinted by our Floquet scheme (Supplementary Information section A3). **b**, The average density of photons forms a smooth annulus with no angular structure. We place the fibre off-centre (solid circle) at the radius with the highest density and measure correlations as a function of the angle  $\phi$  to a second location (dashed circle) around the annulus. **c**, The measured angular correlations (circles) for zero time delay exhibit a periodic structure. The maximum of correlation near  $\phi = 0$  reveals that photons are bunched in space due to a Floquet-induced phase difference between photons and polaritons. **d**, After phase correction, the photons avoid one another. The solid curve is a fit to the function  $g_{\text{fit}}^{(2)}(\phi) = g_0 + g_1 \sin^4(3\phi/2)$  (see text). Error bars indicate standard error.

products appear  $R_{3,9} \propto R_6^2$  is quadratic in the rate  $R_6$  at which photons emerge in the initial angular momentum state (Fig. 2d).

To understand the ordering emerging from these collisions, in Fig. 3a, b we consider an idealized energy spectrum for zero, one and two polaritons. We consider a ‘flat Landau level’ where all single-polariton states  $|l\rangle$  are degenerate, with energy  $E_{\text{pol}}$ . When only one single-particle eigenstate (for example, the  $l=6$  mode) is accessible, the interactions between polaritons that arise from their Rydberg components shift the energy and shorten the lifetime of the state  $|66\rangle$  with two polaritons in that mode. Thus, a probe laser that resonantly excites  $|6\rangle$  from the vacuum state  $|\text{vac}\rangle$  does not subsequently excite the  $|6\rangle \rightarrow |66\rangle$  transition, leading to blockade and precluding the formation of multipolariton states.

In this idealized limit, providing the polaritons access to three single-particle eigenstates in the lowest Landau level leads to the emergence of a long-lived two-particle Laughlin state  $|L\rangle$  (Fig. 3b). Our experiments offer a unique opportunity to connect the mathematical form of this Laughlin state to measurements of its microscopic structure. For the modes used in this work, the two-particle Laughlin wavefunction in real space is  $\psi_L(z_1, z_2) \propto z_1^3 z_2^3 (z_1^3 - z_2^3)^2 \exp(-|z_1|^2/4 - |z_2|^2/4)$ ,

where  $z_k \equiv x_k + iy_k$  is a complex number reflecting the position  $(x_k, y_k)$  of particle  $k$ . Expanding the polynomial prefactor lets us write this state in angular-momentum space as  $|L\rangle = \frac{1}{\sqrt{3.1}}|66\rangle - \sqrt{\frac{2.1}{3.1}}|39\rangle$ , where  $|mn\rangle$  is the state with two polaritons of angular momenta  $m\hbar$  and  $n\hbar$  (Supplementary Information section B4). As the wavefunction goes to zero when the particles are at the same position ( $\psi_L(z_1=z_2)=0$ ), the Laughlin state enables two particles to avoid each other while remaining in the lowest Landau level. In the angular momentum basis, this avoidance arises from destructive interference between the  $|66\rangle$  and  $|39\rangle$  two-particle amplitudes for co-located particles. Similar two-particle Laughlin states can be formed in any set of three evenly spaced angular-momentum states (Supplementary Information section B4).

The spatial anticorrelation of polaritons in the Laughlin state suppresses the interaction energy and interaction-induced decay present in other two-particle states. Thus, for ideal polaritons, simply shining a laser into this atom-cavity system would cause a polaritonic Laughlin state to form inside, because all other two-polariton states are blockaded (including the ‘anti-Laughlin’ superposition state  $|AL\rangle = \sqrt{\frac{2.1}{3.1}}|66\rangle + \frac{1}{\sqrt{3.1}}|39\rangle$ ).

The collisions observed in Fig. 2 are the first hint of such ordering. As detailed in the Methods section ‘Forming the Landau level with Floquet polaritons’, the finite optical depth of our polaritons prevents them from reaching a polaritonic Laughlin state before decaying. Surprisingly, we can distil a high-purity photonic Laughlin state outside of the cavity from the polaritonic state produced within the cavity: this is achieved by choosing Floquet conditions to make the  $l=3$  and  $l=9$  polaritons more photon-like than the  $l=6$  polariton (apparent from their larger width in Fig. 2b), and correcting the phase imparted to the outgoing photons by the Floquet modulation (as detailed in Supplementary Information section A1).

To definitively test for the formation of these photonic Laughlin states, we experimentally investigate the correlations of the emerging photon pairs<sup>9</sup>. First, when we detect all output photons regardless of their spatial mode, the ‘all-mode’ correlations  $g_{\text{all}}^{(2)}(\tau)$  reveal only weak blockade (Fig. 3c). This observation confirms that photon pairs can traverse the three-mode cavity whereas they were blocked in its single-mode counterpart; determining the structure of the pairs requires more detailed measurements.

We gain deeper insight by examining the correlations  $g_{jk}^{(2)}$  between photons with angular momenta  $l=j$  and  $l=k$ , again using the setup shown in Fig. 2a. The correlations  $g_{66}^{(2)}$  between photons with  $l=6$  have a non-zero value  $g_{66}^{(2)}(0) = 0.32(3)$  at zero time delay, indicating a substantial population in the pair state  $|66\rangle$  (Fig. 3d). However, their blockade is still stronger than that of the all-mode correlations  $g_{\text{all}}^{(2)}$ , indicating that the two-photon state has a large contribution from pairs not in  $|66\rangle$ . Most remaining pairs are accounted for by examining  $g_{39}^{(2)}$ , which exhibits a prominent peak at zero time (Fig. 3e). The peak height indicates that photons are  $g_{39}^{(2)}(0) = 22(2)$  times more likely to appear in both modes simultaneously than expected for uncorrelated photons arriving with the same individual rates. This bunching arises because photons in these modes are produced together from polariton collisions and rarely injected independently. Unlike parametric collision processes observed in nonlinear crystals requiring macroscopic mode populations<sup>33</sup>, collisions in our system appear for just two intracavity polaritons.

We quantitatively compare the observed pair-photon state with the Laughlin state by calculating the two-particle populations  $\rho_{jk}$  with angular momenta  $j$  and  $k$  from the coincidence rates of events where two photons are observed near simultaneously (Fig. 3f, Supplementary Information section A3). Coincidence events corresponding to  $|66\rangle$  or  $|39\rangle$  account for 85(15)% of all observed photon pairs, consistent with angular momentum conservation. Moreover, the ratio  $\rho_{39}/\rho_{66} = 1.5(5)$  of pair populations is near the intended ratio of 2.1 for the Laughlin state. Deviations from ideal populations could arise from the limited lifetime and interaction strength of our polaritons, slight drifts of

system parameters between experiments or because the polaritons with angular momentum  $l=3$  and  $l=9$  are insufficiently photon-like.

We next test for the remaining essential physical feature of Laughlin states: that the photons avoid each other in real space. To measure in real space, we filter the output photons with a single-mode optical fibre (Fig. 4a) that only admits photons located at its tip, enabling us to count photons at that position. As the average density in a state composed of  $|39\rangle$  and  $|66\rangle$  forms a smooth annulus, we translate the fibre to the radius with the highest density (Fig. 4b). A natural method for measuring angular correlations  $g^{(2)}(\phi, \tau=0)$  between photons separated by the angle  $\phi$  around the annulus would be to use two fibres at different positions; our Floquet scheme induces a rapid rotation of the optical field, enabling an equivalent measurement with a single fibre.

While the average density of photon pairs leaving the cavity exhibits no angular structure, we find that the photons are bunched in space (Fig. 4c) owing to a Floquet-induced phase between photons and polaritons. After compensating this phase with linear optics (Supplementary Information section A3), the photons avoid each other as expected for a Laughlin state (Fig. 4d). The spatial correlations should take the form  $g_L^{(2)}(\phi) \propto |\psi_1(z_2 = e^{i\phi}z_1)|^2 \sin^4(3\phi/2)$ , which oscillates with the angle  $\phi$  between the photons with a periodicity of  $120^\circ$  because only every third angular momentum state is accessible. Motivated by this expected form, we fit the observed correlations with  $g_{\text{fit}}^{(2)}(\phi) = g_0 + g_1 \sin^4(3\phi/2)$  allowing for an offset  $g_0$  from perfect spatial antibunching and a reduction of the oscillation amplitude  $g_1$  owing to imperfect state fidelity or detection. The fit yields a small offset  $g_0 = 0.11(19)$  and an oscillation amplitude  $g_1 = 0.77(36)$  with a significance of 2.1 standard deviations corresponding to a probability of 0.02 to arise from statistical fluctuations in a system with no dependence of correlation on angle. Combining with the uncompensated data further reduces this probability to  $10^{-5}$  (Supplementary Information section A3).

On their own, the angular-momentum mode populations are suggestive of Laughlin physics but insensitive to the phase and even the purity of the superposition between  $|39\rangle$  and  $|66\rangle$ . However, the observed spatial anticorrelation occurs only for a coherent superposition with a minus sign. Combined, these results indicate that the photon pairs, in a virtual image plane at the twisted cavity mode waist (Supplementary Information section A7), have 76(18)% overlap with a pure Laughlin state. This fidelity is limited primarily by our conservative assumptions about the unmeasured momentum-non-conserving pair populations (Supplementary Information section A4).

This work establishes quantum many-body optics as a critical route to breakthroughs in quantum materials. Feasible technological upgrades, including novel state-preparation schemes, will enable our platform to support polaritonic Laughlin states and eventually explore the statistical phases of anyons (see Methods section ‘The near future of our platform’).

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions

and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2318-5>.

- Anderson, P. W. More is different. *Science* **177**, 393–396 (1972).
- Landau, L. D. & Lifshitz, E. M. *Statistical Physics: Course of Theoretical Physics Vol. 5* (Addison-Wesley, 1958).
- Chen, X., Gu, Z.-C. & Wen, X.-G. Local unitary transformation, long-range quantum entanglement, wave function renormalization, and topological order. *Phys. Rev. B* **82**, 155138 (2010).
- Nayak, C., Simon, S. H., Stern, A., Freedman, M. & Das Sarma, S. Non-Abelian anyons and topological quantum computation. *Rev. Mod. Phys.* **80**, 1083–1159 (2008).
- Laughlin, R. B. Anomalous quantum Hall effect: an incompressible quantum fluid with fractionally charged excitations. *Phys. Rev. Lett.* **50**, 1395–1398 (1983).
- Tsui, D. C., Stormer, H. L. & Gossard, A. C. Two-dimensional magnetotransport in the extreme quantum limit. *Phys. Rev. Lett.* **48**, 1559–1562 (1982).
- Bloch, I., Dalibard, J. & Zwerger, W. Many-body physics with ultracold gases. *Rev. Mod. Phys.* **80**, 885–964 (2008).
- Cooper, N. R., Dalibard, J. & Spielman, I. B. Topological bands for ultracold atoms. *Rev. Mod. Phys.* **91**, 015005 (2019).
- Umucalilar, R., Wouters, M. & Carusotto, I. Probing few-particle Laughlin states of photons via correlation measurements. *Phys. Rev. A* **89**, 023803 (2014).
- Carusotto, I. & Ciuti, C. Quantum fluids of light. *Rev. Mod. Phys.* **85**, 299–366 (2013).
- Ozawa, T. et al. Topological photonics. *Rev. Mod. Phys.* **91**, 015006 (2019).
- Du, X., Skachko, I., Duerr, F., Luican, A. & Andrei, E. Y. Fractional quantum Hall effect and insulating phase of Dirac electrons in graphene. *Nature* **462**, 192–195 (2009).
- Bolotin, K. I., Ghahari, F., Shulman, M. D., Stormer, H. L. & Kim, P. Observation of the fractional quantum Hall effect in graphene. *Nature* **462**, 196–199 (2009); corrigendum 475, 122 (2011).
- Spanton, E. M. et al. Observation of fractional Chern insulators in a van der Waals heterostructure. *Science* **360**, 62–66 (2018).
- Fleischhauer, M., Imamoglu, A. & Marangos, J. Electromagnetically induced transparency: optics in coherent media. *Rev. Mod. Phys.* **77**, 633–673 (2005).
- Peyronel, T. et al. Quantum nonlinear optics with single photons enabled by strongly interacting atoms. *Nature* **488**, 57–60 (2012).
- Jia, N. et al. A strongly interacting polaritonic quantum dot. *Nat. Phys.* **14**, 550–554 (2018).
- Clark, L. W. et al. Interacting Floquet polaritons. *Nature* **571**, 532–536 (2019).
- Schne, N., Ryuu, A., Gromov, A., Sommer, A. & Simon, J. Synthetic Landau levels for photons. *Nature* **534**, 671–675 (2016).
- Birnbaum, K. M. et al. Photon blockade in an optical cavity with one trapped atom. *Nature* **436**, 87–90 (2005).
- Thompson, J. D. et al. Coupling a single trapped atom to a nanoscale optical cavity. *Science* **340**, 1202–1205 (2013).
- Sommer, A., Büchler, H. P. & Simon, J. Quantum crystals and Laughlin droplets of cavity Rydberg polaritons. Preprint at <https://arxiv.org/abs/1506.00341> (2015).
- Hasan, M. Z. & Kane, C. L. Colloquium: topological insulators. *Rev. Mod. Phys.* **82**, 3045–3067 (2010).
- Stern, A. Anyons and the quantum Hall effect—a pedagogical review. *Ann. Phys.* **323**, 204–249 (2008).
- Cooper, N. R. Rapidly rotating atomic gases. *Adv. Phys.* **57**, 539–616 (2008).
- Gemelke, N., Sarajlic, E. & Chu, S. Rotating few-body atomic systems in the fractional quantum Hall regime. Preprint at <https://arxiv.org/abs/1007.2677> (2010).
- Eckardt, A. Colloquium: atomic quantum gases in periodically driven optical lattices. *Rev. Mod. Phys.* **89**, 011004 (2017).
- Tai, M. E. et al. Microscopy of the interacting Harper–Hofstadter model in the two-body limit. *Nature* **546**, 519–523 (2017).
- Wallraff, A. et al. Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics. *Nature* **431**, 162–167 (2004).
- Roushan, P. et al. Chiral ground-state currents of interacting photons in a synthetic magnetic field. *Nat. Phys.* **13**, 146–151 (2017).
- Barik, S. et al. A topological quantum optics interface. *Science* **359**, 666–668 (2018).
- Ningyuan, J. et al. Observation and characterization of cavity Rydberg polaritons. *Phys. Rev. A* **93**, 041802 (2016).
- Pan, J.-W. et al. Multiphoton entanglement and interferometry. *Rev. Mod. Phys.* **84**, 777–838 (2012).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

# Article

## Methods

### The near future of our platform

This work establishes quantum many-body optics in strongly interacting gases of topological polaritons as a critical route to breakthroughs in quantum materials. We apply the unique microscopic control of our photonic platform, providing highly tunable system parameters, energy- and space-resolved particle injection, and the ability to measure correlations in almost any basis using simple linear optics. The present performance, combined with a cavity that can be made precisely degenerate in lieu of the Floquet scheme, should directly support polaritonic Laughlin states (Supplementary Information section A1). Looking ahead, novel state-preparation schemes<sup>34,35</sup> such as dissipative stabilization<sup>36–40</sup> will enable the formation of larger topologically ordered states. Interestingly, the Laughlin state that we have assembled in this work already contains a quasi-hole at its centre (Supplementary Information section B4), a precursor to directly measuring statistical phases of anyons<sup>41–45</sup> or even non-Abelian braiding in the Moore–Read state<sup>46</sup>.

### Making polaritons with cavity Rydberg electromagnetically induced transparency

Coupling atomic gases with multiple modes of optical resonators provides exciting opportunities for studying many-body physics<sup>47–52</sup>. Here we use a multimode optical cavity to generate a synthetic gauge field for light<sup>11,19,53–56</sup>. Making photons interact strongly in this cavity enables us to study strongly correlated fractional quantum Hall states of light<sup>9,57–63</sup>.

In our system, photonic interactions are mediated by Rydberg atoms<sup>64</sup> through a scheme called cavity Rydberg electromagnetically induced transparency<sup>15–17,32,65–74</sup>. First, we couple cavity photons to the  $5S_{1/2} \rightarrow 5P_{3/2}$  transition of the gas of  $^{87}\text{Rb}$  atoms at the waist of our primary cavity, which we name the ‘science’ cavity, with collective coupling strength  $g$  (Extended Data Fig. 1a). These experiments operate in the regime of resonant collective strong coupling between the atoms and the cavity, in which the collective cooperativity  $\eta_l = 4g_l^2/(\kappa\Gamma)$  in the relevant modes, with angular momentum  $l$ , collective atom-cavity coupling  $g_l$ , cavity decay rate  $\kappa$  and atomic excited state decay rate  $\Gamma$ , satisfies  $\eta_l \gg 1$  (ref. <sup>75</sup>). The  $5P_{3/2}$  state is subsequently coupled to a highly excited Rydberg level ( $111D_{5/2}$ ) by an additional 480-nm field with coupling strength  $\Omega$ . As detailed in Methods section ‘Experimental setup’ and in Supplementary Information section A6, we use a buildup cavity crossed with the science cavity to achieve sufficient intensity of that Rydberg coupling field while it covers a large area.

As a result of these couplings, photons no longer propagate in the cavity on their own as they would in vacuum. Instead each cavity mode hosts three flavors of polaritons—quasiparticles composed of hybrids between a photon and a collective excitation of the atomic gas<sup>32</sup> (Extended Data Fig. 1b). The nature of these collective excitations is detailed in Supplementary Information section B1. The two ‘bright’ polaritons are largely composed of a photon and a collective  $5P_{3/2}$  excitation; we do not employ bright polaritons in this work because they are short-lived due to rapid decay of the  $5P_{3/2}$  state at  $\Gamma = 2\pi \times 6$  MHz. We utilize the third ‘dark’ polariton throughout the main text of this work. Dark polaritons are purely a superposition of a cavity photon and a collective Rydberg excitation. Because they do not have any  $5P_{3/2}$  component, they are long-lived. Moreover, dark polaritons inherit strong interactions from their large collective Rydberg component<sup>17,64</sup>.

A typical scan of the polariton features corresponding to a single cavity mode is shown in Extended Data Fig. 1c. That spectrum, taken of the  $l=6$  cavity mode, directly reveals the narrow dark polariton excitation at frequency  $f_6$  with a linewidth of 60 kHz, flanked by the two bright polariton resonances, which are  $\pm\sqrt{g^2 + \Omega^2}$  separated in frequency and have linewidths of 3.7(1) MHz. For the displayed data, the

coupling strengths are  $g = 2\pi \times 13$  MHz and  $\Omega = 2\pi \times 1.3$  MHz and the effective Rydberg state linewidth (including decoherence effects) is observed to be  $\Gamma_R = 2\pi \times 50$  kHz.

### Experimental setup

The primary cavity used in our experiments, which couples with the  $5S_{1/2} \rightarrow 5P_{3/2}$  transition, is the so-called science cavity, which consists of four mirrors in a twisted (non-planar) configuration<sup>19,76</sup>. The cavity finesse is  $\mathcal{F} = 1,950$ , yielding a linewidth of  $\kappa = 2\pi \times 1.4$  MHz. The modes are approximately Laguerre–Gaussian in the lower waist where they intersect with the atomic cloud; the fundamental mode has a lower waist size of 19  $\mu\text{m}$ . These parameters yield a peak coupling strength with a single atom of  $g_{\text{single}} = 2\pi \times 0.58$  MHz, corresponding to a cooperativity of  $\eta = 0.16$  per atom<sup>75</sup>. The science cavity is crossed with a build-up cavity for increasing the intensity of the Rydberg coupling beam at 480 nm; the build-up makes it possible to achieve a sufficient intensity over the wide area spanning the science cavity modes up to  $l = 9$  in which we create polaritons. Note that the polarization eigenmodes of both cavities are circular. For more details on the cavity structure, see Supplementary Information section A6.

The science cavity was designed to be tuned to a length at which every third angular momentum state would be degenerate, forming a photonic Landau level<sup>19</sup>. However, as detailed in Supplementary Information section A6, we found that cubic intracavity aberrations arising from astigmatism destabilized the cavity modes when the length was tuned to this degeneracy point. Therefore, we instead operate at a cavity length far enough from degeneracy to create a 70-MHz splitting between every third angular momentum state and utilize the Floquet scheme described below in the next section to create degenerate polaritons that are protected from the effects of intracavity aberrations (Supplementary Information section B2).

We mediate interactions between the photons in the science cavity modes using a gas of cold  $^{87}\text{Rb}$  atoms loaded at the lower waist of the cavity. There are  $6(1) \times 10^3$  atoms distributed nearly uniformly over the region spanned by the first ten modes of the science cavity. The gas is cooled to a temperature below 1  $\mu\text{K}$  and polarized into the lowest-energy spin state ( $|F=2, m_F=-2\rangle$ ) with spin projection  $m_F$  within the hyperfine manifold  $F$  using degenerate Raman sideband cooling<sup>77</sup>. For more details on the atom trapping configuration, atomic polarization and experiment procedure, see Supplementary Information section A5.

### Forming the Landau level with Floquet polaritons

This section explains our scheme for forming the Landau level with Floquet polaritons, which provides a number of crucial features that made this work possible. Most importantly, Floquet polaritons helped us mitigate the effects of intracavity aberrations. As detailed in Supplementary Information section A6, aberrations prevented us from making the bare cavity modes with angular momenta  $l=3, l=6$  and  $l=9$  degenerate. At the cavity length that would otherwise make them degenerate, aberrations mix the modes together, splitting them apart in energy and causing them to decay rapidly. Thankfully, Floquet polaritons are protected from the broadening caused by intracavity aberrations, as detailed in Supplementary Information section B2. The Floquet scheme also provides the ability to efficiently measure angular correlation functions as detailed in Supplementary Information section A3 and enables frequency discrimination to improve the selectivity of our mode sorters discussed in Supplementary Information section A2.

The essence of our Floquet scheme is depicted in Extended Data Fig. 2 and detailed in ref. <sup>18</sup>. Briefly, a sinusoidal modulation of the energy of the intermediate  $5P_{3/2}$  state splits it into three bands that exist at energies separated by the modulation frequency  $f_{\text{mod}}$ . Cavity photons can excite the atom through any of these bands, thereby enabling the atoms to couple with modes whose energies are also split by  $f_{\text{mod}}$ .

To implement the Floquet scheme, we first used temperature and piezo tuning to increase the length of the science cavity and induce a

splitting of  $f_{\text{cav}} \approx 70$  MHz between every third angular momentum mode (Extended Data Fig. 3a). Then, to enable the  $l = 3$ ,  $l = 6$  and  $l = 9$  modes to simultaneously support dark polaritons, we fine-tuned the cavity length to make the  $l = 6$  mode degenerate with the bare  $5S_{1/2} \rightarrow 5P_{3/2}$  transition and modulated the intermediate state at  $f_{\text{mod}} \approx 70$  MHz to create sidebands resonant with the  $l = 3$  and  $l = 9$  states (Extended Data Fig. 3b). By making it possible for the atoms to strongly couple with each of the cavity modes, the Floquet scheme enables all three modes to host polaritons (Extended Data Fig. 3c). Note that the splitting  $f_{\text{cav}}$ , and therefore the required modulation frequency  $f_{\text{mod}}$ , varied over a couple of megahertz from day to day due to variation in the cavity temperature.

Cavity spectroscopy reveals the polariton eigenmodes in all three modes simultaneously (Extended Data Fig. 4a). Because the cavity modes are separated by  $f_{\text{cav}} \approx 70$  MHz, the photonic components of the polaritons also remain separated by approximately  $f_{\text{cav}}$ . Therefore, to spectroscopically characterize the set of polaritons for mode  $l$ , we scan the probe frequency near the frequency  $f_l$  of that mode and simultaneously use a digital micromirror device to spatially mode match the probe laser with the cavity mode<sup>78</sup>, producing the spectra shown in Extended Data Fig. 4a. The asymmetric heights and frequency splittings of the bright polaritons relative to the dark polariton for  $l = 3$  and  $l = 9$  are caused by shifts due to coupling with off-resonant bands of the  $5P_{3/2}$  state. Moreover, the smaller splitting of the bright polaritons with  $l = 3$  and  $l = 9$  (the ‘sideband’ modes) relative to  $l = 6$  (the ‘carrier’ mode) results from smaller atom-cavity couplings  $g_3 \approx g_9 = 0.35g_6$  on those modes relative to the central mode. The relative strengths of these couplings are determined by the modulation amplitude; here, we chose to make the sideband couplings weaker to make the corresponding dark polaritons more photon-like (see Supplementary Information section A1 for details).

While the spectroscopic features are all clearly separated in the scan of the probe frequency  $f$ , the dark polaritons can be tuned to have equal quasifrequency  $\tilde{f}$ , where  $\tilde{f} = (f \bmod f_{\text{mod}})$  is the probe frequency modulo the modulation frequency. In a Floquet model, the energy is only defined up to multiples of the modulation energy; thus, two states with the same quasienergy (equivalently, quasifrequency) behave as though they are degenerate<sup>27</sup>. In the example shown in Extended Data Fig. 4a, the dark polaritons are separated in probe frequency by an amount equal to the modulation frequency  $f_{\text{mod}} = 73.3$  MHz; thus, their quasifrequencies are identical.

It is not trivial to find conditions at which the quasifrequencies of the dark polaritons are equal. The primary challenges are the anharmonicity in the cavity mode spectrum and the off-resonant shifts caused by weak couplings to non-resonant Floquet bands. These effects typically prevent the dark polariton energies from matching under conditions that might naively seem suitable; in particular, when  $f_{\text{cav}} = f_{\text{mod}}$ , the quasifrequency  $\tilde{f}_9$  ( $\tilde{f}_3$ ) of dark polaritons with  $l = 9$  ( $l = 3$ ) will typically be too large (small) due to the off-resonant shifts<sup>18</sup>.

The dark polariton quasifrequency  $\tilde{f}_l = \tilde{\delta}_c^l \cos^2 \theta_l + \tilde{\delta}_r \sin^2 \theta_l$  in each mode is a weighted average of the cavity mode quasifrequency  $\tilde{\delta}_c^l$  with the Rydberg quasifrequency  $\tilde{\delta}_r$ ; the weight is determined by the dark-state rotation angle  $\theta_l$ , which satisfies  $\tan \theta_l = g_r / \Omega$  (see Supplementary Information section B5 in ref. <sup>18</sup>). A smaller ratio  $g_r / \Omega$  increases the contribution from the cavity photon and thus makes the polariton more ‘photon-like’; in the opposite case, the polariton is more ‘Rydberg-like’<sup>32</sup>.

We adjust the Rydberg quasifrequency  $\tilde{\delta}_r$  and the modulation frequency  $f_{\text{mod}}$  to tune the dark polaritons into degeneracy. The Rydberg quasifrequency is controlled by the frequency of the Rydberg coupling laser. Because the  $l = 6$  polaritons are more Rydberg-like than the  $l = 3$  and  $l = 9$  polaritons, the quasifrequency  $\tilde{f}_6$  of  $l = 6$  polaritons increases more rapidly than the others with  $\tilde{\delta}_r$ . Thus, the quasifrequencies  $\tilde{f}_9$  and  $\tilde{f}_3$  decrease relative to  $\tilde{f}_6$  as  $\tilde{\delta}_r$  increases, as depicted in Extended Data

Fig. 4b. Note that, in Fig. 2c, we varied the Rydberg quasifrequency  $\tilde{\delta}_r$  to scan the energy mismatch. Adjusting the modulation frequency  $f_{\text{mod}}$  shifts the cavity mode quasifrequencies  $\tilde{\delta}_c^3$  and  $\tilde{\delta}_c^9$  oppositely, because they couple to the atoms through opposite modulation sidebands of the  $5P_{3/2}$  state (Extended Data Fig. 3c).

These two adjustment parameters are sufficient to tune the three dark polariton features into degeneracy, as demonstrated in Extended Data Fig. 4d. In practice, to reach degeneracy, we repeatedly measure the quasifrequency spectrum of the dark polaritons while adjusting parameters. We modify the frequency of the Rydberg laser until the average of the sideband quasifrequencies matches the carrier, that is,  $(\tilde{f}_3 + \tilde{f}_9)/2 = \tilde{f}_6$ . Similarly, we vary  $f_{\text{mod}}$  until  $\tilde{f}_3 = \tilde{f}_9$ . In the end, we are able to satisfy these conditions to an accuracy well within the linewidths of the dark polaritons.

## Data availability

The experimental data presented in this manuscript are available from the corresponding author upon request.

34. Grusdt, F., Letscher, F., Hafezi, M. & Fleischhauer, M. Topological growing of Laughlin states in synthetic gauge fields. *Phys. Rev. Lett.* **113**, 155301 (2014).
35. Ivanov, P. A., Letscher, F., Simon, J. & Fleischhauer, M. Adiabatic flux insertion and growing of Laughlin states of cavity Rydberg polaritons. *Phys. Rev. A* **98**, 013847 (2018).
36. Kapit, E., Hafezi, M. & Simon, S. H. Induced self-stabilization in fractional quantum Hall states of light. *Phys. Rev. X* **4**, 031039 (2014).
37. Hafezi, M., Adhikari, P. & Taylor, J. Chemical potential for light by parametric coupling. *Phys. Rev. B* **92**, 174305 (2015).
38. Umucalilar, R. & Carusotto, I. Generation and spectroscopic signatures of a fractional quantum Hall liquid of photons in an incoherently pumped optical cavity. *Phys. Rev. A* **96**, 053808 (2017).
39. Biella, A. et al. Phase diagram of incoherently driven strongly correlated photonic lattices. *Phys. Rev. A* **96**, 023839 (2017).
40. Ma, R. et al. A dissipatively stabilized Mott insulator of photons. *Nature* **566**, 51–57 (2019); correction **570**, E52 (2019).
41. Paredes, B., Fedichev, P., Cirac, J. & Zoller, P. 1/2-Anyons in small atomic Bose–Einstein condensates. *Phys. Rev. Lett.* **87**, 010402 (2001).
42. Umucalilar, R. & Carusotto, I. Many-body braiding phases in a rotating strongly correlated photon gas. *Phys. Lett. A* **377**, 2074–2078 (2013).
43. Grusdt, F., Yao, N. Y., Abanin, D., Fleischhauer, M. & Demler, E. Interferometric measurements of many-body topological invariants using mobile impurities. *Nat. Commun.* **7**, 11994 (2016).
44. Dutta, S. & Mueller, E. J. Coherent generation of photonic fractional quantum Hall states in a cavity and the search for anyonic quasiparticles. *Phys. Rev. A* **97**, 033825 (2018).
45. Macaluso, E., Comparin, T., Mazza, L. & Carusotto, I. Fusion channels of non-Abelian anyons from angular-momentum and density-profile measurements. *Phys. Rev. Lett.* **123**, 266801 (2019).
46. Regnault, N. & Jolicoeur, T. Quantum Hall fractions for spinless bosons. *Phys. Rev. B* **69**, 235309 (2004).
47. Gopalakrishnan, S., Lev, B. L. & Goldbart, P. M. Emergent crystallinity and frustration with Bose–Einstein condensates in multimode cavities. *Nat. Phys.* **5**, 845–850 (2009).
48. Wickenbrock, A., Hemmerling, M., Robb, G. R., Emary, C. & Renzoni, F. Collective strong coupling in multimode cavity QED. *Phys. Rev. A* **87**, 043817 (2013).
49. Ritsch, H., Domokos, P., Brennecke, F. & Esslinger, T. Cold atoms in cavity-generated dynamical optical potentials. *Rev. Mod. Phys.* **85**, 553–601 (2013).
50. Douglas, J. S. et al. Quantum many-body models with cold atoms coupled to photonic crystals. *Nat. Photon.* **9**, 326–331 (2015).
51. 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).
52. Vaidya, V. D. et al. Tunable-range, photon-mediated atomic interactions in multimode cavity QED. *Phys. Rev. X* **8**, 011002 (2018).
53. Hafezi, M., Mittal, S., Fan, J., Miggall, A. & Taylor, J. Imaging topological edge states in silicon photonics. *Nat. Photon.* **7**, 1001–1005 (2013).
54. Rechtsman, M. C. et al. Photonic Floquet topological insulators. *Nature* **496**, 196–200 (2013).
55. Lim, H.-T., Togan, E., Krone, M., Miguel-Sánchez, J. & Imamoglu, A. Electrically tunable artificial gauge potential for polaritons. *Nat. Commun.* **8**, 14540 (2017).
56. Schine, N., Chalupnik, M., Can, T., Gromov, A. & Simon, J. Electromagnetic and gravitational responses of photonic Landau levels. *Nature* **565**, 173–179 (2019).
57. Hartmann, M. J., Brando, F. G. & Plenio, M. B. Strongly interacting polaritons in coupled arrays of cavities. *Nat. Phys.* **2**, 849–855 (2006).
58. Greentree, A. D., Tahon, C., Cole, J. H. & Hollenberg, L. C. Quantum phase transitions of light. *Nat. Phys.* **2**, 856–861 (2006).
59. Angelakis, D. G., Santos, M. F. & Bose, S. Photon-blockade-induced Mott transitions and XY spin models in coupled cavity arrays. *Phys. Rev. A* **76**, 031805 (2007).
60. Cho, J., Angelakis, D. G. & Bose, S. Fractional quantum Hall state in coupled cavities. *Phys. Rev. Lett.* **101**, 246809 (2008).

# Article

61. Nunnenkamp, A., Koch, J. & Girvin, S. Synthetic gauge fields and homodyne transmission in Jaynes–Cummings lattices. *New J. Phys.* **13**, 095008 (2011).
62. Hayward, A. L., Martin, A. M. & Greentree, A. D. Fractional quantum Hall physics in Jaynes–Cummings–Hubbard lattices. *Phys. Rev. Lett.* **108**, 223602 (2012).
63. Hafezi, M., Lukin, M. D. & Taylor, J. M. Non-equilibrium fractional quantum Hall state of light. *New J. Phys.* **15**, 063001 (2013).
64. Saffman, M., Walker, T. G. & Mølmer, K. Quantum information with Rydberg atoms. *Rev. Mod. Phys.* **82**, 2313–2363 (2010).
65. Fleischhauer, M. & Lukin, M. D. Dark-state polaritons in electromagnetically induced transparency. *Phys. Rev. Lett.* **84**, 5094–5097 (2000).
66. Mohapatra, A., Jackson, T. & Adams, C. Coherent optical detection of highly excited Rydberg states using electromagnetically induced transparency. *Phys. Rev. Lett.* **98**, 113003 (2007).
67. Pritchard, J. D. et al. Cooperative atom-light interaction in a blockaded Rydberg ensemble. *Phys. Rev. Lett.* **105**, 193603 (2010).
68. Guerlin, C., Brion, E., Esslinger, T. & Mølmer, K. Cavity quantum electrodynamics with a Rydberg-blocked atomic ensemble. *Phys. Rev. A* **82**, 053832 (2010).
69. Gorshkov, A. V., Otterbach, J., Fleischhauer, M., Pohl, T. & Lukin, M. D. Photon–photon interactions via Rydberg blockade. *Phys. Rev. Lett.* **107**, 133602 (2011).
70. Dudin, Y. O. & Kuzmich, A. Strongly interacting Rydberg excitations of a cold atomic gas. *Science* **336**, 887–889 (2012).
71. Tiarks, D., Baur, S., Schneider, K., Dürr, S. & Rempe, G. Single-photon transistor using a Förster resonance. *Phys. Rev. Lett.* **113**, 053602 (2014).
72. Gorniakzyk, H., Tresp, C., Schmidt, J., Fedder, H. & Hofferberth, S. Single-photon transistor mediated by interstate Rydberg interactions. *Phys. Rev. Lett.* **113**, 053601 (2014).
73. Boddeda, R. et al. Rydberg-induced optical nonlinearities from a cold atomic ensemble trapped inside a cavity. *J. Phys. B* **49**, 084005 (2016).
74. Georgakopoulos, A., Sommer, A. & Simon, J. Theory of interacting cavity Rydberg polaritons. *Quantum Sci. Technol.* **4**, 014005 (2018).
75. Tanji-Suzuki, H. et al. in *Advances in Atomic, Molecular, and Optical Physics* Vol. 60 (eds Arimondo, E. et al.) 201–237 (Elsevier, 2011).
76. Sommer, A. & Simon, J. Engineering photonic Floquet Hamiltonians through Fabry–Pérot resonators. *New J. Phys.* **18**, 035008 (2016).
77. Kerman, A. J., Vučetić, V., Chin, C. & Chu, S. Beyond optical molasses: 3D Raman sideband cooling of atomic cesium to high phase-space density. *Phys. Rev. Lett.* **84**, 439–442 (2000).
78. Zupancic, P. et al. Ultra-precise holographic beam shaping for microscopic quantum control. *Opt. Express* **24**, 13881–13893 (2016).

**Acknowledgements** We thank L. Feng and M. Jaffe for feedback on the manuscript. This work was supported by AFOSR grant FA9550-18-1-0317 and AFOSR MURI grant FA9550-16-1-0323. N.S. acknowledges support from the University of Chicago Grainger graduate fellowship and C.B. acknowledges support from the NSF GRFP.

**Author contributions** The experiment was designed and built by all authors. N.S. built the primary cavity. L.W.C., N.S. and C.B. collected the data. L.W.C. and N.S. analysed the data. L.W.C. wrote, and all authors contributed to, the manuscript.

**Competing interests** The authors declare no competing interests.

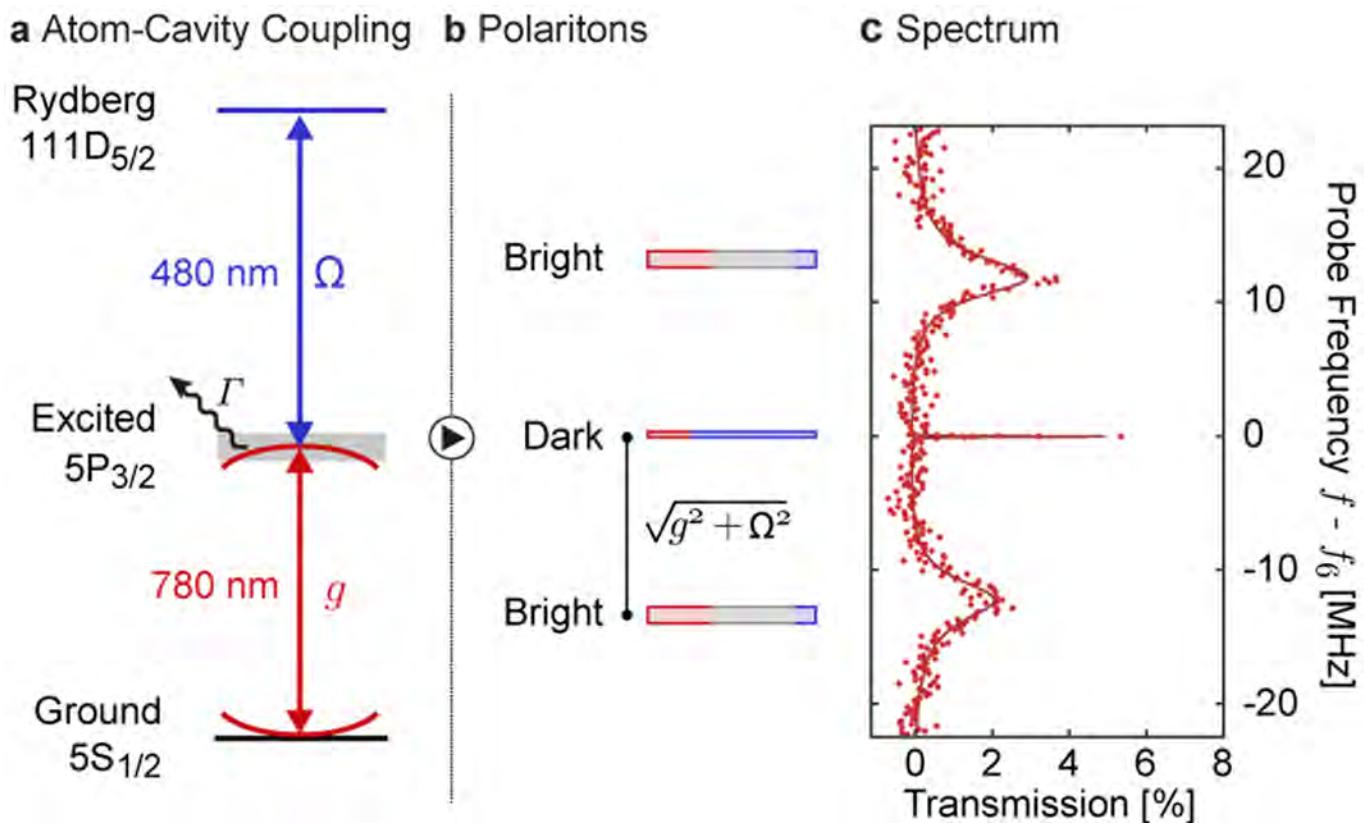
## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2318-5>.

**Correspondence and requests for materials** should be addressed to J.S.

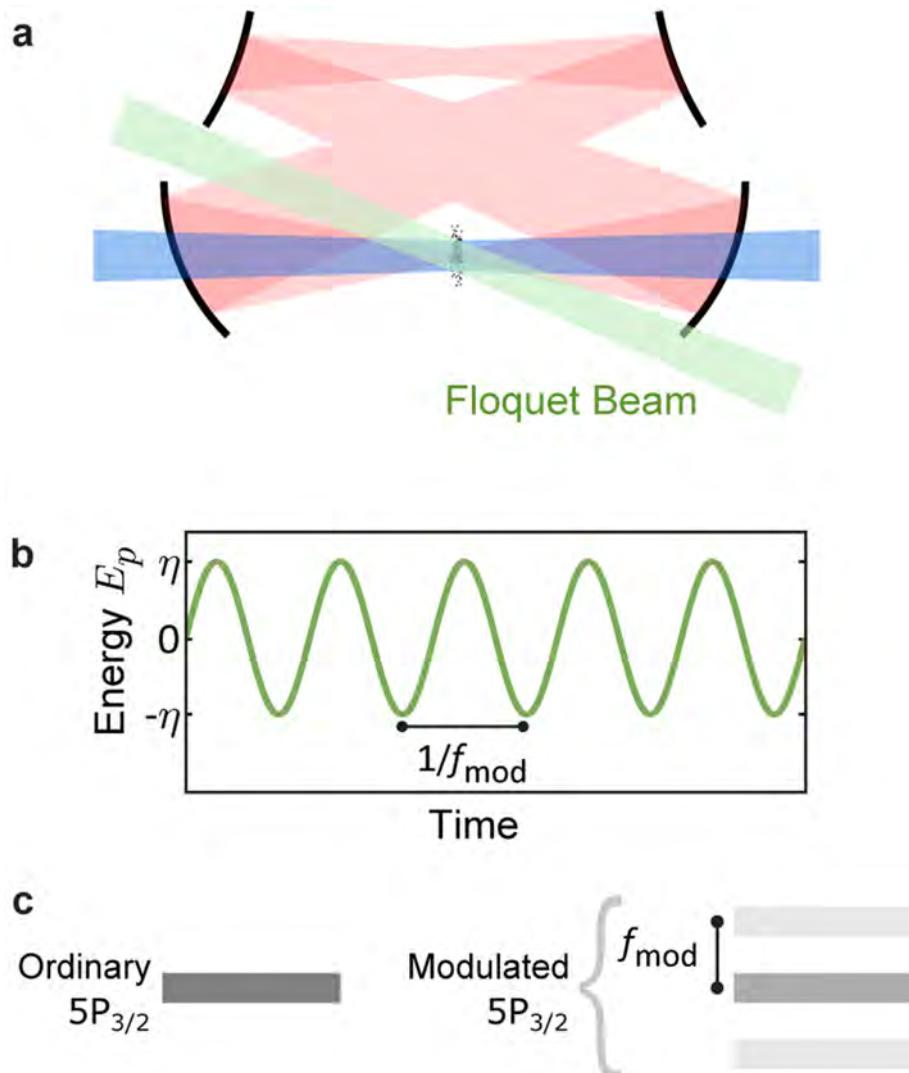
**Peer review information** *Nature* thanks Laura Corman, Oliver Morsch and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



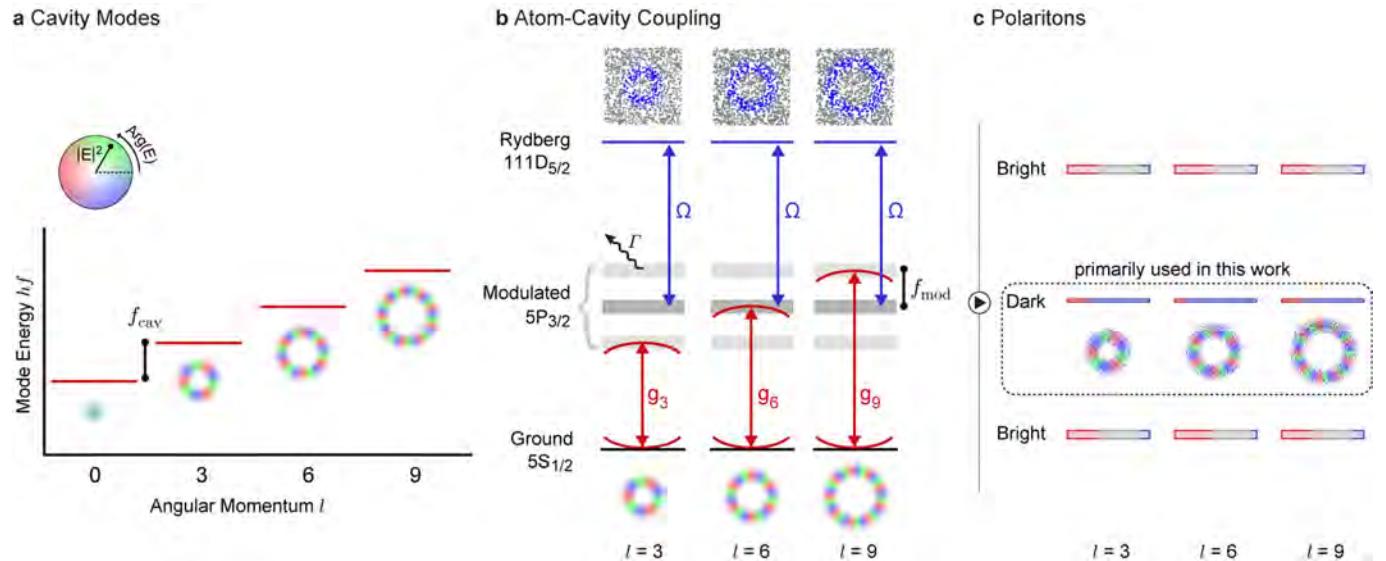
**Extended Data Fig. 1 | Single-mode polariton spectrum.** **a**, Photons at the waist of our twisted optical cavity (red) couple with strength  $g$  to the  $5S_{1/2} \rightarrow 5P_{3/2}$  transition of a gas of cold  $^{87}\text{Rb}$  atoms, which is subsequently coupled with strength  $\Omega$  to the highly excited  $111\text{D}_{5/2}$  Rydberg state using an additional laser (blue). **b**, These couplings cause excitations of the atom-cavity system to propagate as polaritons—quasiparticles combining photons with collective

atomic excitations. **c**, The transmission spectrum of the cavity with atoms present directly reveals the narrow dark polariton flanked by two broad bright polariton peaks. The solid curve shows a fit of the cavity electromagnetically induced transparency spectrum to the measured transmission (see ref.<sup>32</sup>, Supplementary equation (7)).



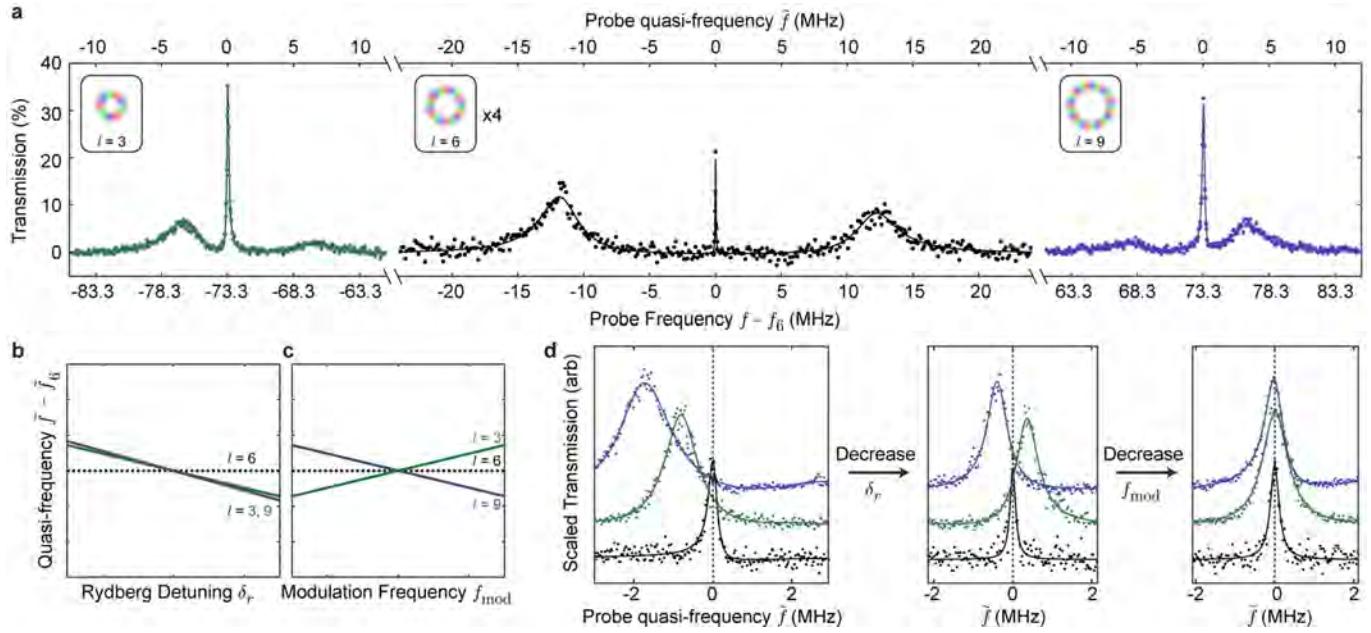
**Extended Data Fig. 2 | Essential features of the Floquet scheme.** **a**, Our Floquet scheme utilizes an additional laser beam (green) incident on the atoms with a wavelength of  $\lambda = 1,529$  nm close to the  $5P_{3/2} \rightarrow 4D$  transition. **b**, This beam induces a sinusoidally modulated a.c. Stark shift  $E_p = \eta \sin(2\pi f_{\text{mod}} t)$  of the  $5P_{3/2}$  state with amplitude  $\eta$  and frequency  $f_{\text{mod}}$ . **c**, As a result of this modulation, the

ordinary  $5P_{3/2}$  state is split into three bands with energies separated by the modulation frequency. The additional bands enable the atoms to couple with cavity photons at frequencies shifted by  $\pm f_{\text{mod}}$  from the ordinary  $5S_{1/2} \rightarrow 5P_{3/2}$  resonance frequency. For more details on the Floquet scheme see ref.<sup>18</sup>.



**Extended Data Fig. 3 | Scheme for forming the Landau level of Floquet polaritons.** **a**, The bare cavity modes are not degenerate in this work, but instead the length of the cavity is increased so that there is a  $f_{\text{cav}} \approx 70$  MHz splitting between every third angular momentum mode. **b**, To form polaritons in three modes, even though only the  $l = 6$  mode is resonant with the un-modulated  $5S_{1/2} \rightarrow 5P_{3/2}$  transition, we utilize the Floquet scheme depicted in Extended Data Fig. 2<sup>18</sup>. Modulating the  $5P_{3/2}$  state at  $f_{\text{mod}} \approx 70$  MHz splits it into three bands (grey), each of which is resonant with one of the three chosen cavity modes. The coupling strengths  $g_l$  to each mode  $l$  are controlled by the

modulation amplitude; in this work,  $g_3 = g_9 = 0.37(4)g_6$ . Note that each mode couples to a unique collective atomic excitation, as depicted at the top (blue atoms are included in the corresponding collective excitation, while grey atoms are not). **c**, This scheme produces polaritons in the  $l = 3, l = 6$  and  $l = 9$  modes. The dark polaritons can be made effectively degenerate (see Extended Data Fig. 4) without making the corresponding cavity modes degenerate, which protects the polaritons from intracavity aberrations (see Supplementary Information section B2).



**Extended Data Fig. 4 | Understanding and controlling polariton spectra with the Floquet scheme.** **a**, Cavity transmission spectrum in the presence of the modulated atoms (see Extended Data Fig. 3), reproducing Fig. 2b. The spectrum was collected in three parts, corresponding to injection of photons into  $l=3$  (left, green),  $l=6$  (middle, black) and  $l=9$  (right, violet). Dark polaritons in the  $l=6$  mode are less photon-like than those in the other two modes, reducing their relative transmission (Supplementary Information section A1); we multiply the  $l=6$  transmission by four to improve visibility. The lower x axis indicates the frequency of the probe laser relative to the  $l=6$  dark polariton resonance at  $f_6$ . The top x axis indicates the quasifrequency  $\tilde{f}$ , proportional to the quasienergy of the polaritons from a treatment using Floquet theory;  $\tilde{f}$  is equal to  $f$  modulo the modulation frequency  $f_{\text{mod}}$ . The solid curves show three independent fits of the cavity electromagnetically induced transparency spectrum to the measured transmission in each angular momentum mode (see ref.<sup>32</sup>, Supplementary equation (7)). **b, c**, Illustration of

the theoretical dependence of the quasifrequencies of the three dark polariton features on the Rydberg beam detuning (**b**) and the modulation frequency  $f_{\text{mod}}$  (**c**). **d**, Example transmission spectra for the  $l=6$  (black, lower),  $l=3$  (green, middle) and  $l=9$  (violet, upper) dark polaritons as a function of quasifrequency. The scans are scaled to make their heights equal and have additional vertical offsets for clarity. Shortly before performing each of the experiments reported in the main text, we collect a sequence of plots similar to those displayed here and adjust the Rydberg detuning and modulation frequency to make all three dark polaritons have the same quasifrequency (right-most plot). The only experiments reported in the main text that did not use this sequence are those shown in Fig. 2c, where instead we varied  $\delta_r$  to intentionally vary the energy mismatch between the polaritons. Throughout this figure, quasifrequencies  $\tilde{f}$  are reported relative to the  $l=6$  dark polariton resonance  $\tilde{f}_6$ . Solid curves provide a guide to the eye.

# Controlling free electrons with optical whispering-gallery modes

<https://doi.org/10.1038/s41586-020-2320-y>

Received: 8 October 2019

Accepted: 6 March 2020

Published online: 3 June 2020

 Check for updates

Ofer Kfir<sup>1</sup>✉, Hugo Lourenço-Martins<sup>1</sup>, Gero Storeck<sup>1</sup>, Murat Sivis<sup>1</sup>, Tyler R. Harvey<sup>1</sup>, Tobias J. Kippenberg<sup>2</sup>, Armin Feist<sup>1</sup> & Claus Ropers<sup>1,3</sup>✉

Free-electron beams are versatile probes of microscopic structure and composition<sup>1,2</sup>, and have revolutionized atomic-scale imaging in several fields, from solid-state physics to structural biology<sup>3</sup>. Over the past decade, the manipulation and interaction of electrons with optical fields have enabled considerable progress in imaging methods<sup>4</sup>, near-field electron acceleration<sup>5,6</sup>, and four-dimensional microscopy techniques with high temporal and spatial resolution<sup>7</sup>. However, electron beams typically couple only weakly to optical excitations, and emerging applications in electron control and sensing<sup>8–11</sup> require large enhancements using tailored fields and interactions. Here we couple a free-electron beam to a travelling-wave resonant cavity mode. The enhanced interaction with the optical whispering-gallery modes of dielectric microresonators induces a strong phase modulation on co-propagating electrons, which leads to a spectral broadening of 700 electronvolts, corresponding to the absorption and emission of hundreds of photons. By mapping the near-field interaction with ultrashort electron pulses in space and time, we trace the lifetime of the microresonator following a femtosecond excitation and observe the spectral response of the cavity. The natural matching of free electrons to these quintessential optical modes could enable the application of integrated photonics technology in electron microscopy, with broad implications for attosecond structuring, probing quantum emitters and possible electron–light entanglement.

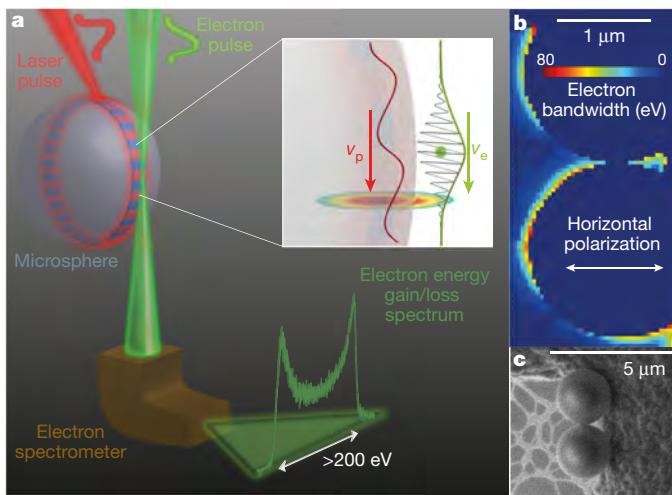
The interaction of free electrons with optical excitations forms the basis of microscopy techniques, yielding fundamental insights into optical properties at the nanoscale. One of the most widespread methods, electron energy loss spectroscopy (EELS), is extensively used to analyse resonant nanostructures<sup>12,13</sup> and enables the measurement of the local photonic density of states<sup>14–18</sup> with nanometre precision. The acquisition of laser-driven electron energy gain spectra (EEGS) yields spatial and spectral mode maps in nanoplasmonics<sup>19–22</sup> and photonic crystals<sup>23</sup>.

Inelastic scattering of electrons and optical excitations of nanostructures naturally cause changes to the free-electron state. In the case of stimulated interactions, the spatial and temporal phase information of a light field is imprinted onto the electronic wavefunction passing through the optical field. Control over the wavefunction of free-electron beams was recently demonstrated in spatial<sup>11</sup>, temporal<sup>24</sup> and spatiotemporal<sup>25,26</sup> forms. One such example, the generation of electron pulses of attosecond-scale duration<sup>27–29</sup>, could combine sub-optical-cycle temporal resolution with the high spatial resolution of transmission electron microscopes (TEMs). The above endeavours harness the enhancement of photon-induced near-field electron microscopy<sup>4,30</sup> (PINEM) resulting from the large polarizability of plasmons. Nonetheless, a limiting factor on the strength of such couplings is typically the sub-optical cycle interaction time of swift electrons across spatially localized fields<sup>8,12</sup>. In analogy to other nonlinear processes,

the coupling can be enhanced by coherently accumulating transition amplitude in more extended structures<sup>30,31</sup>, which requires phase matching, that is, equating the electron group velocity with the optical phase velocity<sup>8,16,32</sup>. Nearly relativistic electrons with energies of around 100–200 keV are naturally phase-matched to travelling waves at visible frequencies in ordinary dielectrics, such as fused silica, which was previously used in a prism geometry for electron acceleration and coherent spectral broadening<sup>33,34</sup>. However, to achieve a stronger coupling between photons and free electrons, which will ultimately lead to substantial entanglement<sup>9,10</sup>, the combination of travelling-wave phase matching with a high density of photonic states is necessary.

Here, we address this challenge and harness whispering-gallery mode (WGM) microcavities to enhance the interaction between light and free electrons. Like their acoustic analogue, optical WGMs are travelling electromagnetic waves confined in dielectric resonators by total internal reflection, and achieve the highest finesse of any optical resonator. The strong field enhancements in WGM resonators leads to multiple applications, such as the manipulation of mechanical modes in cavity optomechanics through radiation pressure<sup>35</sup>, the generation of soliton frequency combs<sup>36</sup> via the third-order nonlinearity of the material, and near-field sensing<sup>37</sup> by measuring dispersive frequency shifts<sup>38</sup>. The modal extension into free space, which is instrumental for atomic and molecular sensing<sup>37,39,40</sup>, motivates the use of these structures for coupling to free electrons.

<sup>1</sup>University of Göttingen, IV Physical Institute, Göttingen, Germany. <sup>2</sup>Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland. <sup>3</sup>Max Planck Institute for Biophysical Chemistry (MPIBPC), Göttingen, Germany. ✉e-mail: ofer.kfir@phys.uni-goettingen.de; claus.ropers@uni-goettingen.de



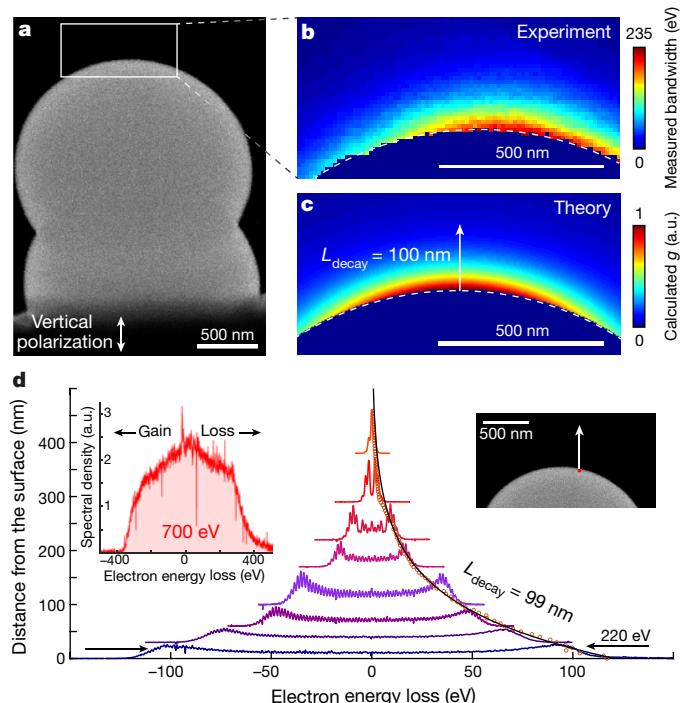
**Fig. 1 | Experimental scheme.** **a**, A laser-driven WGM circulating a microsphere drives stimulated gain and loss in a traversing electron, resulting in an exceptionally broad energy spectrum spanning hundreds of sidebands. Inset, efficient transfer of energy and momentum is enabled by matching the velocities of the electron,  $v_e$ , and the wave in the optical cavity,  $v_p$ , within the evanescent modal field. **b**, Spatial map of interaction strength for two neighbouring spheres with linearly polarized excitation in the UTEM, quantified in terms of electron spectral bandwidth. **c**, Scanning electron micrograph of the two spheres on the support structure.

In this study, we focus on two scenarios using microspheres with diameters of 2  $\mu\text{m}$  and 5  $\mu\text{m}$ . The WGMs in the smaller spheres allow us to reach an extremely strong and coherent modulation of the electron wavefunction, which is manifested in the emergence of electron energy sidebands of 220 eV and 700 eV for highly chirped picosecond and moderately chirped femtosecond optical pulses, respectively. For the larger spheres, the smaller free spectral range allows us to analyse the coupling of multiple WGMs to electrons and to trace the optical cavity ring-down time. These results establish the WGM as a key element in future approaches to coherently manipulate free electrons. Moreover, we expect that WGM-enhanced electron spectroscopy has the potential of ultimately facilitating controlled coupling to individual quantum systems.

The experiments are conducted in an ultrafast TEM (UTEM) featuring electron pulses as short as 200 fs (see illustration in Fig. 1a and details in Methods). These electron pulses are passed near the surface of optically illuminated silica microspheres. The WGMs are excited by optical pump pulses with a centre wavelength of 800 nm. The similar velocities of the electron and WGM phases allow them to efficiently exchange energy through the evanescent light field that permeates vacuum (Fig. 1a inset).

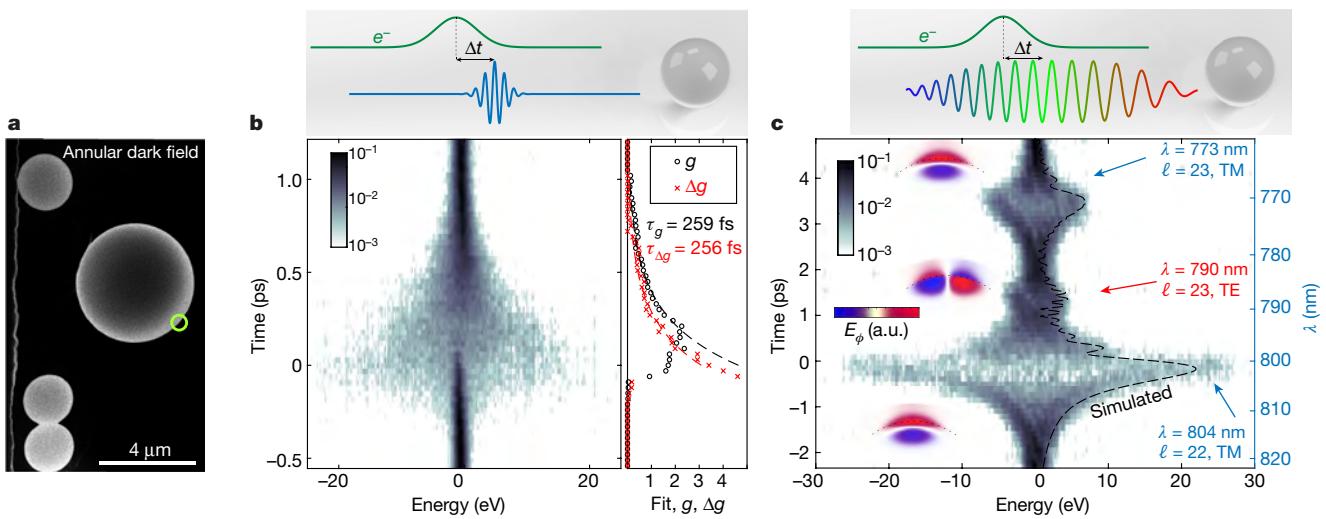
Figure 1b displays a spatial map of the electron–WGM interaction strength, with the colour code representing the width of the electron energy spectrum. The spectrum substantially broadens near the surface of the sphere and, to a lesser extent, near the metallic edges of the support structure (see Methods). The interaction strength is evident along wide sections of the sphere circumference, and is suppressed only in regions that are approximately perpendicular to the incident horizontal polarization of the light. Because the expected transverse profile of a single WGM spans approximately one wavelength, we infer from the wider azimuthal distribution of the interaction map that the driving laser excites multiple modes.

We observe very similar halo-shaped interaction maps on multiple individual spheres and assemblies, with overall magnitudes depending on the specific geometrical arrangement and in-coupling conditions. An azimuthally selective excitation of a WGM is achieved in structures with substantial geometrical anisotropy, such as the doublet displayed



**Fig. 2 | Electron spectral broadening induced by WGMs.** **a**, An annular dark-field image of a microsphere doublet comprising a pair of 2- $\mu\text{m}$ -diameter spheres that are optically pumped from the top (into the page plane). **b**, Measured electron bandwidth as a function of position. **c**, Simulated electron bandwidth from the coherent interaction with a single WGM. a.u., arbitrary units. **d**, Electron spectra measured at distances of 0 nm, 30 nm, 60 nm, 100 nm, 170 nm, 220 nm, 290 nm and 380 nm from the surface of the sphere (shifted accordingly and scaled vertically for clarity). The electron bandwidth reaches 200 eV at the surface (bottom curve; see text). The scanning line is annotated on the right inset. The decay length of the exponential fit (solid lines) of the measured interaction strength (red circles),  $L_{\text{decay}}$ , agrees with the calculated length scale (see Fig. 2c). Pumping the sphere with a shorter pulse extends the electron bandwidth to 700 eV (left inset), albeit at a reduced purity.

in Fig. 2a. Here, a pair of 2- $\mu\text{m}$ -diameter spheres is partially merged along an axis perpendicular to the edge of the supporting copper grid (see Fig. 2a). For this structure, the interaction strength is maximized when the pump excites a WGM with transverse magnetic polarization (TM), that is, for polarization along the doublet axis. The electron energy gain and loss via the PINEM are driven by the polarization component parallel to the electron path. The spatial dependence of the resulting electron bandwidth in Fig. 2b shows very good agreement with the expected PINEM interaction for a single TM-polarized WGM, as simulated in Fig. 2c. Specifically, both the azimuthal extent of about one wavelength and the exponential radial decay with a characteristic length of  $L_{\text{decay}}=0.1 \mu\text{m}$  are reproduced. The experimental evaluation of the decay length is based on a line scan, for which the spectral bandwidth was estimated at each point as a function of its distance from the surface (orange circles in Fig. 2d). It should be noted that the radial dependence of the measured coupling strength is a combination of the mode decay and the phase-matching condition. The circular propagation of the WGM implies that the linear phase velocity increases with the radius; thus, at larger radii the WGM field acquires an increasing velocity mismatch to the electrons. This radially dependent phase matching reduces the observed decay length to 100 nm from a mode decay length of 260 nm. The individual curves in Fig. 2d are spectra recorded for a set of fixed distances, which visualize the broadening of the spectrum near the sphere; the spectra are scaled vertically for clarity. Close to the sphere's surface, the spectral bandwidth extends to a remarkable 220 eV while keeping the double-lobed shape that is characteristic of



**Fig. 3 | Spectral and temporal properties of the interaction between free electrons and WGMs.** **a**, Annular dark-field micrograph of spheres with diameters of 2  $\mu\text{m}$  and 5  $\mu\text{m}$  supported by a thin carbon film. The yellow circle marks the measurement position. **b**, Left, colour map of the electron spectrum (logarithmic scale) measured after a loading of WGMs by a 50-fs-long optical pulse at time  $t = 0$  (see scheme at top). At later times, the electron is driven by light stored in the cavity. Right, the exponential decays of  $g$  and  $\Delta g$  (symbols) evaluated from each spectrum separately, and fits with a cavity lifetime of

260 fs (dashed lines) corresponding to a quality factor of  $Q = 305$ . **c**, WGM selectivity probed by the electron's spectral response. Using a frequency sweep (chirped pulse; see scheme at top), three modes are identified by the increased electron bandwidth as the sweep crosses their resonance frequencies. The simulated cavity response for a sphere with a diameter of 4.765  $\mu\text{m}$  (dashed black line) agrees with the experiment, from which we identify the mode polarizations (TM or TE) and indices  $\ell$ . The colour maps present the PINEM-relevant azimuthal component,  $E_\phi$ , of the identified modes.

a PINEM spectrum for a single, homogeneous interaction strength, representing a high-purity electron energy comb. In this experiment, by temporally stretching an ultrashort optical pulse to 3.5 ps full-width at half-maximum in an SF6-glass bar, the driving-field amplitude is essentially constant during the interaction with the electron pulse, and thus the phase modulation imprinted on the electron wavefunction is uniform and deterministic<sup>24</sup>. Replacing the SF6 bar with a shorter BK7 bar results in an optical pump duration of 400 fs (see Methods for details). Thus, the strength of the optical field is increased, albeit at the cost of some variation of the coupling strength across the electron pulse duration. The top-left inset of Fig. 2d shows an electron spectrum with a considerably wider bandwidth of 700 eV, far beyond what is typically observed in PINEM<sup>4,24,25,27</sup>. By comparing this measured energy gain (350 eV) with the calculated WGM field, we estimate a peak acceleration of 1.4 GeV m<sup>-1</sup>.

To quantify the strength and temporal uniformity of the interaction, we fit the experimental data to the expected PINEM spectrum<sup>30,41</sup>. The interaction parameter,  $g$ , also referred to as the coupling or Rabi parameter<sup>24</sup>, provides the transition amplitude between adjacent electron energy states, and thus quantifies the spectral distribution of the electron beam after the inelastic interaction with the laser field. For a uniform interaction strength, the electrons are in a pure state, fully determined by the Rabi parameter. The probability  $P_k$  for a total gain or loss of  $k$  quanta of the photon energy in such a state is  $P_k = |J_k(2|g|)|^2$ , where  $J_k$  is the Bessel function of the first kind<sup>24,30</sup>. In the limit of  $g \gg 1$ , the electron energy width (or bandwidth) scales as  $4|g|$  times the optical photon energy. In practice, however, one has to account for some variations,  $\Delta g$ , of the interaction strength that originate from the finite size of the electron beam, as well as from the spatial and temporal profile of the laser field. These averaging effects can be described by an incoherent sum of spectra calculated with a normally distributed coupling parameter with a standard deviation of  $\Delta g$ . The relative uncertainty,  $\Delta g/g$ , quantifies the deviation from a pure electron state. This parameter is highly relevant in any coherent use of the modulated electron beam, such as the formation of attosecond electron pulse trains<sup>27–29</sup>. In this case, increased values of  $\Delta g/g$  will result in unwanted broadening of the attosecond electron pulses. For the measurements in Fig. 2d, we

evaluated the relative uncertainty of the interaction strength for the 220- and 700-eV-wide spectra as  $\Delta g/g = 0.09$  and  $\Delta g/g = 0.33$ , respectively.

To characterize the resonance properties in the interaction with the WGMs, we utilize larger spheres with a diameter of 5  $\mu\text{m}$  (see Fig. 3a), for which the optical bandwidth of our laser covers several discrete cavity modes. We first study the temporal decay of the cavity field by measuring the electron spectrum as a function of time delay between the electron pulses and 50-fs optical pump pulses (see Fig. 3b). The electron spectrum rapidly broadens around the temporal overlap ( $t = 0$ ) and narrows down to its original width with an exponential decay time of about 260 fs (see Fig. 3b, right), as expected for a Lorentzian resonance with a quality factor of  $Q = \pi c/\lambda = 305$  (where  $\tau$  is the exponential time constant,  $c$  is the speed of light in vacuum and  $\lambda$  is the vacuum wavelength). This value is somewhat lower than the calculated radiative quality factor ( $Q = 485$ ), probably because of scattering from structural imperfections and the support, suggesting the possibility of improvement in optimized geometries. We note that for an exponentially decaying field, temporal averaging over a given electron pulse duration is expected to result in a constant  $\Delta g/g$ , yielding proportional curves for  $\Delta g$  and  $g$ . In our case,  $\Delta g/g$  is of order unity, which is caused by a temporal beating of several modes (Fig. 3c) with a similar lifetime and by a beating period shorter than the electron pulse duration.

For the spectral mode analysis, we pump the microsphere with a strongly chirped pulse, mapping optical frequencies to arrival times (see details in Methods). The delay-dependent electron spectrogram (Fig. 3c) now shows three broadened regions corresponding to distinct WGM indices. By comparing with the computed time-dependent cavity response for the chirped pulse (dashed line), we find that the strong features represent two TM modes with vacuum wavelength  $\lambda_{\text{TM},\ell=22} = 804 \text{ nm}$  and  $\lambda_{\text{TM},\ell=23} = 773 \text{ nm}$ . The weaker feature is attributed to a transverse electric (TE) mode at  $\lambda_{\text{TE},\ell=23} = 790 \text{ nm}$ , which results from a lower field along the electron trajectory (see colour-coded mode profiles of the relevant out-of-plane field component,  $E_\phi$ ). Aside from reproducing the general features of the measured spectra, the calculation also predicts small temporal oscillations, because the frequency sweep rate in the chirped pulse is only moderately slower than the cavity

ring-down. Stretching the optical pulse further would eliminate these oscillations. Similar oscillations are indeed discernible in some areas of the experimental spectrogram between  $t = 0$  and  $t = 1\text{ ps}$ .

In conclusion, our work combines resonant cavity enhancement with electron–light phase matching to strongly drive electron–light transitions, resulting in a coherent electron spectral broadening spanning hundreds of photon orders. This is achieved by using WGM microresonators, prototypical dielectric structures with numerous applications in spectroscopy and sensing. Although the current study focuses on low quality factors and free-space excitation, it can readily be extended to chip-based microresonators with quality factors greater than 10<sup>7</sup> (ref. <sup>42</sup>) and using phase-matched, high-ideality and fibre-pigtailed nanophotonic waveguides for excitation. Importantly, such an approach could also reach the strong-coupling regime, promoting the possibility for electron–photon entanglement<sup>9</sup>. Moreover, attosecond optical modulation of the phase and density of continuous electron beams in standard electron microscopes appears in reach. Such high-current electron beams dressed by light could transfer optical polarizations with nanometre-sized focal spots, possibly acquiring high-resolution spectroscopic information from resonator-coupled atoms<sup>39,40,43</sup>, molecules<sup>37</sup> and nanoparticles<sup>44</sup>. In a similar vein, continuous-wave probing of the phase response of resonators could enable all-optical real-time detection of electrons by a dispersive phase shift, in analogy to previous all-optical molecule detection<sup>37</sup>. The ability of microresonators to generate optical dissipative solitons may enable the coupling of electrons to tightly localized fields with lengths of only a few optical cycles<sup>36</sup>, relevant for time-gated interaction. Generally, cavity-enhanced and phase-matched near-field electron interactions could allow a merging of electron microscopes and photonic chip-based microresonators, with far-reaching consequences in local quantum control and sensing.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2320-y>.

1. Oh, S. H., Kauffmann, Y., Scheu, C., Kaplan, W. D. & Röhle, M. Ordered liquid aluminum at the interface with sapphire. *Science* **310**, 661–663 (2005).
2. Kimoto, K. et al. Element-selective imaging of atomic columns in a crystal using STEM and EELS. *Nature* **450**, 702–704 (2007).
3. Glaeser, R. M. How good can cryo-EM become? *Nat. Methods* **13**, 28–32 (2016).
4. Barwick, B., Flannigan, D. J. & Zewail, A. H. Photon-induced near-field electron microscopy. *Nature* **462**, 902–906 (2009).
5. Peralta, E. A. et al. Demonstration of electron acceleration in a laser-driven dielectric microstructure. *Nature* **503**, 91–94 (2013).
6. Breuer, J. & Hommelhoff, P. Laser-based acceleration of nonrelativistic electrons at a dielectric structure. *Phys. Rev. Lett.* **111**, 134803 (2013).
7. Carbone, F., Kwon, O.-H. & Zewail, A. H. Dynamics of chemical bonding mapped by energy-resolved 4D electron microscopy. *Science* **325**, 181–184 (2009).
8. Cai, W., Sainidou, R., Xu, J., Polman, A. & García de Abajo, F. J. Efficient generation of propagating plasmons by electron beams. *Nano Lett.* **9**, 1176–1181 (2009).
9. Kfir, O. Entanglements of electrons and cavity photons in the strong-coupling regime. *Phys. Rev. Lett.* **123**, 103602 (2019).
10. Di Giulio, V., Kociak, M. & García de Abajo, F. J. Probing quantum optical excitations with fast electrons. *Optica* **6**, 1524–1534 (2019).
11. Schwartz, O. et al. Laser phase plate for transmission electron microscopy. *Nat. Methods* **16**, 1016–1020 (2019).
12. García de Abajo, F. J. Optical excitations in electron microscopy. *Rev. Mod. Phys.* **82**, 209–275 (2010).
13. Talebi, N. Interaction of electron beams with optical nanostructures and metamaterials: from coherent photon sources towards shaping the wave function. *J. Opt.* **19**, 103001 (2017).
14. García de Abajo, F. J. & Kociak, M. Probing the photonic local density of states with electron energy loss spectroscopy. *Phys. Rev. Lett.* **100**, 106804 (2008).
15. Nelayah, J. et al. Mapping surface plasmons on a single metallic nanoparticle. *Nat. Phys.* **3**, 348–353 (2007).
16. Talebi, N. et al. Excitation of mesoscopic plasmonic tapers by relativistic electrons: phase matching versus eigenmode resonances. *ACS Nano* **9**, 7641–7648 (2015).
17. Hörl, A. et al. Tomographic imaging of the photonic environment of plasmonic nanoparticles. *Nat. Commun.* **8**, 37 (2017).
18. Bosman, M. et al. Surface plasmon damping quantified with an electron nanoprobe. *Sci. Rep.* **3**, 1312 (2013).
19. García de Abajo, F. J. & Kociak, M. Electron energy-gain spectroscopy. *New J. Phys.* **10**, 073035 (2008).
20. Pomarico, E. et al. meV resolution in laser-assisted energy-filtered transmission electron microscopy. *ACS Photonics* **5**, 759–764 (2018).
21. Das, P. et al. Stimulated electron energy loss and gain in an electron microscope without a pulsed electron gun. *Ultramicroscopy* **203**, 44–51 (2019).
22. Asenjo-García, A. & García de Abajo, F. J. Plasmon electron energy-gain spectroscopy. *New J. Phys.* **15**, 103021 (2013).
23. Wang, K. et al. Coherent interaction between free electrons and a photonic cavity. *Nature* <https://doi.org/10.1038/s41586-020-2321-x> (2020).
24. Feist, A. et al. Quantum coherent optical phase modulation in an ultrafast transmission electron microscope. *Nature* **521**, 200–203 (2015).
25. Madan, I. et al. Holographic imaging of electromagnetic fields via electron-light quantum interference. *Sci. Adv.* **5**, eaav8358 (2019).
26. Vanacore, G. M. et al. Ultrafast generation and control of an electron vortex beam via chiral plasmonic near fields. *Nat. Mater.* **18**, 573–579 (2019).
27. Priebe, K. E. et al. Attosecond electron pulse trains and quantum state reconstruction in ultrafast transmission electron microscopy. *Nat. Photon.* **11**, 793–797 (2017).
28. Morimoto, Y. & Baum, P. Diffraction and microscopy with attosecond electron pulse trains. *Nat. Phys.* **14**, 252–256 (2018).
29. Kozák, M., Schönenberger, N. & Hommelhoff, P. Ponderomotive generation and detection of attosecond free-electron pulse trains. *Phys. Rev. Lett.* **120**, 103203 (2018).
30. Park, S. T., Lin, M. & Zewail, A. H. Photon-induced near-field electron microscopy (PINEM): theoretical and experimental. *New J. Phys.* **12**, 123028 (2010).
31. Constant, E. et al. Optimizing high harmonic generation in absorbing gases: model and experiment. *Phys. Rev. Lett.* **82**, 1668–1671 (1999).
32. Breuer, J., McNeur, J. & Hommelhoff, P. Dielectric laser acceleration of electrons in the vicinity of single and double grating structures—theory and simulations. *J. Phys. B* **47**, 234004 (2014).
33. Kozák, M. et al. Acceleration of sub-relativistic electrons with an evanescent optical wave at a planar interface. *Opt. Express* **25**, 19195–19204 (2017).
34. Dahan, R. et al. Observation of the stimulated quantum Cherenkov effect. Preprint at <https://arxiv.org/abs/1909.00757> (2019).
35. Aspelmeier, M., Kippenberg, T. J. & Marquardt, F. Cavity optomechanics. *Rev. Mod. Phys.* **86**, 1391–1452 (2014).
36. Brasch, V. et al. Photonic chip-based optical frequency comb using soliton Cherenkov radiation. *Science* **351**, 357–360 (2016).
37. Vollmer, F. & Arnold, S. Whispering-gallery-mode biosensing: label-free detection down to single molecules. *Nat. Methods* **5**, 591–596 (2008).
38. Arnold, S., Khoshimra, M., Teraoka, I., Holler, S. & Vollmer, F. Shift of whispering-gallery modes in microspheres by protein adsorption. *Opt. Lett.* **28**, 272–274 (2003).
39. Aoki, T. et al. Observation of strong coupling between one atom and a monolithic microresonator. *Nature* **443**, 671–674 (2006).
40. Shomroni, I. et al. All-optical routing of single photons by a one-atom switch controlled by a single photon. *Science* **345**, 903–906 (2014).
41. García de Abajo, F. J., Asenjo-García, A. & Kociak, M. Multiphoton absorption and emission by interaction of swift electrons with evanescent light fields. *Nano Lett.* **10**, 1859–1863 (2010).
42. Liu, J. et al. Ultralow-power chip-based soliton microcombs for photonic integration. *Optica* **5**, 1347 (2018).
43. Bechler, O. et al. A passive photon-atom qubit swap operation. *Nat. Phys.* **14**, 996–1000 (2018).
44. He, L., Özdemir, S. K., Zhu, J., Kim, W. & Yang, L. Detecting single viruses and nanoparticles using whispering gallery microlasers. *Nat. Nanotechnol.* **6**, 428–432 (2011).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

# Article

## Methods

### Experimental details

The experimental system is a TEM (JEOL JEM 2100F) modified to use a pulsed photoelectron beam, with pulses as short as 200 fs. The technical details of the UTEM are given in ref.<sup>45</sup>. The pump laser is an amplified Ti:sapphire system (Coherent RegA) that provides pulses centred at a wavelength of 800 nm, with a bandwidth of 34 nm full-width at half-maximum (FWHM), a repetition rate of 600 kHz and an average power of 150 mW entering the TEM. For experiments with a highly chirped pumping laser pulse (3.5 ps FWHM), the beam passes through a 19-cm-long bar made of Schott dense flint glass, SF6, resulting in a group velocity dispersion (GDD) of approximately  $3.8 \times 10^4$  fs<sup>2</sup>. For the moderately chirped pulse (400 fs FWHM), we used a 10-cm-long bar of BK7 glass that provided a GDD of approximately  $4.5 \times 10^3$  fs<sup>2</sup>. In practice, we calculated the full dispersion relation of our pulse in the glass bars using the dispersion coefficients in ref.<sup>46</sup>. For short pulses (50 fs; see Fig. 3b), no glass bar was used in the beam path.

The optical beam is nearly co-propagating with the electron beam (6° off axis) and focuses down to a characteristic mode size of 10 μm FWHM. The timing between the electron pulse and the laser is controlled by a delay stage.

The experiment (see illustration in Fig. 1a) investigates the electron interaction with WGMs circling inside silica microspheres with diameters of 2 μm and 5 μm (SSD5003 and SSD5000, respectively, Bangs Laboratories). For the sample preparation, the spheres are immersed in ethanol and are randomly distributed by drop-casting on a lacy-carbon support (ultrathin carbon/lacy support on 400 mesh, Ted Pella). The experiments use scanning TEM for a systematic acquisition of the electron spectrum and spectral bandwidth (Gatan Enfinium spectrometer) and an annular dark-field detector.

### Simulation details of the PINEM map in Fig. 2c

The calculation uses electrons accelerated to 200 keV that interact locally with a WGM in a 2-μm-diameter sphere having a vacuum wavelength of  $\lambda = 806$  nm, which was found to be the mode closest to the laser wavelength using the WGMode toolbox<sup>47</sup>. Furthermore, the WGMode toolbox was used to calculate the  $E_\phi$  component of the electric field on the equatorial plane, from which the PINEM-relevant field in space and time was calculated by adding the temporal and azimuthal phase  $e^{i(\omega t - \ell\phi)}$ , where  $\omega$  is the angular frequency of the mode and  $\phi$  is the azimuthal angle. The contribution to the PINEM<sup>30</sup> signal at each coordinate ( $x, y$ ) was integrated over the electron trajectory,  $t(z)$ , as  $g = \frac{q}{2\pi\hbar} \int_{-\infty}^{\infty} E_\phi(x, y, z, t(z)) dz$  ( $\hbar$ , reduced Planck constant;  $q$ , electron charge). This dimensionless integral is the electron energy gain/loss, normalized by twice the photon energy. The effect of other components of the field is negligible. Although a large path is integrated numerically, the main contribution is from distances close to the sphere's surface, near the maximal field amplitude,  $E_{\max}$ . At this proximity, the field travels approximately as a phasor,  $E_\phi \approx E_{\max} e^{i(\omega t - \ell\phi)/R}$ , where  $R$  is the sphere's radius. For a perfect phase matching, the electron's velocity fits the WGM phase velocity at the surface,  $v_p = \omega R/\ell$ . Near the equator, this simplified phase velocity term can be compared to phase-matching experiments with flat vacuum–dielectric interfaces<sup>33,34,48</sup>, vacuum–grating interfaces<sup>6,32,49</sup> and electron–nanoparticle interactions<sup>16,18,22,30,50</sup>. The role of an imperfect phase matching is similar to energy conversion processes in nonlinear and extreme nonlinear optics<sup>31,51</sup>. The evaluation of the abovementioned modal decay length of 260 nm was retrieved by an exponential fit to the electric field magnitude  $|E|$  for mode number  $\ell = 8$  near the sphere's surface. Farther from the sphere, the field is accurately described by the Hansen solutions of the spherical vectorial wave equation<sup>52,53</sup>. The solution was calculated using the WGMode toolbox for 2-μm-diameter spheres with a refractive index of  $n = 1.4533$  (ref.<sup>46</sup>). Figure 2c can be reproduced from the  $E_\phi$  field component using a MATLAB script (see 'Data availability').

### Calculation of the expected spectral response of the WGM in Fig. 3c

The only fitting parameter for the expected WGM-driven electron spectrum (dashed line in Fig. 3c) is the sphere's diameter, for which the best fit was found to be 4.765 μm. This value is only 2% smaller than the measured diameter. To evaluate the resonator response, we used the tabulated refractive index for fused silica,  $n = 1.4533$  (ref.<sup>46</sup>) and the experimentally measured lifetime of 260 fs (see Fig. 3b). The measurement results in Fig. 3b suggest a uniform temporal decay for all the participating modes. This is reasonable because the decay is dominated by scattering, which is faster than the optical leakage rate. The wavelength of the resonances was calculated using the WGMode package<sup>47</sup>, and the relative strength of each resonance was based on the maximal field components,  $E_\phi$ , of the different modes. For the TM modes,  $E_\phi$  is maximal at the centre, whereas for the TE mode,  $E_\phi$  peaks at a slightly shifted position compared with the plane of circumference (see red and blue colour maps in Fig. 3c). The temporal dispersion of the pump is calculated using the measured bandwidth of our laser (34 nm) and the full chromatic dispersion in the long SF6-glass bar (see experimental details above). The arrival time of the central pumping wavelength ( $\lambda = 800$  nm) was calibrated on the basis of PINEM experiments on metallic surfaces<sup>26,45,54</sup> and was determined as time zero.

### Data availability

The data supporting the findings of this study are available within the paper and at the Open Science Framework repository at [https://osf.io/5da8g/?view\\_only=779f3a157219431bb51e48bc3fd47f47](https://osf.io/5da8g/?view_only=779f3a157219431bb51e48bc3fd47f47). Source data for Figs. 1–3 are provided with the paper.

45. Feist, A. et al. Ultrafast transmission electron microscopy using a laser-driven field emitter: femtosecond resolution with a high coherence electron beam. *Ultramicroscopy* **176**, 63–73 (2017).
46. Polynitskiy, M. N. *Refractive Index Database* <http://refractiveindex.info> (2015).
47. Balac, S. WGMode: a Matlab toolbox for whispering gallery modes volume computation in spherical optical micro-resonators. *Comput. Phys. Commun.* **243**, 121–134 (2019).
48. Garcia de Abajo, F. J. Multiple excitation of confined graphene plasmons by single free electrons. *ACS Nano* **7**, 11409–11419 (2013).
49. Yang, Y. et al. Maximal spontaneous photon emission and energy loss from free electrons. *Nat. Phys.* **14**, 894–899 (2018); correction **14**, 967 (2018).
50. Hörl, A., Trügler, A. & Hohenester, U. Tomography of particle plasmon fields from electron energy loss spectroscopy. *Phys. Rev. Lett.* **111**, 076801 (2013).
51. Boyd, R. W. *Nonlinear Optics* (Academic Press, 2003).
52. Hansen, J. E. *Spherical Near-field Antenna Measurements* (IET, 1988).
53. Balac, S. & Féron, P. *Whispering Gallery Modes Volume Computation in Optical Micro-Spheres*. Report 01279396 (HAL, 2014); <https://hal.archives-ouvertes.fr/hal-01279396>.
54. Vananore, G. M. et al. Attosecond coherent control of free-electron wave functions using semi-infinite light fields. *Nat. Commun.* **9**, 2694 (2018); correction **10**, 1069 (2019).

**Acknowledgements** We acknowledge the joint effort of the UTEM team in Göttingen and especially the support of M. Möller, J. H. Gaida, T. Danz and T. Domröse. We thank J. Liu for productive discussions. O.K. gratefully acknowledges funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement number 752533. T.R.H. acknowledges the support of a postdoctoral fellowship from the Alexander von Humboldt Foundation and its sponsor, the German Federal Ministry for Education and Research. This work was funded by the Deutsche Forschungsgemeinschaft (DFG) through the Collaborative Research Center 'Atomic Scale Control of Energy Conversion' (DFG-SFB 1073, project A05) and the Priority Program 'Quantum Dynamics in Tailored Intense Fields' (DFG-SPP 1840).

**Author contributions** O.K. conceived the experiment. C.R. directed the study. O.K. and H.L.-M. conducted the experiment with contributions from A.F. and T.R.H.; O.K., H.L.-M., G.S. and M.S. prepared the samples. O.K. analysed the data with contributions from A.F.; O.K. and C.R. wrote the manuscript with contributions from T.J.K., H.L.-M., A.F. and M.S. and based on discussions with all authors.

**Competing interests** The authors declare no competing interests.

#### Additional information

**Correspondence** and requests for materials should be addressed to O.K. or C.R.

**Peer review information** *Nature* thanks Nahid Talebi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

## Article

# Coherent interaction between free electrons and a photonic cavity

<https://doi.org/10.1038/s41586-020-2321-x>

Received: 26 August 2019

Accepted: 17 March 2020

Published online: 3 June 2020

 Check for updates

Kangpeng Wang<sup>1</sup>, Raphael Dahan<sup>1</sup>, Michael Shentcis<sup>1</sup>, Yaron Kauffmann<sup>2</sup>, Adi Ben Hayun<sup>1</sup>, Ori Reinhardt<sup>1</sup>, Shai Tsesses<sup>1</sup> & Ido Kaminer<sup>1</sup>✉

Advances in the research of interactions between ultrafast free electrons and light have introduced a previously unknown kind of quantum matter, quantum free-electron wavepackets<sup>1–5</sup>. So far, studies of the interactions of cavity-confined light with quantum matter have focused on bound electron systems, such as atoms, quantum dots and quantum circuits, which are considerably limited by their fixed energy states, spectral range and selection rules. By contrast, quantum free-electron wavepackets have no such limits, but so far no experiment has shown the influence of a photonic cavity on quantum free-electron wavepackets. Here we develop a platform for multidimensional nanoscale imaging and spectroscopy of free-electron interactions with photonic cavities. We directly measure the cavity-photon lifetime via a coherent free-electron probe and observe an enhancement of more than an order of magnitude in the interaction strength relative to previous experiments of electron–photon interactions. Our free-electron probe resolves the spatiotemporal and energy–momentum information of the interaction. The quantum nature of the electrons is verified by spatially mapping Rabi oscillations of the electron spectrum. The interactions between free electrons and cavity photons could enable low-dose, ultrafast electron microscopy of soft matter or other beam-sensitive materials. Such interactions may also open paths towards using free electrons for quantum information processing and quantum sensing. Future studies could achieve free-electron strong coupling<sup>6,7</sup>, photon quantum state synthesis<sup>8</sup> and quantum nonlinear phenomena such as cavity electro-optomechanics<sup>9</sup>.

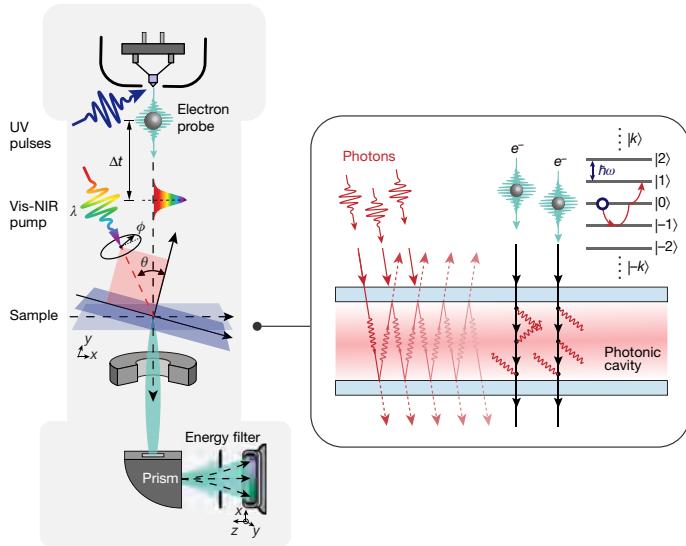
The ultrafast interaction between free electrons and laser pulses, mediated by evanescent light fields, has enabled the development of powerful tools for exploring femtosecond dynamics at the nanoscale<sup>1–3</sup>. This type of interaction, known as photon-induced near-field electron microscopy (PINEM), has been used to demonstrate the laser-driven quantum walk of free electrons<sup>4</sup>, attosecond electron pulse trains<sup>10–12</sup>, the transfer of angular momentum from the optical near field to free electrons<sup>13</sup> and the imaging of plasmons<sup>1,5</sup> at (laser-induced) millielectronvolt energy resolution<sup>10</sup>. After the interaction, the electron splits into a quantum superposition of states that are equally spaced by the photon energy  $\hbar\omega$  ( $\hbar$ , reduced Planck constant;  $\omega$ , frequency)<sup>2–4</sup>—that is, the electron spans a ‘ladder’ of energy levels that constitutes a novel type of quantum matter. One could then be inspired to position such a free electron inside a cavity with a large quality factor ( $Q$ ), changing the nature of this quantum coherent excitation in the same manner as for bound electron systems in cavities.

In comparison with bound electrons, free electrons can interact at higher energy scales and shorter timescales. Unlike bound electrons, which are constrained to discrete energy levels, free electrons occupy a continuum of energies that can be quantized in a tunable way into the electron energy ladder. This tunability could lead to fundamental

effects in quantum nonlinear optics, quantum state synthesis and quantum correlations of electrons and photons. Nevertheless, experimental electron–photon interactions reported so far have not shown any influence of a photonic cavity, in either enhanced interaction strength or duration, because of intrinsic optical losses and low  $Q$  factors. In this study, we propose a platform for nanoscale electron–cavity–photon interactions (Fig. 1), in which we experimentally achieved the strongest coherent electron–light interaction reported so far. The strength of the interaction allowed us to measure the photon lifetime in the cavity via a free-electron probe and to obtain a PINEM interaction using low pulse energies down to the picojoule scale. We demonstrated the interaction with a low-loss, dielectric photonic-crystal cavity, and measured a spatial analogue of free-electron Rabi oscillations as a function of position in the photonic-crystal cavity mode. The Rabi oscillations verify the quantum nature of electrons in our system and confirm that a purely classical theory is not sufficient to describe light–matter interactions in our experiment.

In addition to the time dynamics of the interaction, we are able to resolve the photonic band structure as a function of energy, momentum and polarization, simultaneously capturing the spatial distribution of the cavity modes at deep-subwavelength resolution. These capabilities

<sup>1</sup>Solid State Institute and Faculty of Electrical Engineering, Technion – Israel Institute of Technology, Haifa, Israel. <sup>2</sup>Department of Materials Science and Engineering, Technion – Israel Institute of Technology, Haifa, Israel. ✉e-mail: kaminer@technion.ac.il



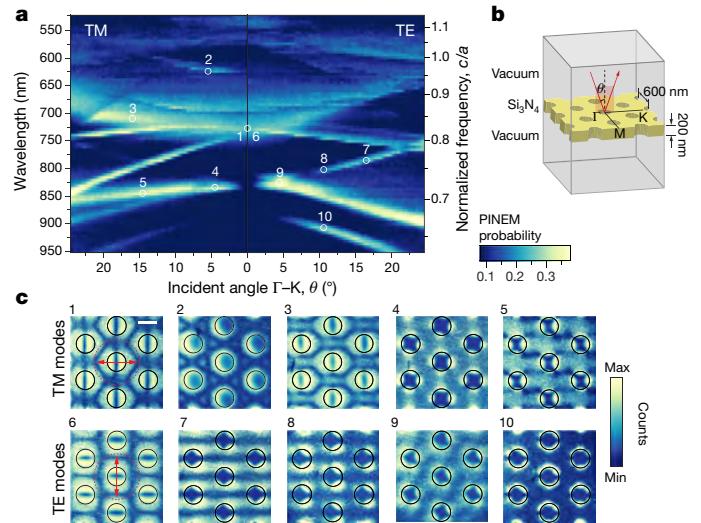
**Fig. 1 | Quantum interaction of free electrons with a photonic cavity in the UTEM.** The UTEM setup offers five degrees of freedom to measure the interactions: the delay  $\Delta t$  between the light pump and the electron probe, the pump light wavelength  $\lambda$  and polarization  $\phi$ , the sample tilt angle  $\theta$  and the electron spatial distribution in the  $x$ - $y$  plane after the interaction (elaborated in Extended Data Fig. 1). Inset, schematic of the interaction; each electron is dressed into an infinite quantum energy ladder equally spaced by the photon energy  $\hbar\omega$ . UV, ultraviolet; Vis-NIR, visible–near infrared.

offer multidimensional characterization of nanostructures beyond the limits of incoherent broadening and suggest a path towards the low-dose excitation and probing of soft matter and other fragile samples. Our results pave the way to strong coupling of free electrons with cavity photons<sup>6,7</sup> and to previously unknown types of electron–photon and electron–electron entanglement<sup>6,14</sup>. The photonic density of states in the cavity could also shape spontaneous radiation emission by free electrons, enabling regimes of Compton scattering, free-electron lasing and Cherenkov effects<sup>15–19</sup>. The preliminary results of this work were presented in May 2019<sup>20,21</sup>.

## Results

We utilize quantum free-electron wavepackets to probe an optical cavity (specifically, a photonic-crystal membrane) with multidimensional capabilities installed at an ultrafast transmission electron microscope (UTEM)<sup>1,4,5,11,22</sup> (Extended Data Fig. 1). As in other such microscopes, a femtosecond laser pulse is split into two parts: one used to excite the sample (pump) and the second one to generate the electron pulse (probe) that interacts with the sample. The probe electron is used to image the optical-field distribution in real space by electron energy filtering (Extended Data Fig. 1a), resolving the near field of nanostructures at deep-subwavelength resolution. In our setup, we also vary the sample tilt relative to the laser pump (Extended Data Fig. 1b) to enable coupling to photonic cavity resonances. The resonances are fully characterized by changing the pump laser wavelength (Extended Data Fig. 1c) and polarization (Extended Data Fig. 1d). Using a controllable delay time between the pump laser pulse and the probe electron pulse (Extended Data Fig. 1e), we can image the dynamics of light confined in the photonic cavity at ultrafast timescales. More information about the setup is provided in Methods.

Before investigating the strength of the electron–cavity–photon interaction, we used the electron probe to characterize our cavity, a triangular photonic crystal (Fig. 2). First, we measured the band structure of the photonic crystal (Fig. 2a) for both transverse magnetic (TM) and transverse electric (TE) polarizations in the wavelength range



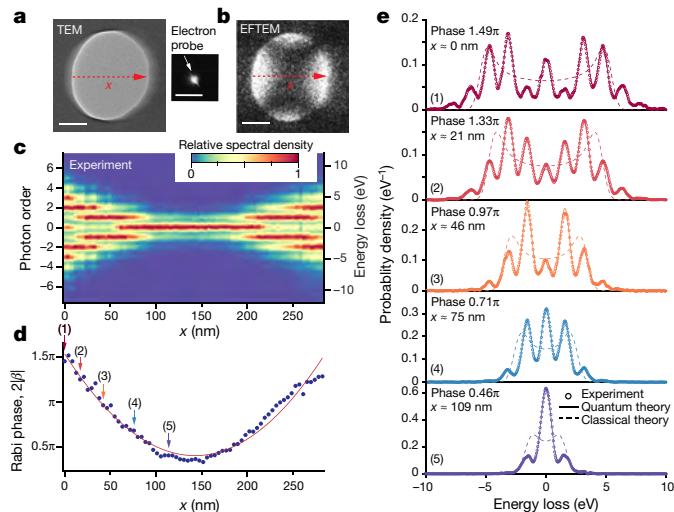
**Fig. 2 | Reconstruction of band structure and direct imaging of the Bloch modes of the photonic crystal.** **a**, Band structure measured by scanning over incident laser angles and wavelengths. The normalized frequency is given by  $c/a$ , where  $a$  is the periodicity of the photonic crystal. Each data point in the map is a separate EELS measurement of the electron–light interaction at zero delay time. The colour scale represents the interaction probability. **b**, Layout of the photonic crystal and incident pump laser pulse. **c**, Bloch modes of the photonic crystal measured at the angles and wavelengths marked in **a** (see Methods). Scale bar, 300 nm.

525–950 nm and at incident angles of  $0^\circ$ – $24.4^\circ$  (see Methods). Our spectral resolution was limited only by the laser linewidth of 5–10 nm in the visible range and by our choice of a 5-nm-wavelength step size. Our angular resolution was limited by our choice of a stage tilt step of  $0.2^\circ$ . The measured band structure (Fig. 2a) agrees very well with the simulated one (Extended Data Fig. 2a), which was calculated using the finite-difference time domain (FDTD) method (see Methods). The measurement and simulation were performed along the  $\Gamma$ –K axis. The slight artefacts in the measurement appeared because of a small polarization impurity in our setup (Extended Data Fig. 3e).

For every point measured in the band structure, we were able to probe the spatial distribution of the electric field (typical examples in Fig. 2c). The Bloch modes of the photonic crystal were measured directly by the electron interaction with the electric-field component parallel to the electron velocity<sup>1–3</sup>, at a deep-subwavelength resolution of ~30 nm (Fig. 2c). Our measurement was also performed at superpositions of TE and TM polarizations, demonstrating the rotation of the spatial distribution of the photonic-crystal Bloch modes (Extended Data Fig. 2c).

To demonstrate the quantum nature of the electrons in our system, we recorded quantum coherent oscillations in the electron energy spectrum (Fig. 3). This observation was achieved by stretching the exciting laser pulse to ~2 ps, focusing our electron probe to a ~35-nm-diameter spot (Fig. 3a) and scanning over an optical Bloch mode (Fig. 3b; Methods). This phenomenon is analogous to quantum Rabi oscillations and free-electron Rabi oscillations<sup>4</sup>, as expected from previous work on PINEM-related phenomena<sup>4,23</sup>. Figure 3e shows that our measured results cannot be explained by a fully classical description (see Supplementary Note 4) of both the light and the electron (dashed lines), in which the electron is treated as a point particle moving under the (relativistic) Newton–Lorentz equation of motion. Instead, the data match with theory when considering the electron as a quantum particle (solid lines).

Reassured by the quantum nature of the electrons in our system, we measured the photon lifetime of a cavity mode directly inside the cavity using the electrons (Fig. 4). This feat was made possible by the ultrafast temporal resolution of our system and the relatively high quality factor



**Fig. 3 | Mapping of quantum coherent electron–light interactions in a photonic-crystal cavity mode, showing a spatial-analogue of Rabi oscillations.** **a**, TEM image of a hole in the photonic-crystal membrane (shown in Fig. 2b), where the red arrow indicates the trajectory of the 35-nm electron probe (right) used in the measurements shown in **c–e**. **b**, Energy-filtered TEM (EFTEM) image showing the profile of the photonic-crystal Bloch mode. **c**, Electron energy spectrum as a function of electron probe position  $x$ , showing the oscillations of the occupation probability at each electron energy level. Each spectrum is normalized to its maxima for every  $x$  position. **d**, Rabi phase  $2|\beta|$  as a function of the position of the electron probe. The solid line is a visual guide. **e**, EELS (circles) measured along the  $x$  axis at the locations marked as (1)–(5) in **d**. Solid and dashed lines represent fits using the quantum theory (equation (1)) and the classical theory (see Supplementary Note 4), respectively. The good match of the quantum theory fit with the experimental data shows the coherent quantum nature of the free-electron probe. Scale bars, 100 nm.

achievable in photonic-crystal modes. Figure 4a compares the electron energy loss spectra (EELS) obtained from interactions with a high- or low- $Q$  photonic-crystal mode, presented as a function of the delay between the laser and electron pulses. Depending on the lifetime of the photon in the cavity, the spectrum exhibits typical time-symmetric behaviour (for low  $Q$ ) or time-asymmetric behaviour (for high  $Q$ ) around zero delay time. We note that the slight time–energy tilt in Fig. 4a–b results from the electron chirp<sup>24</sup>.

To extract the photon dynamics and lifetime, we subtracted the background of the non-interacting electrons at each time delay<sup>4,24</sup> (see Supplementary Note 5) and thus eliminated various noise factors, such as camera dark noise. The extracted dynamics is shown in Fig. 4b, where the interaction time difference is visible: the high- $Q$  spectrum is elongated, whereas the low- $Q$  spectrum persists only during the overlap of the pump and probe pulses. By fitting the experimental results to theory (comparison in Fig. 4b), we found the photon lifetime  $\tau$  of the high- $Q$  spectrum to be approximately 340 fs and its quality factor to be  $Q = \pi\tau c/\lambda \approx 384$  ( $\lambda$ , wavelength;  $c$ , speed of light in vacuum). We further corroborated this  $Q$  value using the spectral linewidths extracted from numerical simulations (see Supplementary Note 6).

We also confirmed the persistence of light in the cavity after the initial excitation pulse in Fig. 4c, which presents the interaction probability as a function of delay time (the probabilities in Fig. 4c are the sum of non-zero PINEM orders in Fig. 4b). The interaction time of the high- $Q$  mode (blue dots) is found to be longer than the convolution (blue shading) of the durations of the electron and laser pulses. The results also show that the maximal point of interaction (~40% probability) is shifted by ~207 fs, which further confirms that the total optical energy in the cavity increases over time. Such a measurement can also be performed in other photonic cavities, such as microspheres that support whispering-gallery modes<sup>25</sup>.

This high  $Q$  value leads to an enhancement in the interaction strength of more than one order of magnitude relative to the highest PINEM-related value reported so far<sup>26</sup> (Fig. 5a shows the comparison). Figure 5b presents the dependence of the interaction probability on the incident laser pulse energy, which remains visible for pulse energies as low as 100 pJ (2.67  $\mu\text{J cm}^{-2}$  fluence). The inset of Fig. 5a also shows the total interaction probability as a function of pulse energy, revealing a cavity enhancement of 10 compared to an aluminium film that we measured, and an enhancement of 30 compared to a gold film reported in the literature<sup>26</sup>.

We used recent theoretical advancements<sup>6,7</sup> to develop a fully quantum, time-dependent theory for the interaction of an electron with a cavity photon (see Supplementary Note 1). In our experiment, we can assume that the photonic state is a coherent state, that is, a classical field. At the same time, the electron is a quantum wavepacket initiated by a single energy state. These hypotheses enable several simplifications, from which we reach the following expression (see Supplementary Note 2) that agrees well with our experimental results in Figs. 3, 4:

$$P_k(E, t) = J_k^2 \left[ 2|\beta| \left( \frac{\Theta(t)e^{-t/\tau}}{\tau} \right) * e^{-(t/\sigma)^2} \right] * G(t - \zeta E, \sigma_e), \quad (1)$$

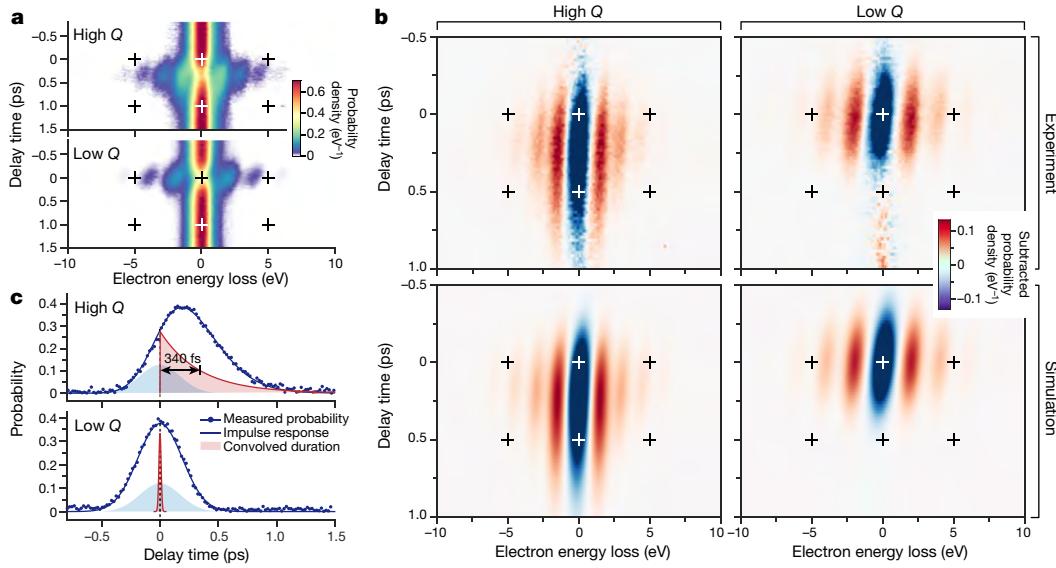
where  $*$  denotes convolution,  $E$  is the electron energy,  $t$  is the delay time and  $J_k$  is the  $k$ th-order Bessel function of the first kind.  $\beta$  represents the PINEM field<sup>3</sup> (its modulus also connects to half the Rabi phase<sup>4</sup> in Fig. 3), which scales roughly as the square root of the cavity lifetime, showing the benefit of a high- $Q$  cavity. The term  $e^{-(t/\sigma)^2}$  accounts for the Gaussian field amplitude of the excitation laser pulse.  $G(t, \sigma) = \frac{1}{\sigma\sqrt{\pi}} e^{-(t/\sigma)^2}$  describes the electron pulse duration, into which we substitute the intrinsic chirp coefficient  $\zeta$ . The standard deviations  $\sigma_e, \sigma_l$  of the electron and laser pulses, respectively, depend on the full-width at half-maximum (FWHM) pulse durations  $\tau_e, \tau_l$  via  $\sigma_{e,l} = \tau_{e,l}/(2\sqrt{\ln 2})$ .  $\Theta(t)$  is the Heaviside step function. We note that in the limit  $\tau \rightarrow 0$ , equation (1) converges back to the conventional PINEM theory<sup>2,3</sup> and can also be derived directly from it (see Supplementary Note 2). The algorithm using equation (1) to fit the experimental results to theory is described in Supplementary Note 3.

## Discussion

In summary, we directly measured the lifetime of cavity photons via a free-electron probe and achieved coherent electron–photon interaction at picojoule pulse energies. Simultaneously, we were able to record the complete real-space and energy–momentum-space information of our sample. Our work will help to promote additional near-field imaging capabilities in UTEMs, such as imaging electromagnetic waves that are deep inside<sup>27</sup> or decay outside<sup>4</sup> materials, without probe-induced distortions<sup>28</sup> (see discussion and additional references in Methods).

The substantial enhancement of the electron–photon interaction by using large- $Q$  cavity modes also suggests a path towards low-dose excitation of soft matter<sup>29</sup> and other beam-sensitive samples (for example, halide perovskites; see Fig. 5e). The low excitation dose is achieved through the field enhancement that results from the cavity resonance and the small number of electrons in ultrafast electron microscopy experiments<sup>1</sup>. The high multidimensional resolution of our method may additionally be used to explore electromagnetic structures at the nanometre<sup>30</sup> or even picometre<sup>31</sup> scale, avoiding undesired collective effects and background signals (for example, incoherent broadening) and potentially unveiling the mysteries associated with the dynamics of the single atoms therein<sup>31</sup>.

If the interaction strength is further increased to the level at which few-photon states will be enough to shift the electron energy, quantum state synthesis will be able to create new quantum optical states with exotic photon statistics<sup>8</sup> (see Fig. 5c, d and Supplementary Note 1). Then, successive interactions could show considerable changes in

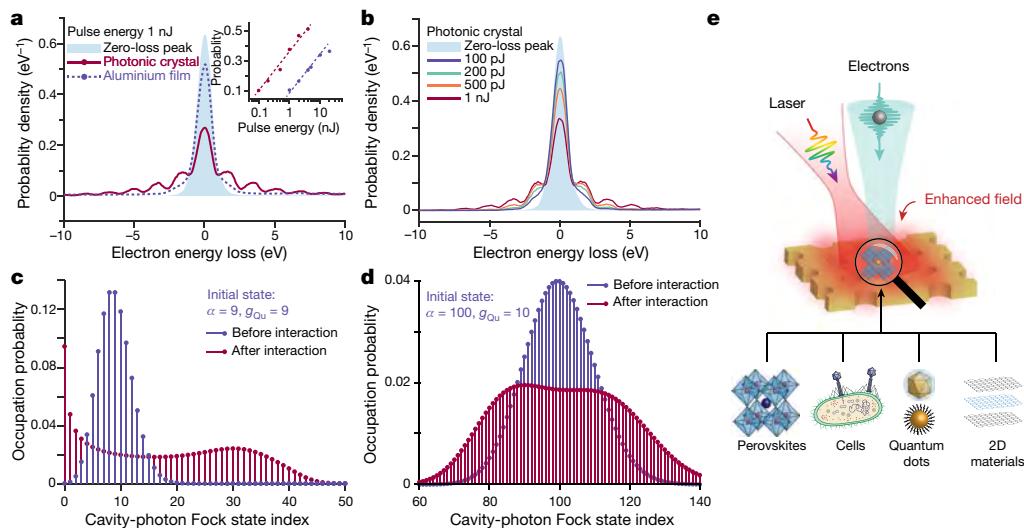


**Fig. 4 | Direct measurement of photon lifetime and dynamics using a free electron.** **a**, Time-resolved EELS maps obtained from a high- $Q$  mode (point 4 in Fig. 2a) and a low- $Q$  mode (point 2 in Fig. 2a). **b**, Time-resolved difference EELS map obtained after subtracting the zero-loss electrons. The high- $Q$  map (top left) is elongated and time-asymmetric, in contrast to the shorter and time-symmetric low- $Q$  map (top right). Our simulations using equation (1) match the experiment, both when accounting for the cavity-photon lifetime (bottom left) and when neglecting it entirely (bottom right). **c**, Experimental

(dots) and simulated (solid curve) probabilities of the electron–cavity–photon interaction as a function of delay time, obtained from the integration over all positive values in **b** at each delay time. The blue shading indicates the combined durations of the electron and laser pulses, highlighting the asymmetric features (top). The impulse response of the high- and the low- $Q$  modes is presented in red, showing an exponential decay with a lifetime of ~340 fs, as well as a shorter lifetime, below our measurement resolution (effectively a Dirac delta function).

the reaction of electrons to a photonic state, in a manner akin to photon blockade. A particularly promising route to increase the interaction strength is matching the electron and photon phase velocities over long distances<sup>32</sup>. Furthermore, the strength of the interaction between mechanical and optical modes can also be greatly enhanced

by resonators with very large  $Q$  values<sup>33</sup>, even when using nanosecond or continuous-wave light. This suggests the possible coupling of mechanical, optical and electronic degrees of freedom in UTEMs, which in the quantum limit would be a coupling of free electrons, photons and phonons—that is, cavity electro-optomechanics<sup>9</sup>.



**Fig. 5 | Enhanced interaction of electrons with picojoule laser pulses in a photonic crystal cavity, and possible applications in quantum state synthesis and ultrafast electron microscopy of sensitive materials.**

**a**, Electron energy spectra at normal incidence (pulse energy 1 nJ, wavelength 730 nm) showing a substantial enhancement of the electron–cavity–photon interaction (solid line) in comparison to a 31-nm-thick aluminium film (dashed line) and non-interacting electrons (blue shaded area). Inset, total interaction probability as a function of laser pulse energy, showing an order-of-magnitude decrease in the required energy for the photonic crystal compared with the aluminium sample. **b**, Interaction with the photonic crystal for different laser

pulse energies, showing that the interaction persists even at energies as low as 100 pJ ( $2.67 \mu\text{J cm}^{-2}$  fluence on the sample; see Supplementary Note 7). **c, d**, Simulated quantum state synthesis using the electron–cavity–photon interaction, where  $g_{\text{Qu}}$  is the single-photon coupling strength<sup>6</sup>, provided that the cavity holds a coherent state  $|\alpha\rangle$  before the interaction ( $\beta = g_{\text{Qu}}\alpha$ ; Supplementary Note 1). **e**, Prospects of using photonic-cavity structures as platforms for low-dose excitations of beam-sensitive samples for ultrafast multidimensional spectroscopy and microscopy, such as perovskite materials, frozen cells, quantum dots and two-dimensional (2D) materials.

# Article

In conclusion, much like their bound counterparts, the physics of free electrons can be modified by their insertion into a cavity. We may now think about free electrons as possible qubits to hold or transfer quantum information<sup>14,34</sup>. Free-electron spontaneous emission energies and rates may be considerably altered by the presence of the cavity's density of states, resulting in new regimes for Compton scattering, Cherenkov radiation and even free-electron lasing<sup>15–19</sup>. Perhaps even the photon statistics of the emission may be engineered, producing quantum light sources from free electrons<sup>8,19</sup>. Electron microscopy could be further developed for quantum sensing, such that the electron–photon interaction provides indirect, non-destructive sensing schemes for minute objects (such as molecules and proteins) placed upon the optical cavity inside electron microscopes.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2321-x>.

1. Barwick, B., Flannigan, D. J. & Zewail, A. H. Photon-induced near-field electron microscopy. *Nature* **462**, 902–906 (2009).
2. García de Abajo, F. J., Asenjo-García, A. & Kociak, M. Multiphoton absorption and emission by interaction of swift electrons with evanescent light fields. *Nano Lett.* **10**, 1859–1863 (2010).
3. Park, S. T., Lin, M. & Zewail, A. H. Photon-induced near-field electron microscopy (PINEM) theoretical and experimental. *New J. Phys.* **12**, 123028 (2010).
4. Feist, A. et al. Quantum coherent optical phase modulation in an ultrafast transmission electron microscope. *Nature* **521**, 200–203 (2015).
5. Piazza, L. U. C. A. et al. Simultaneous observation of the quantization and the interference pattern of a plasmonic near-field. *Nat. Commun.* **6**, 6407 (2015).
6. Kfir, O. Entanglements of electrons and cavity-photons in the strong coupling regime. *Phys. Rev. Lett.* **123**, 103602 (2019).
7. Di Giulio, V., Kociak, M. & García de Abajo, F. J. Probing quantum optical excitations with fast electrons. *Optica* **6**, 1524–1534 (2019).
8. Brattke, S., Varcoe, B. T. H. & Walther, H. Generation of photon number states on demand via cavity quantum electrodynamics. *Phys. Rev. Lett.* **86**, 3534–3537 (2001).
9. Aspelmeyer, M., Kippenberg, T. J. & Marquardt, F. Cavity optomechanics. *Rev. Mod. Phys.* **86**, 1391–1452 (2014).
10. Priebe, K. E. et al. Attosecond electron pulse trains and quantum state reconstruction in ultrafast transmission electron microscopy. *Nat. Photonics* **11**, 793–797 (2017).
11. Morimoto, Y. & Baum, P. Diffraction and microscopy with attosecond electron pulse trains. *Nat. Phys.* **14**, 252–256 (2018).
12. Kozák, M., Schönenberger, N. & Hommelhoff, P. Ponderomotive generation and detection of attosecond free-electron pulse trains. *Phys. Rev. Lett.* **120**, 103203 (2018).
13. Vanacore, G. M. et al. Ultrafast generation and control of an electron vortex beam via chiral plasmonic near fields. *Nat. Mater.* **18**, 573–579 (2019).
14. Mechel, C. et al. Imaging the collapse of electron wave-functions: the relation to plasmonic losses. In *CLEO: QELS\_Fundamental Science FF3M.6* (OSA Publishing, 2019).
15. Friedman, A., Gover, A., Kurizki, G., Ruschin, S. & Yariv, A. Spontaneous and stimulated emission from quasi-free electrons. *Rev. Mod. Phys.* **60**, 471–535 (1988).
16. García de Abajo, F. J. Multiple excitation of confined graphene plasmons by single free electrons. *ACS Nano* **7**, 11409–11419 (2013).
17. Schächter, L. *Beam-Wave Interaction in Periodic and Quasi-Periodic Structures* (Springer, 2013).
18. Roques-Carmes, C., Rivera, N., Joannopoulos, J. D., Soljačić, M. & Kaminer, I. Nonperturbative quantum electrodynamics in the Cherenkov effect. *Phys. Rev. X* **8**, 041013 (2018).
19. Rivera, N., Wong, L. J., Joannopoulos, J. D., Soljačić, M. & Kaminer, I. Light emission based on nanophotonic vacuum forces. *Nat. Phys.* **15**, 1284–1289 (2019).
20. Wang K. et al. Transmission nearfield optical microscopy (TNOM) of photonic crystal Bloch modes. In *CLEO: QELS\_Fundamental Science JTh5B.9* (OSA Publishing, 2019).
21. Wang, K. et al. Coherent interaction between free electrons and cavity photons. Preprint at <https://arxiv.org/abs/1908.06206> (2019).
22. Pomarico, E. et al. meV resolution in laser-assisted energy-filtered transmission electron microscopy. *ACS Photonics* **5**, 759–764 (2018).
23. Vanacore, G. M. et al. Attosecond coherent control of free-electron wave functions using semi-infinite light fields. *Nat. Commun.* **9**, 2694 (2018); correction **10**, 1069 (2019).
24. Park, S. T., Kwon, O.-H. & Zewail, A. H. Chirped imaging pulses in four-dimensional electron microscopy: femtosecond pulsed hole burning. *New J. Phys.* **14**, 053046 (2012).
25. Kfir, O. et al. Controlling free electrons with optical whispering-gallery modes. *Nature* <https://doi.org/10.1038/s41586-020-2320-y> (2020).
26. Liu, H., Baskin, J. S. & Zewail, A. H. Infrared PINEM developed by diffraction in 4D UEM. *Proc. Natl Acad. Sci. USA* **113**, 2041–2046 (2016).
27. Lummen, T. T. et al. Imaging and controlling plasmonic interference fields at buried interfaces. *Nat. Commun.* **7**, 13156 (2016).
28. Rotenberg, N. & Kuipers, L. Mapping nanoscale light fields. *Nat. Photonics* **8**, 919–926 (2014).
29. VandenBussche, E. J. & Flannigan, D. J. Reducing radiation damage in soft matter with femtosecond-timed single-electron packets. *Nano Lett.* **19**, 6687–6694 (2019).
30. Cognée, K. G. et al. Mapping complex mode volumes with cavity perturbation theory. *Optica* **6**, 269–273 (2019).
31. Benz, F. et al. Single-molecule optomechanics in “picocavities”. *Science* **354**, 726–729 (2016).
32. Nehemia, S. et al. Observation of the stimulated quantum Cherenkov effect. Preprint at <https://arxiv.org/abs/1909.00757> (2019).
33. Carmon, T. & Vahala, K. J. Modal spectroscopy of optoexcited vibrations of a micron-scale on-chip resonator at greater than 1 GHz frequency. *Phys. Rev. Lett.* **98**, 123901 (2007).
34. Reinhardt, O., Mechel, C., Lynch, M. & Kaminer, I. Free electron qubits. In *CLEO: QELS\_Fundamental Science FF1F.6* (OSA Publishing, 2019).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Methods

### Ultrafast electron microscopy

The experiments were performed on a UTEM that is based on a JEOL JEM-2100 Plus TEM with a LaB<sub>6</sub> electron gun and an acceleration voltage varying from 40 kV to 200 kV (schematic in Fig. 1). The UTEM<sup>1,4,5,11,22,35,36</sup> is a pump–probe setup that uses femtosecond light pulses to excite the sample and ultrafast electron pulses to probe the sample’s transient state. To this end, a 1,030-nm, -220-fs laser (Carbide, Light Conversion) operating at 1 MHz repetition rate is split into two pulses. The first pulse is converted to ultraviolet light via two stages of second-harmonic generation and is then guided to the TEM cathode by an aluminium mirror inserted in the TEM column. This process generates ultrafast electron pulses. These electron pulses travel along the z axis, penetrate the sample and provide the imaging. The second pulse is converted into variable wavelengths (and shorter pulse durations) by an optical parametric amplifier (OPA) for pumping the sample. This pulse is then guided by an additional aluminium mirror in the TEM column and impinges on the sample from the top with a small angle of -4.4° relative to the z axis in the x-z plane. The delay time between the electron pulse and the OPA pulse is controlled by a motorized stage. The Si<sub>3</sub>N<sub>4</sub> photonic-crystal sample (Ted Pella, Pelco #21588-10; see Extended Data Fig. 2b) is installed on a double-tilt TEM sample holder that allows tilting around the x and y axes at angles from -20° to 20°. To analyse the electron energy spectrum after the interaction, a post-column EELS system (Gatan) is installed in the TEM. This system also provides the energy-filtered TEM (EFTEM) capability of the EELS system for real-space imaging. The inclusion of all the above multidimensional capabilities in one setup is extremely useful for full characterization of nanoscale objects, for example, alleviating risks of losing the region of interest during the transfer of the sample between setups.

### Band structure reconstruction

To map the band structure, we operated the UTEM in TEM mode at 80 keV electron energy and parallel illumination. The EELS are collected over a range of wavelengths from 525 nm to 950 nm and incident angles from 0° to 24.4° with a zero-loss peak at -1.1 eV. Each measured EELS is centred and normalized to a probability of 1 to reduce noise from fluctuations in the electron current (Extended Data Fig. 3a). Then, the interaction probability with the optical near field is calculated by integrating the electron energy spectra outside a range that is twice the FWHM of the zero-loss peak (Extended Data Fig. 3b–d).

### Spatial-analogue Rabi oscillations

This phenomenon requires each electron to interact with a relatively homogeneous light field in both time and space. Therefore, we focused the electron beam (200 keV) to a deep subwavelength spot (~35 nm diameter, operating at the convergent-beam electron diffraction mode of the TEM) and stretched the pump laser pulses to ~2.0 ps using a 1.6-cm-thick ZnSe<sub>2</sub> plate (pulse duration measured by an autocorrelator). The photonic crystal was installed on the TEM holder, with the holder’s x axis corresponding to the  $\Gamma$ –M direction of the photonic crystal. Then the sample was tilted by 15.2° along the x axis, allowing a laser incident angle of 19.6°. The pump laser pulses were TM-polarized and had 0.5 mW power, 500 kHz repetition rate and a wavelength of 764 nm. To avoid a sample drift by heating, the laser was kept on during the entire measurement. The electron probe’s position was controlled by the deflectors of the electron microscope (beam shift knobs).

### Cavity-photon lifetime and field enhancement

The EELS are collected as a function of delay time with 200-keV electrons in TEM mode. We measured the reference zero-loss peak by probing the photonic-crystal sample with the electron pulse a few picoseconds before the laser excitation. This reference zero-loss peak is used as the background of the non-interacting electrons in

the time-resolved EELS map (Fig. 4a). After subtracting the zero-loss electrons, the time-resolved difference map was obtained (Fig. 4b). The field enhancement (Fig. 5a) was measured by comparing the maximal achieved electron interaction strength for a photonic-crystal mode to the laser excitation of an evaporated aluminium film (Ted Pella, Pelco Product #619) used as a reference sample.

### FDTD simulations

The band structure of the photonic crystal was simulated by the three-dimensional FDTD method using Lumerical FDTD software. We modelled the photonic-crystal structure as illustrated in Fig. 2c and used a plane-wave illumination source, imitating our experimental conditions. This source was placed above the photonic crystal using the broadband fixed-angle source technique, covering the entire unit cell. The temporal response of the system was recorded by time monitors inside the photonic-crystal membrane and then converted to the frequency domain via fast Fourier transform. The band structure, consisting of the frequency response over a range of incident angles, was fully mapped by a series of simulations with the source at different tilt angles. The TM/TE mode response was obtained by changing the light source to either p or s polarization.

Such simulations can also be used to obtain the theoretical Bloch mode profiles at each frequency and angle. Extended Data Fig. 2d presents some typical mode profiles that correspond well to the experimental ones in Fig. 2c.

### Imaging photonic-crystal Bloch modes

We used EFTEM at 200 keV to image the light field with deep-subwavelength resolution while providing sufficient electrons that penetrate the Si<sub>3</sub>N<sub>4</sub> membrane. The images were acquired in EFTEM mode using a slit in the energy spectrum with width ~10 eV and centred at ~10 eV (energy gain side; see Extended Data Fig. 1a). To reduce the contribution of scattered electrons, an objective aperture was applied during image exposure. We found a count loss of approximately 87.5% for electrons that penetrate the Si<sub>3</sub>N<sub>4</sub> membrane compared to electrons that move through the holes. To show the light field in the photonic-crystal membrane more clearly, post-image processing was performed to enhance the contrast of the image in the membrane area (see the following section). Consequently, the signal-to-noise ratio was lower in the membrane area.

Photonic Bloch modes were previously imaged in photonic lattices<sup>37,38</sup> above the diffraction limit, or by observing the local density of states with near-field probes on photonic-crystal waveguides<sup>39,40</sup> and slabs<sup>41,42</sup>. Our method possesses the characterization abilities of different near-field setups in space<sup>28,39–41,43,44</sup>, time<sup>28,45–49</sup>, energy–momentum space<sup>42,50–53</sup> and polarization<sup>42,54</sup>, as well as combinations thereof<sup>41,42,46,54–57</sup>. In this respect, our work<sup>20,21</sup> has been developed in parallel with other efforts<sup>38,59</sup> to pursue the integration of all the above capabilities at a comparable or better resolution than the current state of the art.

### Post-image processing of photonic-crystal Bloch modes

This section presents an example of the image processing applied on each acquired image. A raw image acquired by EFTEM is presented in Extended Data Fig. 4a, showing a typical photonic-crystal Bloch mode (Fig. 2c, image 1). Such raw images need to be post-processed because changes in the sample tilt angle and the electron transmission may cause image distortion and contrast issues. We applied three major processing steps: (1) geometric corrections for the sample tilt; (2) image cropping for a better presentation of the Bloch mode features; and (3) contrast balancing between the holes and the Si<sub>3</sub>N<sub>4</sub> membrane.

The first step included contrast increase in the raw image by adjusting the image intensity map, enabling better results of the image binarizing process using adaptive thresholding (Extended Data Fig. 4b). Next, the binarized image was analysed using a circular Hough transform

# Article

to obtain the central coordinates and true radii of the holes (Extended Data Fig. 4c). From the obtained hole positions, we defined the (moving) control points on the raw image (green dots on the left of Extended Data Fig. 4d), as well as the (fixed) control points on a reference undistorted image (green dots on the right of Extended Data Fig. 4d). By matching the shifted control points with the fixed ones (green lines on the right of Extended Data Fig. 4d), one can find the geometric distortion using image registration. This geometric distortion is represented by a  $3 \times 3$  transformation matrix, according to which the image was realigned (Extended Data Fig. 4e).

After correcting the geometric distortion of the raw image, we cropped the image to show a single unit cell (Extended Data Fig. 4f). The coordinates used for cropping were also inferred from the hole positions. In the final part of image processing, an image mask was used to sort pixels of the cropped image into hole pixels and membrane pixels (Extended Data Fig. 4g). This mask was obtained by comparing it to a binarized reference image of the same area of the sample. Extended Data Fig. 4g shows the images of the holes and the membrane. Finally, the contrast of the pixels in the membrane area can be enhanced relative to the hole areas by quantifying the relative electron transmission probability. The contrast was enhanced to match that of the holes, and both parts were combined to obtain a final image of a Bloch mode.

## Data availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

35. Flannigan, D. J. & Lindenberg, A. M. Atomic-scale imaging of ultrafast materials dynamics. *MRS Bull.* **43**, 485–490 (2018).
36. Madan, I. et al. Holographic imaging of electromagnetic fields via electron-light quantum interference. *Sci. Adv.* **5**, eaav8358 (2019).
37. Bartal, G. et al. Brillouin zone spectroscopy of nonlinear photonic lattices. *Phys. Rev. Lett.* **94**, 163902 (2005).
38. Mandelik, D., Eisenberg, H. S., Silberberg, Y., Morandotti, R. & Aitchison, J. S. Band-gap structure of waveguide arrays and excitation of Floquet-Bloch solitons. *Phys. Rev. Lett.* **90**, 053902 (2003).
39. Abashin, M. et al. Near-field characterization of propagating optical modes in photonic crystal waveguides. *Opt. Express* **14**, 1643–1657 (2006).
40. Engelen, R. J. P. et al. Local probing of Bloch mode dispersion in a photonic crystal waveguide. *Opt. Express* **13**, 4457–4464 (2005).
41. Sapienza, R. et al. Deep-subwavelength imaging of the modal dispersion of light. *Nat. Mater.* **11**, 781–787 (2012).
42. Peng, S. et al. Probing the band structure of topological silicon photonic lattices in the visible spectrum. *Phys. Rev. Lett.* **122**, 117401 (2019).
43. Adamo, G. et al. Light well: a tunable free-electron light source on a chip. *Phys. Rev. Lett.* **103**, 113901 (2009).
44. Sannomiya, T., Saito, H., Junesch, J. & Yamamoto, N. Coupling of plasmonic nanopore pairs: facing dipoles attract each other. *Light Sci. Appl.* **5**, e16146 (2016).
45. Stockman, M. I., Kling, M. F., Kleineberg, U. & Krausz, F. Attosecond nanoplasmic-field microscope. *Nat. Photonics* **1**, 539–544 (2007).

46. Man, M. K. L. et al. Imaging the motion of electrons across semiconductor heterojunctions. *Nat. Nanotechnol.* **12**, 36–40 (2017).
47. Petek, H. & Ogawa, S. Femtosecond time-resolved two-photon photoemission studies of electron dynamics in metals. *Prog. Surf. Sci.* **56**, 239–310 (1997).
48. Kubo, A. et al. Femtosecond imaging of surface plasmon dynamics in a nanostructured silver film. *Nano Lett.* **5**, 1123–1127 (2005).
49. Spektor, G. et al. Revealing the subfemtosecond dynamics of orbital angular momentum in nanoplasmonic vortices. *Science* **355**, 1187–1191 (2017).
50. Hyun, J. K., Couillard, M., Rajendran, P., Liddell, C. M. & Muller, D. A. Measuring far-ultraviolet whispering gallery modes with high energy electrons. *Appl. Phys. Lett.* **93**, 243106 (2008).
51. Cha, J. J. et al. Mapping local optical densities of states in silicon photonic structures with nanoscale electron spectroscopy. *Phys. Rev. B* **81**, 113102 (2010).
52. Tarhan, i. i. & Watson, G. H. Photonic band structure of fcc colloidal crystals. *Phys. Rev. Lett.* **76**, 315–318 (1996).
53. Honda, M. & Yamamoto, N. High-Q band edge mode of plasmonic crystals studied by cathodoluminescence. *Appl. Phys. Lett.* **104**, 081112 (2014).
54. Brenny, B. J. M., Beggs, D. M., van der Wel, R. E. C., Kuipers, L. & Polman, A. Near-infrared spectroscopic cathodoluminescence imaging polarimetry on silicon photonic crystal waveguides. *ACS Photonics* **3**, 2112–2121 (2016).
55. Meuret, S. et al. Complementary cathodoluminescence lifetime imaging configurations in a scanning electron microscope. *Ultramicroscopy* **197**, 28–38 (2019).
56. Yamamoto, N. Development of high-resolution cathodoluminescence system for STEM and application to plasmonic nanostructures. *Microscopy* **65**, 282–295 (2016).
57. Takeuchi, K. & Yamamoto, N. Visualization of surface plasmon polariton waves in two-dimensional plasmonic crystal by cathodoluminescence. *Opt. Express* **19**, 12365–12374 (2011).
58. Sola Garcia, M., Schilder, N., Meuret, S., Coenen, T. & Polman, A. Time-, and phase-resolved cathodoluminescence spectroscopy. In *Q-sort International Conference on Quantum Imaging and Electron Beam Shaping* 87–88 (2019).
59. Dani, K. Imaging the motion of charge with time-resolved photoemission electron microscopy. In *CLEO: QELS\_Fundamental Science FW4M.1* (OSA Publishing, 2019).

**Acknowledgements** The research was supported by the ERC starting grant NanoEP 851780 and the Israel Science Foundation grants 3334/19 and 831/19. K.W. is partially supported by a fellowship from the Lady Davis Foundation. S.T. acknowledges support by the Adams Fellowship Program of the Israel Academy of Science and Humanities. I.K. acknowledges the support of the Azrieli Faculty Fellowship. The experiments were performed on the UTEM of AdQuanta group of I.K., which is installed in the electron microscopy center (MIKA) of the Department of Materials Science and Engineering at the Technion. We thank IDES Ltd and especially S. T. Park for support, advice and discussions. We are grateful to Y. Yang, M. Soljačić, S. G. Johnson, J. D. Joannopoulos and M. Segev for discussions. I.K. wholeheartedly acknowledges the support of R. Magid and B. Magid, whose donation made the purchase of the UTEM possible; without their help, all the experiments presented here would not have been possible.

**Author contributions** K.W., R.D., M.S. and Y.K. carried out the experiments. A.B.H., O.R. and S.T. developed the quantum description. K.W. produced the figures and proposed and implemented the simulations and the data analysis. I.K. conceived the research. All authors provided substantial input to all aspects of the project and to the writing of the paper.

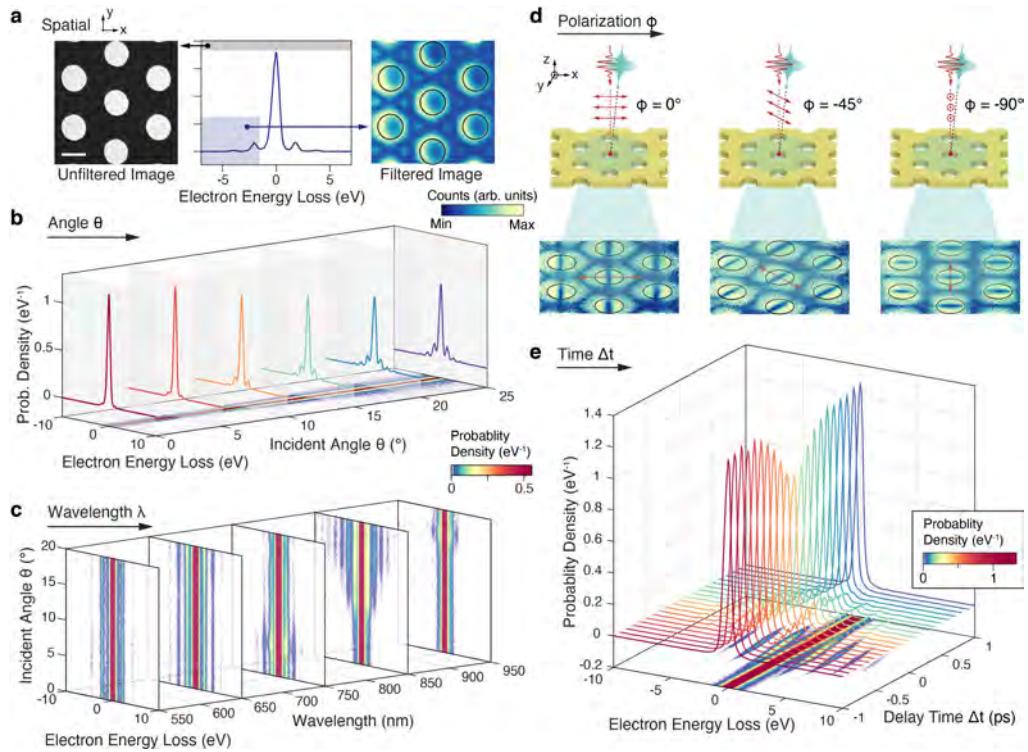
**Competing interests** The authors declare no competing interests.

### Additional information

**Correspondence and requests for materials** should be addressed to I.K.

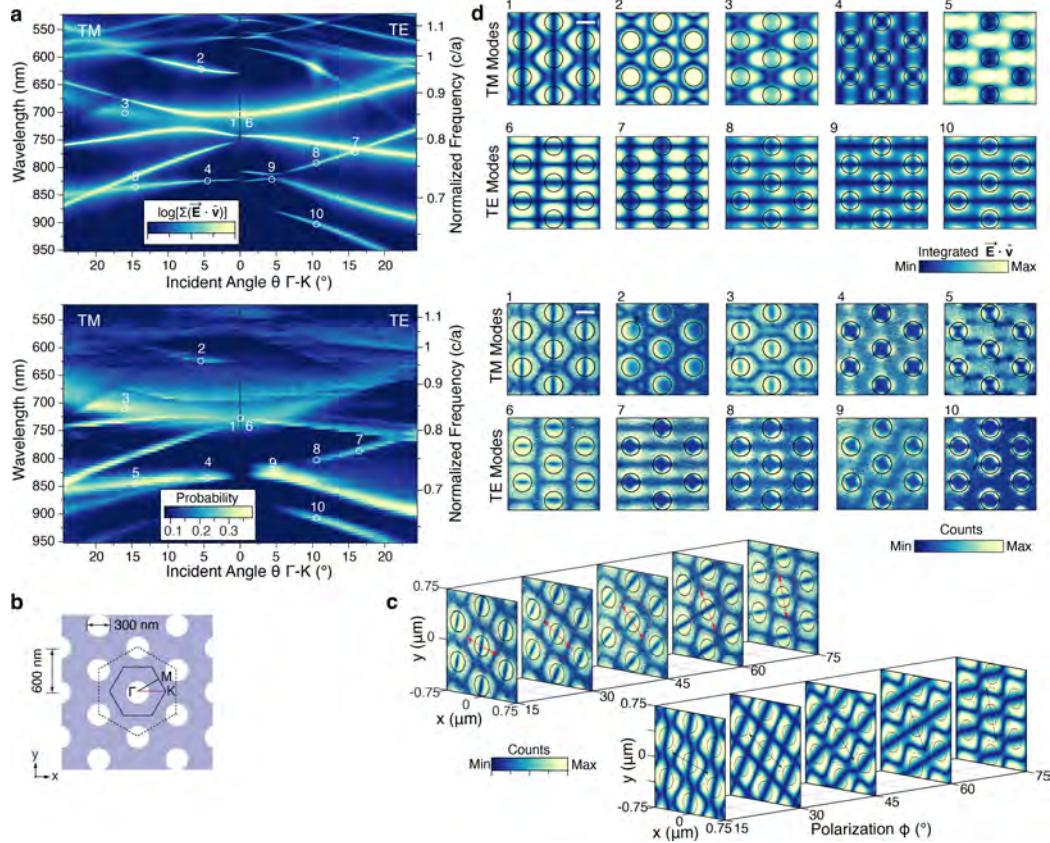
**Peer review information** *Nature* thanks Albert Polman and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



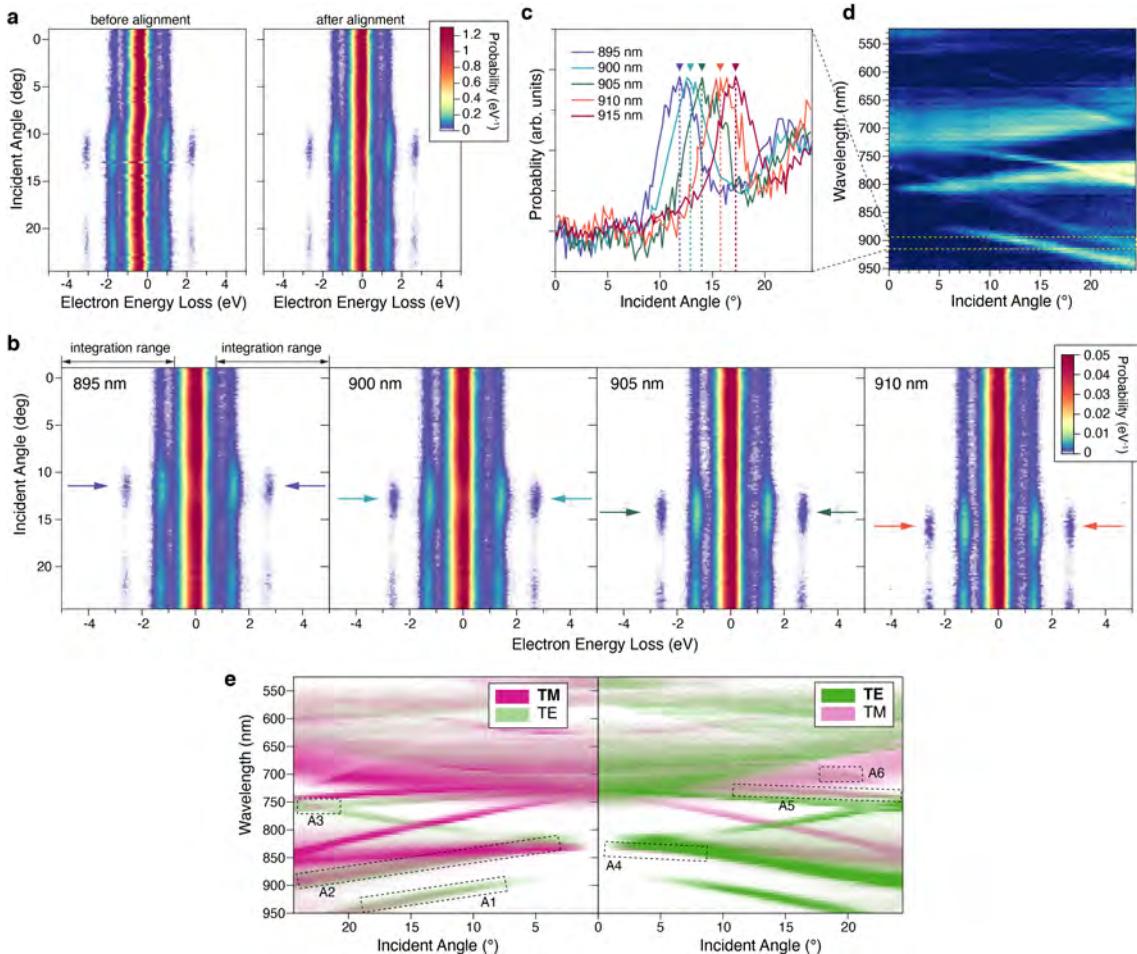
**Extended Data Fig. 1 | Ultrafast TEM for multidimensional spectroscopy (space, momentum, energy, polarization and time).** **a**, Electron microscopy images of a photonic-crystal membrane used for proof-of-concept demonstrations throughout this work. Imaging of matter and light fields with the electron energy filter disabled (left) and enabled (right). The middle plot shows a typical EELS; energy-filtered electrons used for imaging are marked by the blue shaded area. Scale bar, 300 nm. **b**, Multiple EELS over a range of sample

tilt angles, showing the angle-resolving capability of the technique; the bottom plane displays an angle-resolved EELS map assembled from individual spectra. **c**, EELS map showing the capability of mapping the electron–photon interaction for a range of wavelengths. **d**, Snapshots showing the capability of imaging excitations of different incident light polarizations. **e**, Multiple EELS showing the femtosecond time resolution of the interaction; the bottom plane displays a time-resolved EELS map assembled from individual spectra.



**Extended Data Fig. 2 | Comparison between simulations and measurements of photonic-crystal band structures and images of Bloch modes, also showing the polarization control of Bloch modes.** **a**, Simulated band structure (upper panel) by the FDTD method (see Methods) matching the measured band structure (lower panel; same as Fig. 2a). **b**, Illustration of the photonic crystal on a 200-nm-thick  $\text{Si}_3\text{N}_4$  membrane. **c**, Measured (top) and

FDTD-simulated (bottom) photonic-crystal Bloch mode images rotating with the polarization direction (double arrows) at normal incidence and 730 nm wavelength (points 1 and 6 in a). **d**, Simulated images (top) of photonic-crystal Bloch modes corresponding to experimental ones (lower panel; same as Fig. 2c), where  $\vec{E}$  is the electric field of the light and  $\vec{v}$  is the unit vector of the electron velocity. Scale bars, 300 nm.



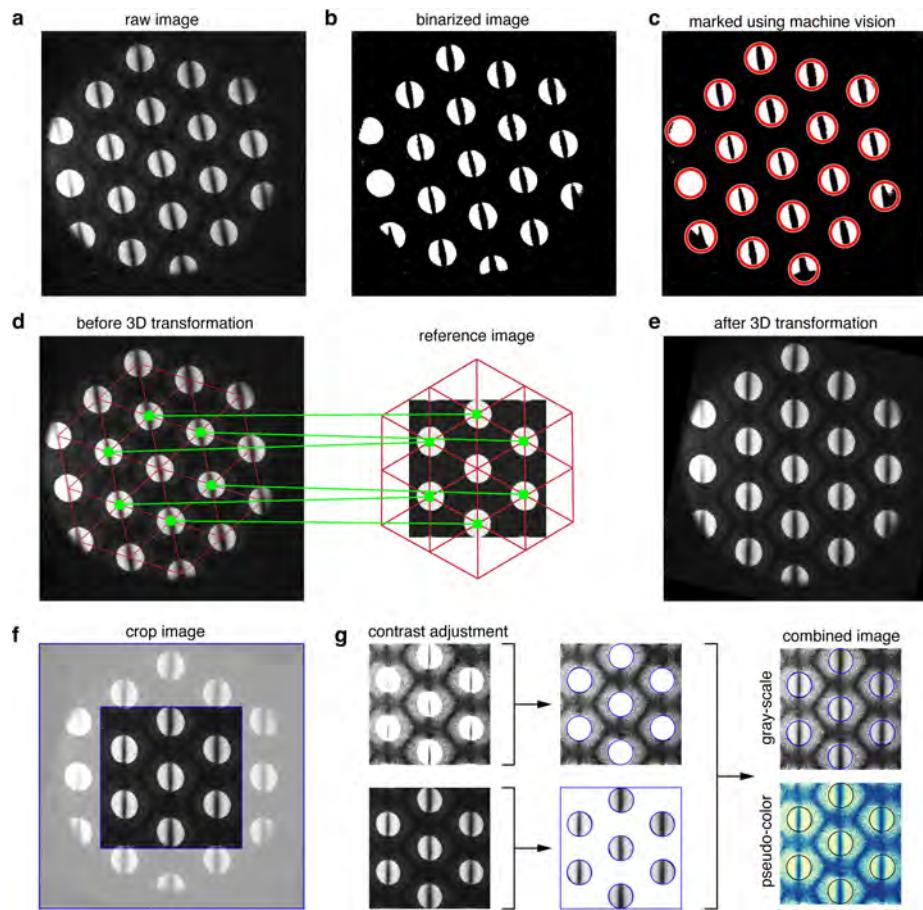
### Extended Data Fig. 3 | EELS alignment and reconstruction of the photonic-crystal band structure from a series of EELS measurements.

**a**, Example of an EELS map before (left) and after (right) alignment of the zero-loss peak (after normalizing the total probability to 1). The drift of the spectra along the x axis may occur because of magnetic fluctuations, and it was corrected by tracking the position of the zero-loss peak in each spectrum.

**b**, Examples of EELS maps as a function of angle, collected at excitation wavelengths of 895 nm, 900 nm, 905 nm and 910 nm, showing that the location of a resonance peak shifts from -12° to -17°. The double arrows indicate the integration energy range used to extract the signal, which was chosen to exclude the contribution from the zero-loss peak and was defined as twice the FWHM of the zero-loss peak. **c**, Curves showing the electron-photon interaction probability as a function of incident angle, obtained from the

integration of the EELS angle maps along the energy axis. The triangles indicate the peak shift as a function of wavelength. **d**, Band structure reconstructed from a series of measurements. The dashed box indicates the pixels shown in the other panels. **e**, Artefacts in the measured photonic-crystal band structure. The polarization impurity in the laser pulse excitation causes a partial mix of the TM and TE bands. The dashed boxes indicate artefacts in each polarization. The TM (pink) and TE (light green) band structures are equivalent to the data presented in Fig. 2a. To illustrate that the artefacts stem from mixed polarization, we overlaid the TE bands (light green) on the TM bands in the left panel. The regions (A1–A3) where the light green mixes with pink indicate the existence of TE polarization in the TM band structure. Similar regions (A4–A6) were also found in the TE band structure.

# Article



**Extended Data Fig. 4 | Image post-processing revealing the optical-field distribution of a single photonic-crystal unit cell.** **a**, Raw image captured by the camera in greyscale. **b**, Black and white image binarized by adaptive thresholding, showing the features of the etched holes in the photonic-crystal membrane. **c**, The photonic-crystal holes are recognized by machine vision using the circular Hough transform. **d**, Calculation of the geometric distortion.

The control points in the red grids are generated from the recognized hole centres. The green lines indicate the control point pairs used for image registration. **e**, Aligned image recovered from the geometric distortion. **f**, Cropped image. **g**, Contrast adjustment on selected image areas: recovery of the field amplitude by normalizing the electron current in the  $\text{Si}_3\text{N}_4$  membrane relative to the holes.

## Article

# Design of robust superhydrophobic surfaces

<https://doi.org/10.1038/s41586-020-2331-8>

Received: 1 February 2019

Accepted: 1 April 2020

Published online: 3 June 2020



Dehui Wang<sup>1</sup>, Qiangqiang Sun<sup>1</sup>, Matti J. Hokkanen<sup>2,3</sup>, Chenglin Zhang<sup>1</sup>, Fan-Yen Lin<sup>4</sup>, Qiang Liu<sup>5</sup>, Shun-Peng Zhu<sup>5</sup>, Tianfeng Zhou<sup>6</sup>, Qing Chang<sup>7</sup>, Bo He<sup>7</sup>, Quan Zhou<sup>3</sup>, Longquan Chen<sup>8</sup>, Zuankai Wang<sup>9</sup>, Robin H. A. Ras<sup>2,10</sup> & Xu Deng<sup>1</sup>✉

The ability of superhydrophobic surfaces to stay dry, self-clean and avoid biofouling is attractive for applications in biotechnology, medicine and heat transfer<sup>1–10</sup>. Water droplets that contact these surfaces must have large apparent contact angles (greater than 150 degrees) and small roll-off angles (less than 10 degrees). This can be realized for surfaces that have low-surface-energy chemistry and micro- or nanoscale surface roughness, minimizing contact between the liquid and the solid surface<sup>11–17</sup>. However, rough surfaces—for which only a small fraction of the overall area is in contact with the liquid—experience high local pressures under mechanical load, making them fragile and highly susceptible to abrasion<sup>18</sup>. Additionally, abrasion exposes underlying materials and may change the local nature of the surface from hydrophobic to hydrophilic<sup>19</sup>, resulting in the pinning of water droplets to the surface. It has therefore been assumed that mechanical robustness and water repellency are mutually exclusive surface properties. Here we show that robust superhydrophobicity can be realized by structuring surfaces at two different length scales, with a nanostructure design to provide water repellency and a microstructure design to provide durability. The microstructure is an interconnected surface frame containing ‘pockets’ that house highly water-repellent and mechanically fragile nanostructures. This surface frame acts as ‘armour’, preventing the removal of the nanostructures by abradants that are larger than the frame size. We apply this strategy to various substrates—including silicon, ceramic, metal and transparent glass—and show that the water repellency of the resulting superhydrophobic surfaces is preserved even after abrasion by sandpaper and by a sharp steel blade. We suggest that this transparent, mechanically robust, self-cleaning glass could help to negate the dust-contamination issue that leads to a loss of efficiency in solar cells. Our design strategy could also guide the development of other materials that need to retain effective self-cleaning, anti-fouling or heat-transfer abilities in harsh operating environments.

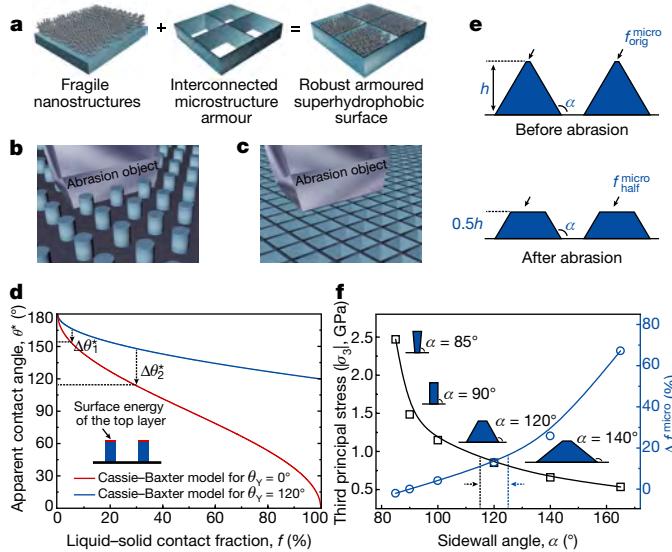
Minimizing the contact area between liquid and solid is a widely used strategy to enhance superhydrophobicity; however, it results in fragile surface textures and poor resistance to wear<sup>20</sup>. Various approaches have been explored to address this problem—for example strengthening the bonding between the coating and the substrate by using an adhesion layer<sup>21,22</sup>, bearing the abrasion force by randomly introducing discrete microstructures<sup>23–26</sup> and allowing abrasion by sacrificing the upper layers of a self-similar structure<sup>27–29</sup>—but these have resulted in only modest improvements in robustness (Supplementary Video 1). As an alternative strategy, we considered the features of mechanical durability and non-wettability separately and implemented them at two different length scales, fabricating nanostructures to impart water repellency and a microstructure to act as ‘armour’ to resist abrasion (Fig. 1a, Supplementary Fig. 1, Supplementary Video 2, Supplementary Discussion section 2.1). As the

first feature of our design, the microstructure consists of an interconnected frame that prevents the nanostructures from being removed by abradants that are larger than the frame size (Fig. 1b, c). The interconnectivity also enhances mechanical robustness, as observed for various natural structures including springtail skin and honeycomb. Additional design features must also be considered to ensure that the non-wettability of the surface is not compromised. Here we explore the relationship between the liquid–solid contact fraction  $f$ , the Young’s contact angle  $\theta_Y$  and the apparent contact angle  $\theta^*$  using the Cassie–Baxter model:

$$\cos \theta^* = f(1 + \cos \theta_Y) - 1 \quad (1)$$

In the Cassie–Baxter wetting state, the role of the Young’s contact angle is investigated by plotting  $\theta^*$  against  $f$  for different values of  $\theta_Y$

<sup>1</sup>Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, China. <sup>2</sup>Department of Applied Physics, Aalto University School of Science, Espoo, Finland. <sup>3</sup>Department of Electrical Engineering and Automation, Aalto University School of Electrical Engineering, Espoo, Finland. <sup>4</sup>Bruker Nano Surfaces Division, Santa Barbara, CA, USA. <sup>5</sup>School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu, China. <sup>6</sup>Key Laboratory of Fundamental Science for Advanced Machining, Beijing Institute of Technology, Beijing, China. <sup>7</sup>Hanergy Chengdu R&D Center, Chengdu, China. <sup>8</sup>School of Physics, University of Electronic Science and Technology of China, Chengdu, China. <sup>9</sup>Department of Mechanical Engineering, City University of Hong Kong, Hong Kong, China. <sup>10</sup>Department of Bioproducts and Biosystems, Aalto University School of Chemical Engineering, Espoo, Finland. ✉e-mail: robin.ras@aalto.fi; dengxu@uestc.edu.cn



**Fig. 1 | Design of the microstructure armour.** **a**, Schematic showing the strategy for enhancing the mechanical stability of the superhydrophobic surface by housing water-repellent nanostructures within a protective microstructure 'armour'. **b, c**, Schematics showing the mechanism by which abrasion damages discrete and interconnected microstructures. **b**, In the case of discrete microstructures, the abrading object can be easily inserted in between the microstructures and can damage both the nanostructure and the microstructure. **c**, Protection afforded by the topology of interconnected microstructures. Abrasion objects that are larger than the frame are blocked by the microstructure. **d**, Relationship between the apparent contact angle  $\theta^*$  (°) and the liquid–solid contact fraction  $f$  for an ideal Cassie–Baxter state at two different values of the Young's contact angle  $\theta_y = 0^\circ$  and  $\theta_y = 120^\circ$ . **e**, Cross-section showing the change of contact area on the top of the framework structures when the height  $h$  is fractured to half of its original value by abrasion (see Supplementary Fig. 2 and Supplementary Discussion section 2.2 for details). **f**, Influence of mechanical stability and change of the liquid–solid contact fraction  $\Delta f^{\text{micro}}$  as function of the sidewall angle ( $\alpha$ ) (see Supplementary Fig. 3 and Discussion section 2.2 for details).

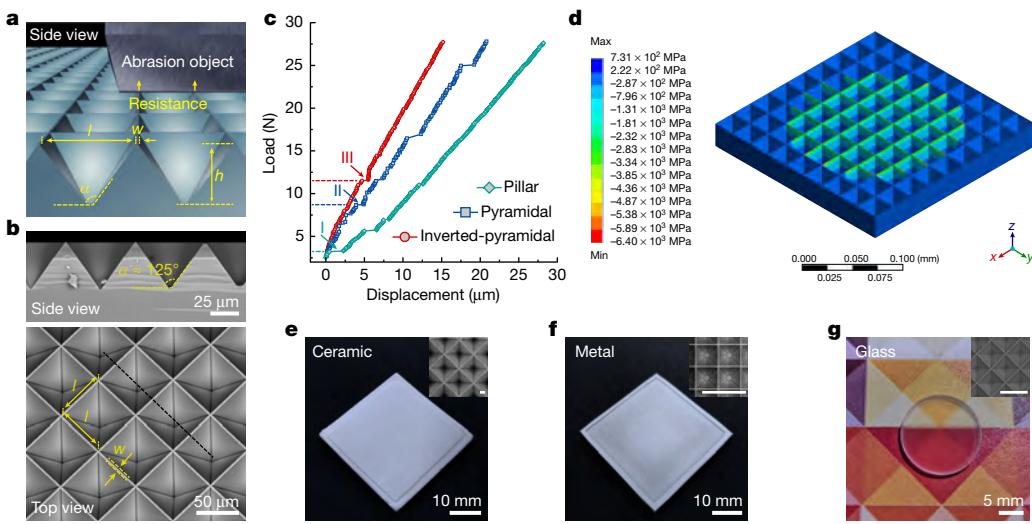
(Fig. 1d). The difference ( $\Delta\theta^*$ ) between the  $\theta^*$  values of the hydrophobic surface ( $\theta_y = 120^\circ$ ) and the hydrophilic surface ( $\theta_y = 0^\circ$ ) decreases with decreasing values of  $f$  (Fig. 1d). This indicates that the contribution of the surface chemistry of the material ( $\theta_y$ ) to the liquid repellency ( $\theta^*$ ) decreases upon minimizing  $f$ . That is, even if the top surface were to be altered from hydrophobic to hydrophilic during abrasion, the surface could still repel water if  $f$  is very small.

The mechanical stability of microstructures is dominated by their geometry. To optimize the robustness, we adjusted the angle  $\alpha$  between the sidewall and the substrate of the microstructures (Fig. 1e, f) while keeping the top contact area constant. From a structural mechanics viewpoint, increasing  $\alpha$  is usually an effective way to strengthen the structural stability of the architecture. To confirm this principle, microstructures with different values of  $\alpha$  were modelled and the stress distributions under fixed load were simulated using multipurpose finite-element analysis (Supplementary Fig. 2). The third principal stress ( $|σ_3|$ ) reduces significantly, and therefore the stability of the microstructures improves markedly, as  $\alpha$  increases (Fig. 1f). Conversely, the liquid–solid contact fraction of the microstructures,  $f^{\text{micro}}$ , increases to  $f^{\text{micro}}$  when half of the height is abraded (Fig. 1e, f, Supplementary Fig. 3, Supplementary Discussion section 2.2). The increase of  $\Delta f^{\text{micro}} = f_{\text{half}}^{\text{micro}} - f_{\text{orig}}^{\text{micro}}$  with  $\alpha$  means that the liquid–solid contact area increases—that is, the liquid adhesion force increases—for larger values of  $\alpha$ . As shown in Fig. 1f, an optimum regime emerges around  $\alpha \approx 120^\circ$  in which both superhydrophobicity and mechanical stability can be balanced and guaranteed. The second and third design features of this strategy are therefore a low  $f^{\text{micro}}$  and an  $\alpha$  value of approximately  $120^\circ$ .

Following these three design features, we fabricated a microstructure framework consisting of microscale inverted-pyramidal cavities. Using parameters including the width of the ridge  $w$  (the distance between the adjacent holes), the width of the cavity  $l$  and the height  $h$  (Fig. 2a), the liquid–solid contact fraction ( $f^{\text{micro}}$ ) can be tailored according to the following equation:

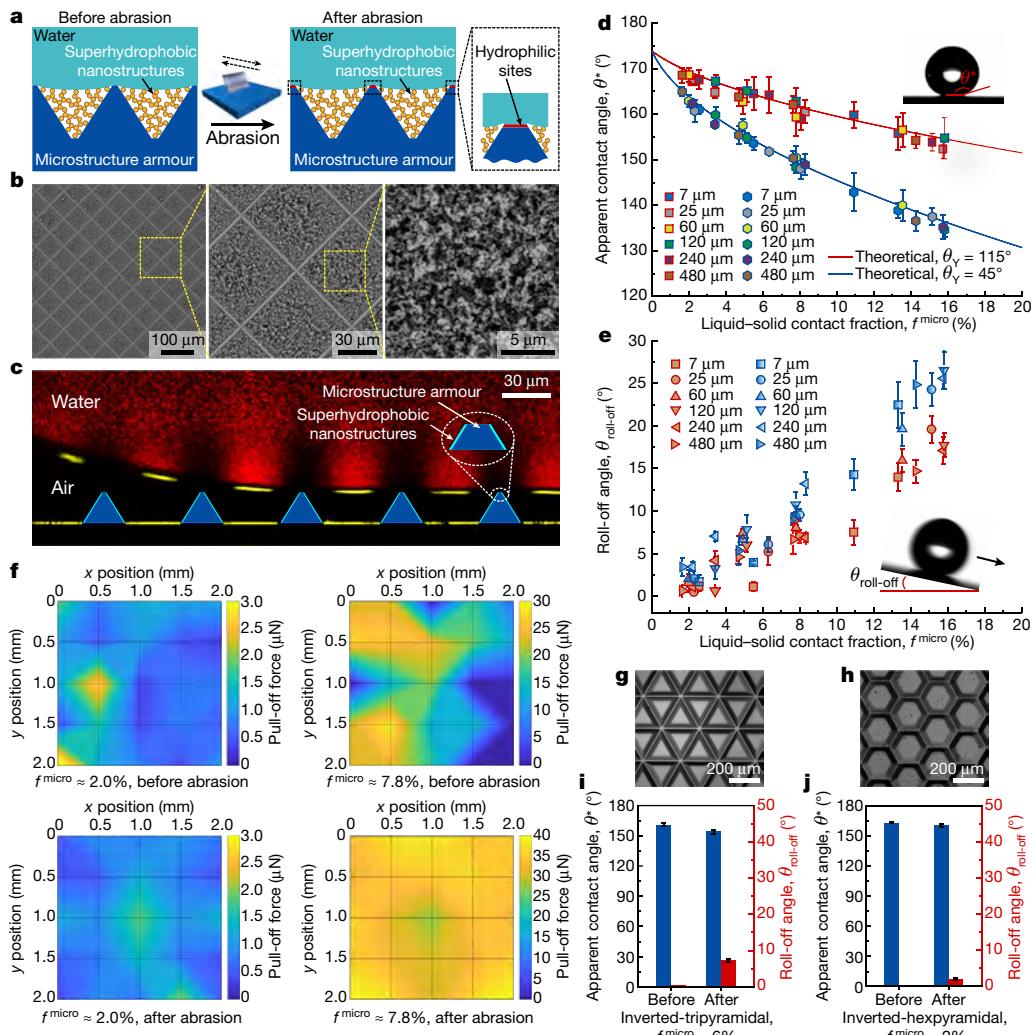
$$f^{\text{micro}} = \frac{2wl + w^2}{(w + l)^2} \quad (2)$$

Inverted-pyramidal microstructures with  $\alpha \approx 125^\circ$  were manufactured on silicon substrates by photolithography (Fig. 2b, Supplementary



**Fig. 2 | Mechanical stability of the microstructure armour.** **a**, The microstructure framework consists of an array of microscale inverted-pyramidal cavities. **b**, Scanning electron micrographs of the inverted-pyramidal structures on silicon substrates. **c**, Mechanical characterization of microstructures of different geometries on silicon substrates by microindentation. The first apparent breakpoints are indicated

as I, II and III for pillar, pyramidal and inverted-pyramidal structures, respectively. **d**, Simulated stress distribution on the inverted-pyramidal microstructures. **e–g**, Photographs of the microstructure armour on ceramic (e), metal (f) and glass (g) substrates. The insets show the corresponding scanning electron micrographs, scale bars 50  $\mu\text{m}$ .

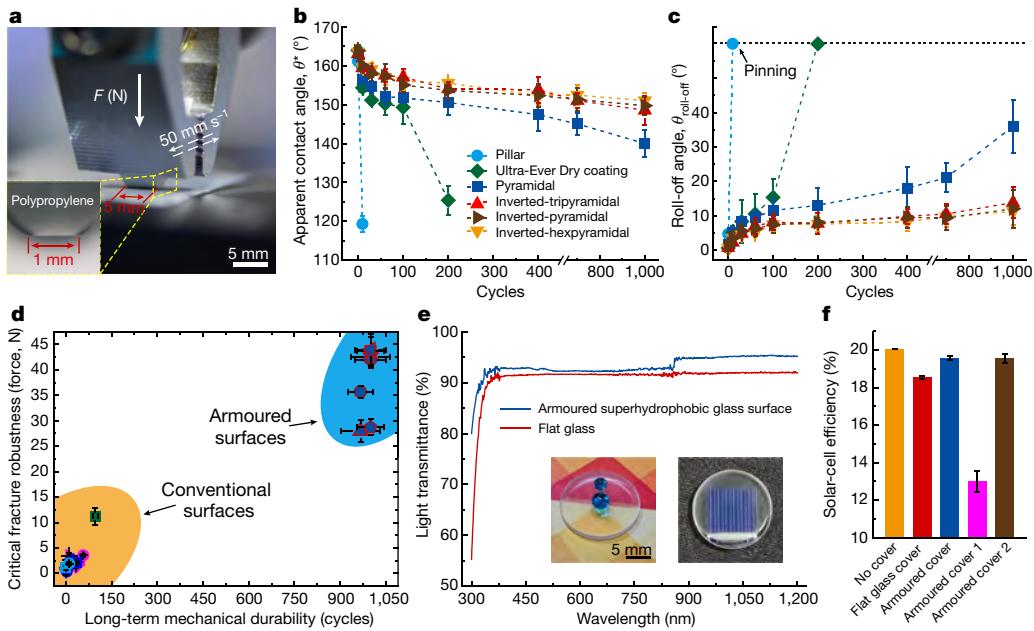


**Fig. 3 | Evaluation of water repellency after abrasion.** **a**, Schematic showing the mechanism by which the armoured superhydrophobic surface repels water, before and after abrasion. **b**, Scanning electron micrographs of silica fractal nanostructures housed within the silicon microstructure frame after abrasion. **c**, Confocal microscopy image of a water drop on an armoured superhydrophobic surface after abrasion. **d**, The experimental  $\theta^*$  before (square symbols) and after (hexagonal symbols) abrasion on armoured superhydrophobic surfaces compared with the theoretical model (lines; equation (3)). **e**, Comparison of  $\theta_{\text{roll-off}}$  before (red) and after (blue) abrasion.

**f**, Pull-off force maps showing the adhesion of water drops onto armoured ( $w + l = 60 \mu\text{m}$ ) superhydrophobic surfaces with different  $f^{\text{micro}}$  values before and after abrasion. **g**, **h**, Scanning electron micrographs showing the inverted-tripyratinal structures on a silicon substrate (**g**) and the inverted-hexpyramidal structures on an anodized aluminium alloy substrate (**h**). **i**, **j**, Apparent contact angle (blue) and roll-off angle (red) of droplets on the silicon inverted-tripyratinal (**i**) and anodized aluminium inverted-hexpyramidal (**j**) surfaces before and after abrasion. Data are mean  $\pm$  s.d. from at least five independent measurements.

Figs. 4, 5). For comparison, pillar and pyramidal microstructures were fabricated on silicon surfaces with the same  $f^{\text{micro}}$  as the inverted-pyramidal microstructures (Supplementary Fig. 6). The mechanical stability of the surface textures was assessed by microindentation testing. The load–displacement curves of these surfaces exhibited breakpoints (Fig. 2c), which were attributed to saltation during loading and subsequent displacement as a result of the fracture of the microstructures. The fractures at the first apparent breakpoints are shown in Supplementary Fig. 7. The inverted-pyramidal structures were found to resist the highest loads tested, experiencing only minor damage. These results agree well with those from the finite-element simulation (Fig. 2d, Supplementary Fig. 8), and verify our design principle. We also fabricated inverted-pyramidal structures on ceramic, metal and transparent glass substrates using embossing technology (Fig. 2e–g, Supplementary Figs. 9–13). From an engineering perspective, this approach can also be applied to curved substrates and is scalable using roll-to-plate printing technology (Supplementary Fig. 10c, d, Supplementary Video 3).

The microstructured surface exhibits superhydrophobicity after the integration of a nanostructured coating. Fluorinated fractal nano-clusters of silica were used as a model superhydrophobic nanomaterial (Fig. 3a, Supplementary Figs. 14, 15). Upon repeated scraping using a steel blade, the microstructure showed excellent resistance to the vertical pressure and the shear force, and the fractal nanostructures in between the microstructure framework remained intact (Fig. 3b). It is notable that the abrasion removes the fluorinated silane layer from the top of the microstructures, which alters the local wetting properties of the surface from hydrophobic ( $\theta_y = 115 \pm 1^\circ$ ) to hydrophilic ( $\theta_y = 45 \pm 0.5^\circ$ , Supplementary Table 1). Using laser scanning confocal microscopy we confirmed that the air–water–solid composite interface at the micro-scale is very stable, because the air–liquid–solid three-phase contact line is supported by nanoscale superhydrophobic materials (Fig. 3a, c). The water-repellent nanostructures can prevent the sagging of the liquid–air interface caused by the Laplace pressure, and the entire system remained in the constrained equilibrium Cassie–Baxter state



**Fig. 4 | Mechanical stability of the armoured superhydrophobic surface.**

**a**, Photograph showing the linear abrasion setup. A normal load was around 12 MPa. **b, c**, The influence of linear abrasion cycles under the same load on  $\theta^*$  (**b**) and  $\theta_{\text{roll-off}}$  (**c**) for various superhydrophobic coatings. The key is the same for both panels. Data are mean  $\pm$  s.d. from at least five independent measurements. **d**, Comparison of the mechanical stabilities of different superhydrophobic surfaces. See Supplementary Information section 1.11c for details. The shaded areas encompass superhydrophobic surfaces with mechanical stabilities of the same order of magnitude (see Extended Data Fig. 4 for details). Data are mean  $\pm$  s.d. from at least nine independent measurements. **e**, Transmission

spectra of armoured glass (blue) fabricated using a flat glass substrate (red). The photographs show the transparency and the superhydrophobicity of the armoured glass substrate (left) and a solar cell assembled with an armoured glass superhydrophobic surface as the cover plate (right).

**f**, The energy-conversion efficiency of the robust self-cleaning solar panel. ‘Armoured cover 1’ is the solar cell covered with armoured glass and contaminated by dust; ‘armoured cover 2’ is the contaminated solar cell after the self-cleaning process (see Supplementary Fig. 35 for details). Data are mean  $\pm$  s.d. from three independent measurements.

(Fig. 3a, c, Supplementary Fig. 16). To systematically evaluate the effect of abrasion on superhydrophobicity, we designed a series of superhydrophobic surfaces according to our strategy with different open cavity widths  $l$  and liquid–solid contact fractions  $f^{\text{micro}}$ , and measured the static contact angle  $\theta^*$  and the roll-off angle  $\theta_{\text{roll-off}}$  for the surfaces before and after abrasion. All experimental data are consistent with the theoretical model equation (3) (Fig. 3d).

$$\cos \theta^* = f^{\text{micro}} (\cos \theta_Y^{\text{micro}} + 1) + f^{\text{nano}} (\cos \theta_Y^{\text{nano}} + 1) - 1 \quad (3)$$

We defined  $f^{\text{micro}}$  as the liquid–solid contact fraction of the microstructure and  $f^{\text{nano}}$  as the liquid–solid contact fraction of the nanostructures (Fig. 3a, Supplementary Discussion section 2.3). The values of both  $\theta^*$  and  $\theta_{\text{roll-off}}$  show that these surfaces can maintain high levels of water repellency after abrasion if  $f^{\text{micro}}$  is lower than 8% (Fig. 3d, e). These results are consistent with the ideal Cassie–Baxter model, providing further evidence that non-wettability is independent of the size of inverted-pyramidal structures. However, the smaller the microstructures, the more extensive are the changes to the liquid–solid contact fraction  $\Delta f^{\text{micro}}$  after the same abrasion fracture (Supplementary Fig. 17). As such, the size of the microstructure armour could be tailored for various practical applications.

To further understand the effect of the liquid–solid contact fraction on controlling the non-wettability after abrasion, pull-off force maps of the superhydrophobic surfaces were obtained using scanning droplet adhesion microscopy. Surfaces with two different values of  $f^{\text{micro}}$  (-2% and -7.8%) were measured before and after abrasion. As shown in Extended Data Fig. 1 and Fig. 3f, the pull-off forces before abrasion were similar for both surfaces. After abrasion, the damage to the hydrophobic layer on the top of the armour resulted in an increase

in pull-off force at the same  $f^{\text{micro}}$ . However, the pull-off forces on the high- $f^{\text{micro}}$  surface increased more rapidly than those on the low- $f^{\text{micro}}$  surface, consistent with the trend of  $\theta_{\text{roll-off}}$  (Fig. 3e). Extended Data Fig. 2 illustrates a water jet impinging onto the surface before and after abrasion, with an incidence angle of 28° and a volume velocity of 6 ml min<sup>-1</sup>. These results agree with our previous wettability measurements, showing that for lower values of  $f^{\text{micro}}$  the deflected angles are higher—that is, less energy dissipation occurs when water bounces away from the surface. Similar results were obtained when analysing the impact of water droplets, further supporting this principle (Extended Data Fig. 3, Supplementary Video 4).

To demonstrate that the interconnected frame architecture—consisting of individual cavities with a large sidewall angle—could be used as a general strategy through which to achieve superior performance, we fabricated inverted-triangular-pyramidal (tripyramidal) and inverted-hexagonal-pyramidal (hexpyramidal) structures on silicon, metal and ceramic substrates (Fig. 3g, h, Supplementary Figs. 18–20). Finite-element modelling demonstrated that the stress distribution on these interconnected frame architectures is relatively uniform, showing mechanical robustness comparable with that of the inverted-pyramidal structure. As shown in Fig. 3i, j, after repeated scraping using a steel blade, the inverted-triangular-pyramidal and hexpyramidal interconnected architectures maintained robust superhydrophobic properties.

In real-world applications, surfaces are exposed to repeated abrasion. We therefore examined the long-term mechanical durability of the superhydrophobic surfaces formed from different microstructures. The abrasion was performed using a polypropylene probe as the indenter with a defined vertical pressure and reciprocating linear abrasion (Fig. 4a, Supplementary Fig. 21, Supplementary Video 5). As shown in Fig. 4b, c, the surfaces maintained a static contact angle of

greater than 150° and a roll-off angle of less than 12° even after 1,000 abrasion cycles, providing resistance to the shear force and protection for the silica nanomaterials inside (Supplementary Figs. 21, 22). To illustrate the mechanical durability of our armoured superhydrophobic surfaces, we determined the critical fracture force—that is, the maximum force required to destroy the superhydrophobicity—and the maximum number of abrasion cycles that could be tolerated without a loss of performance, and compared the results with those for conventional superhydrophobic surfaces (Supplementary Fig. 23, Supplementary Methods section 1.1c). More than 1,000 abrasion cycles could be tolerated, which is 10 times higher than for conventional superhydrophobic surfaces (Fig. 4d). The mechanical robustness of our superhydrophobic surfaces was demonstrated using tape-peeling tests, Taber abrasion tests (American Society for Testing and Materials standard) and scratch tests using an ultra-sharp object (Supplementary Figs. 24–26). We also conducted more severe durability tests, including the thermal stability (100 °C for 16 days), the resistance to chemical corrosion (immersion in aqua regia or 2.5 M of NaOH solution for 4 h), the effects of high-speed impact of a water jet (water jet at 32.6 m s<sup>-1</sup>, Weber number ( $We$ ) ≈ 36,478, Supplementary Video 6) and the tolerance to condensation-induced failure in high-humidity environments (Supplementary Figs. 27–34). Our armoured surfaces were found to maintain their superhydrophobicity even under extremely harsh conditions.

These findings demonstrate the concept of an armoured structure to improve the mechanical stability of superhydrophobic surfaces. Our strategy of considering and fabricating the required features of the surface separately enables us to balance the mechanical robustness, non-wettability and optical transparency (Fig. 4e, Supplementary Video 7). We used this strategy to create a robust and transparent self-cleaning topcoat for solar cells. This topcoat enables high energy-conversion efficiency to be maintained through the passive removal of dust contamination, which could lead to large savings in terms of freshwater, labour and cost compared with the traditional cleaning process<sup>30</sup> (Fig. 4e, f, Extended Data Figs. 5, 6, Supplementary Fig. 35). Beyond this initial proof-of-concept illustration, the generality and effectiveness of our design principle could help to move superhydrophobic surfaces towards real-world applications.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2331-8>.

- Wong, T.-S., Sun, T., Feng, L. & Aizenberg, J. Interfacial materials with special wettability. *MRS Bull.* **38**, 366–371 (2013).
- Liu, M., Wang, S. & Jiang, L. Nature-inspired superwettability systems. *Nat. Rev. Mater.* **2**, 17036 (2017).

- Bhushan, B. & Jung, Y. C. Natural and biomimetic artificial surfaces for superhydrophobicity, self-cleaning, low adhesion, and drag reduction. *Prog. Mater. Sci.* **56**, 1–108 (2011).
- Schmidt, D. L., Brady, R. F., Jr, Lam, K., Schmidt, D. C. & Chaudhury, M. K. Contact angle hysteresis, adhesion, and marine biofouling. *Langmuir* **20**, 2830–2836 (2004).
- Jung, S., Tiwari, M. K., Doan, N. V. & Poulikakos, D. Mechanism of supercooled droplet freezing on surfaces. *Nat. Commun.* **3**, 615 (2012).
- Bahng, J. H. et al. Anomalous dispersions of ‘hedgehog’ particles. *Nature* **517**, 596–599 (2015).
- Rastogi, V. et al. Synthesis of light-diffracting assemblies from microspheres and nanoparticles in droplets on a superhydrophobic surface. *Adv. Mater.* **20**, 4263–4268 (2008).
- Lee, S. Y., Rahmawan, Y. & Yang, S. Transparent and superamphiphobic surfaces from mushroom-like micropillar arrays. *ACS Appl. Mater. Interfaces* **7**, 24197–24203 (2015).
- Miljkovic, N. & Wang, E. N. Condensation heat transfer on superhydrophobic surfaces. *MRS Bull.* **38**, 397–406 (2013).
- Pan, S. et al. Coatings super-repellent to ultralow surface tension liquids. *Nat. Mater.* **17**, 1040–1047 (2018).
- Gao, L. & McCarthy, T. J. The “lotus effect” explained: two reasons why two length scales of topography are important. *Langmuir* **22**, 2966–2967 (2006).
- Tuteja, A. et al. Designing superoleophobic surfaces. *Science* **318**, 1618–1622 (2007).
- Quéré, D. Wetting and roughness. *Annu. Rev. Mater. Res.* **38**, 71–99 (2008).
- Butt, H.-J. et al. Characterization of super liquid-repellent surfaces. *Curr. Opin. Colloid Interface Sci.* **19**, 343–354 (2014).
- Liu, T. L. & Kim, C. J. Turning a surface superrepellent even to completely wetting liquids. *Science* **346**, 1096–1100 (2014).
- Sbragaglia, M. et al. Spontaneous breakdown of superhydrophobicity. *Phys. Rev. Lett.* **99**, 156001 (2007).
- Herminghaus, S. Roughness-induced non-wetting. *Europhys. Lett.* **52**, 165 (2000).
- Tian, X., Verho, T. & Ras, R. H. A. Moving superhydrophobic surfaces toward real-world applications. *Science* **352**, 142–143 (2016).
- Milionis, A., Loth, E. & Bayer, I. S. Recent advances in the mechanical durability of superhydrophobic materials. *Adv. Colloid Interface Sci.* **229**, 57–79 (2016).
- Bhushan, B. & Nosonovsky, M. Scale effects in friction using strain gradient plasticity and dislocation-assisted sliding (microslip). *Acta Mater.* **51**, 4331–4345 (2003).
- Zhang, W. et al. Facile design and fabrication of superwetting surfaces with excellent wear-resistance. *ACS Appl. Mater. Interfaces* **9**, 15776–15784 (2017).
- Lu, Y. et al. Repellent materials. Robust self-cleaning surfaces that function when exposed to either air or oil. *Science* **347**, 1132–1135 (2015).
- Zimmermann, J., Reifler, F. A., Fortunato, G., Gerhardt, L.-C. & Seeger, S. A simple, one-step approach to durable and robust superhydrophobic textiles. *Adv. Funct. Mater.* **18**, 3662–3669 (2008).
- Azimi, G., Dhiman, R., Kwon, H. M., Paxson, A. T. & Varanasi, K. K. Hydrophobicity of rare-earth oxide ceramics. *Nat. Mater.* **12**, 315–320 (2013).
- Kondrashov, V. & Rühe, J. Microcones and nanograss: toward mechanically robust superhydrophobic surfaces. *Langmuir* **30**, 4342–4350 (2014).
- Zhang, Y., Ge, D. & Yang, S. Spray-coating of superhydrophobic aluminum alloys with enhanced mechanical robustness. *J. Colloid Interface Sci.* **423**, 101–107 (2014).
- Peng, C., Chen, Z. & Tiwari, M. K. All-organic superhydrophobic coatings with mechanochemical robustness and liquid impalement resistance. *Nat. Mater.* **17**, 355–360 (2018).
- Jin, H., Tian, X., Ikkala, O. & Ras, R. H. A. Preservation of superhydrophobic and superoleophobic properties upon wear damage. *ACS Appl. Mater. Interfaces* **5**, 485–488 (2013).
- Deng, X., Mammen, L., Butt, H.-J. & Vollmer, D. Candle soot as a template for a transparent robust superamphiphobic coating. *Science* **335**, 67–70 (2012).
- Bergin, M. H., Ghoroi, C., Dixit, D., Schauer, J. J. & Shindell, D. T. Large reductions in solar energy production due to dust and particulate air pollution. *Environ. Sci. Technol. Lett.* **4**, 339–344 (2017).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

# Article

## Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

31. Liimatainen, V. et al. Mapping microscale wetting variations on biological and synthetic water-repellent surfaces. *Nat. Commun.* **8**, 1798 (2017).
32. Song, J., Xu, W. & Lu, Y. One-step electrochemical machining of superhydrophobic surfaces on aluminum substrates. *J. Mater. Sci.* **47**, 162–168 (2011).
33. Gao, X., Tong, W., Ouyang, X. & Wang, X. Facile fabrication of a superhydrophobic titanium surface with mechanical durability by chemical etching. *RSC Adv.* **5**, 84666–84672 (2015).
34. Song, J., Huang, W., Liu, J., Huang, L. & Lu, Y. Electrochemical machining of superhydrophobic surfaces on mold steel substrates. *Surf. Coat. Technol.* **344**, 499–506 (2018).
35. Yao, X. et al. Bioinspired ribbed nanoneedles with robust superhydrophobicity. *Adv. Funct. Mater.* **20**, 656–662 (2010).
36. Xu, W., Song, J., Sun, J., Lu, Y. & Yu, Z. Rapid fabrication of large-area, corrosion-resistant superhydrophobic Mg alloy surfaces. *ACS Appl. Mater. Interfaces* **3**, 4404–4414 (2011).
37. Miljkovic, N., Preston, D. J., Enright, R. & Wang, E. N. Electrostatic charging of jumping droplets. *Nat. Commun.* **4**, 2517 (2013).
38. Wang, N., Xiong, D., Deng, Y., Shi, Y. & Wang, K. Mechanically robust superhydrophobic steel surface with anti-icing, UV-durability, and corrosion resistance properties. *ACS Appl. Mater. Interfaces* **7**, 6260–6272 (2015).
39. Sun, J. et al. Electrochemical fabrication of superhydrophobic Zn surfaces. *Appl. Surf. Sci.* **315**, 346–352 (2014).
40. Hosono, E., Fujihara, S., Honma, I. & Zhou, H. Superhydrophobic perpendicularly nanopin film by the bottom-up process. *J. Am. Chem. Soc.* **127**, 13458–13459 (2005).
41. Luo, Y. et al. Design and fabrication of a tip-like ZnO nanotube array structure with condensate microdrop self-propelling function. *ChemNanomat* **2**, 1018–1022 (2016).
42. Su, F. & Yao, K. Facile fabrication of superhydrophobic surface with excellent mechanical abrasion and corrosion resistance on copper substrate by a novel method. *ACS Appl. Mater. Interfaces* **6**, 8762–8770 (2014).

**Acknowledgements** We thank B. J. Yu, J. L. Yang, Y. Li, and G.W. Zhou for discussions, and J. L. Song for providing the samples of the conventional superhydrophobic surfaces. This work was supported by the National Natural Science Foundation of China (21603026), the European Research Council ERC-2016-CoG (725513-SuperRepel), the Academy of Finland (Centres of Excellence Programme 2014–2019), Aalto University AScI/ELEC Thematic Research Programme, Business Finland Project TUTL-SDAM (6798/31/2017) and the University Research Council Grant (CityU9/CRF/13G).

**Author contributions** D.W. and X.D. conceived the research and designed the experiments. X.D. and R.H.A.R. supervised the research. D.W., Q.S., M.J.H., C.Z., F.-Y.L., T.Z. and Q.C. carried out the experiments. Q.L., S.-P.Z., Z.W., L.C., Q.Z. and B.H. built the analytical models. All authors analysed and interpreted the data and wrote the paper.

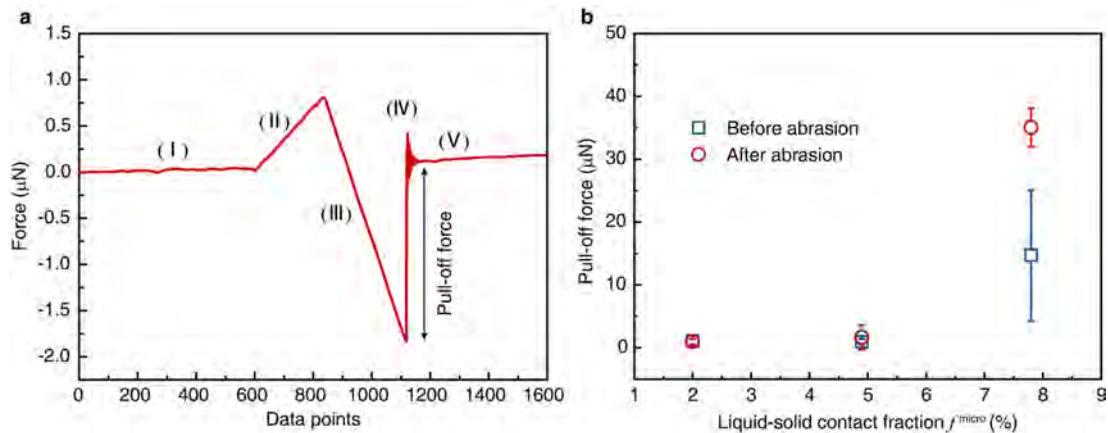
**Competing interests** X.D., D.W. and R.H.A.R. are inventors on a patent application relating to the surfaces described in this work.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2331-8>.

**Correspondence and requests for materials** should be addressed to R.H.A.R. or X.D.

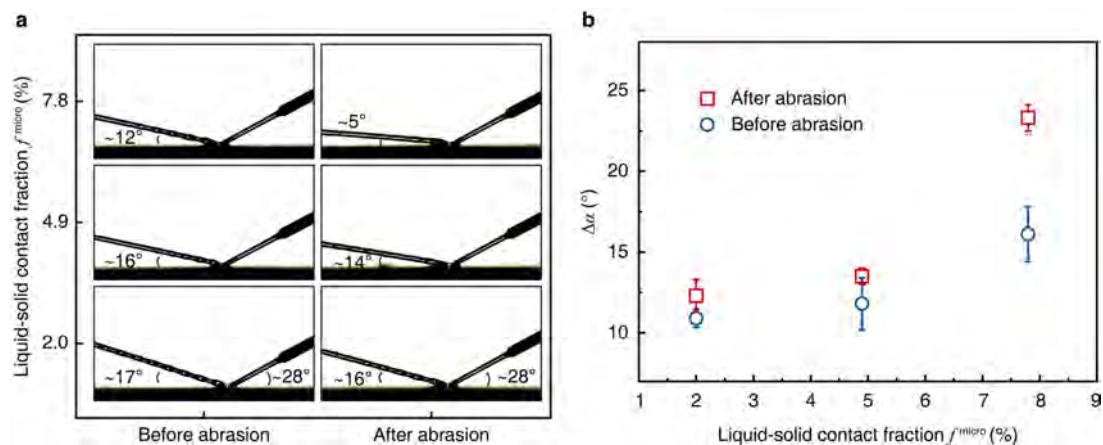
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**Extended Data Fig. 1 | Scanning droplet adhesion microscopy.** **a**, Droplet adhesion force curve recorded during scanning droplet adhesion microscopy<sup>31</sup> sampling of the superhydrophobic surface. Before making contact with the surface, the droplet is freely suspended from the disc (I). After contact is established, the force begins to increase steeply, indicating that the droplet is being pushed against the surface (II). After the peak force, the stage is retracted; the force becomes negative as the liquid meniscus is elongated (III).

Ultimately, the droplet detaches from the test surface, which induces a brief period of oscillations (IV). The pull-off force is defined as the difference between the baseline force after detachment and the global minimum of the curve (V). **b**, Pull-off forces for water drops on superhydrophobic surfaces, before and after abrasion, as a function of  $f^{\text{micro}}$ . Pull-off forces were recorded from an area of  $2.0 \text{ mm} \times 2.0 \text{ mm}$ , with  $500\text{-}\mu\text{m}$  spacing between each measurement point. Data are mean  $\pm$  s.d. from 25 independent measurements.

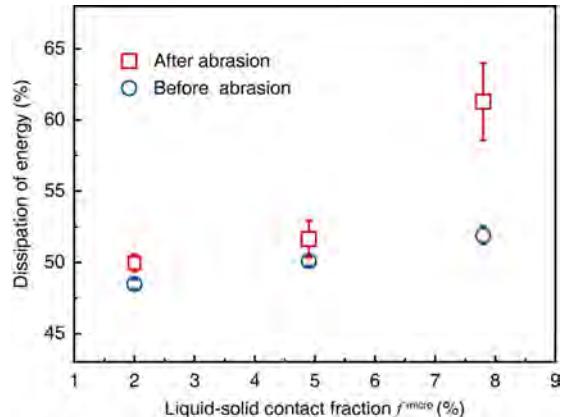
# Article



**Extended Data Fig. 2 | The impact of liquid jets on the hydrophobic surface.**

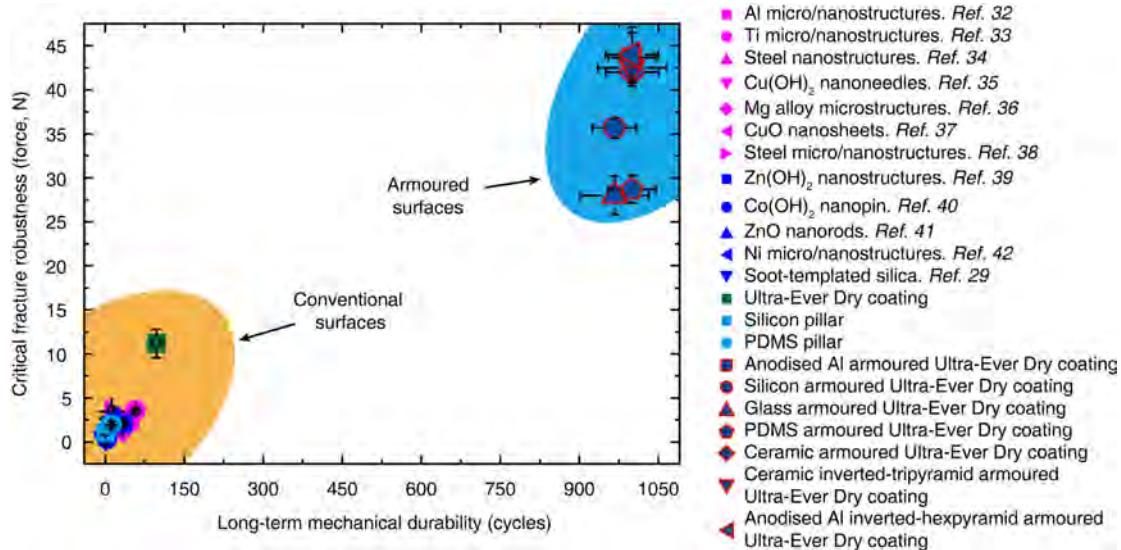
**a**, Deflection of water jets from the superhydrophobic surface before (left) and after (right) abrasion. **b**, The change of angle ( $\Delta\alpha$ ) when water jets (at a flow rate of approximately  $6 \text{ ml min}^{-1}$ ) are deflected by the superhydrophobic surfaces

before and after abrasion.  $\Delta\alpha = \alpha_{\text{incident}} - \alpha_{\text{deflected}}$ ;  $\Delta\alpha$  is plotted as a function of  $f^{\text{micro}}$ . Inverted-pyramidal structures ( $w + l = 60 \mu\text{m}$ ) on silicon substrates were used for this experiment. Data are mean  $\pm$  s.d. from at least five independent measurements.



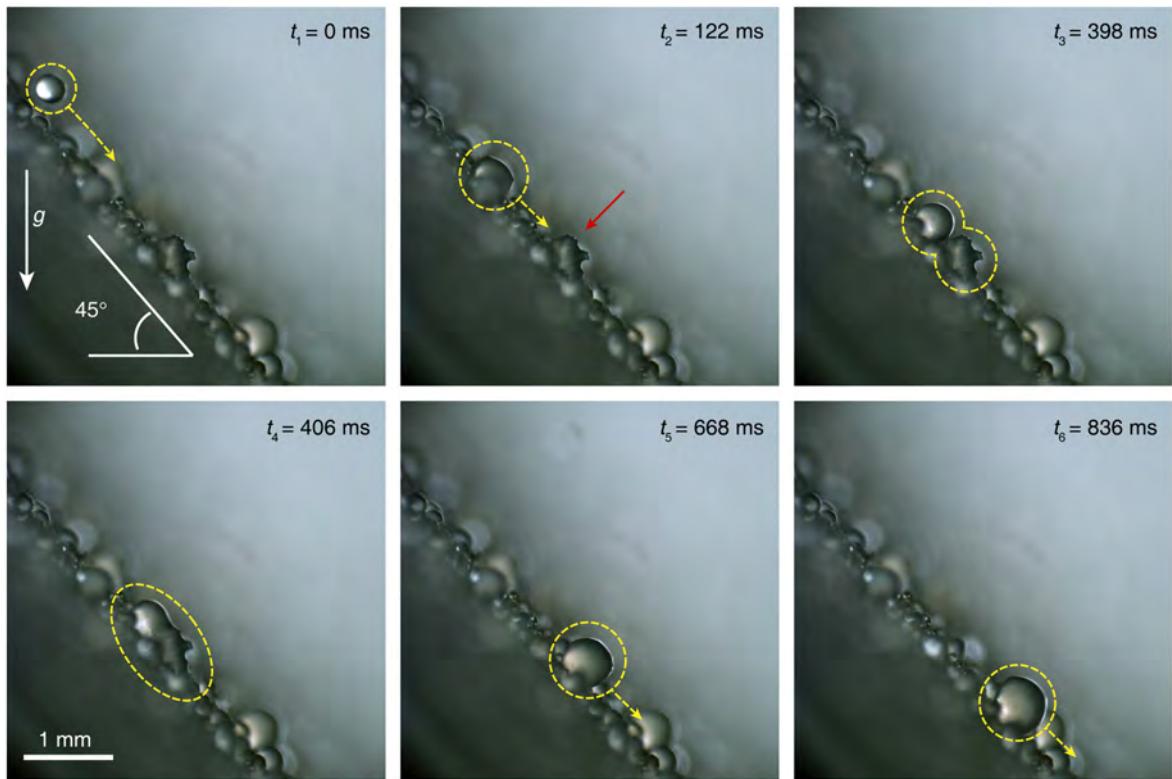
**Extended Data Fig. 3 | The impact of water droplets on the hydrophobic surface.**

Energy dissipation of the water drops (5.5  $\mu\text{l}$ , 14.0 mm of height) after impact on the armoured superhydrophobic surfaces before and after abrasion, plotted as a function of  $f^{\text{micro}}$ . Inverted-pyramidal microstructures ( $w+l=60 \mu\text{m}$ ) on silicon substrates were used for this experiment. Data are mean  $\pm$  s.d. from at least five independent measurements.



**Extended Data Fig. 4 | Comparison of the mechanical stability of various superhydrophobic surfaces.** Symbols of the same colour area indicate superhydrophobic surfaces with mechanical stabilities of the same order of magnitude. The x axis represents the resistance to linear abrasion by polypropylene with a normal load (3 N), the y axis represents the resistance to scratching by an alloy tip with specified normal load. Each data point

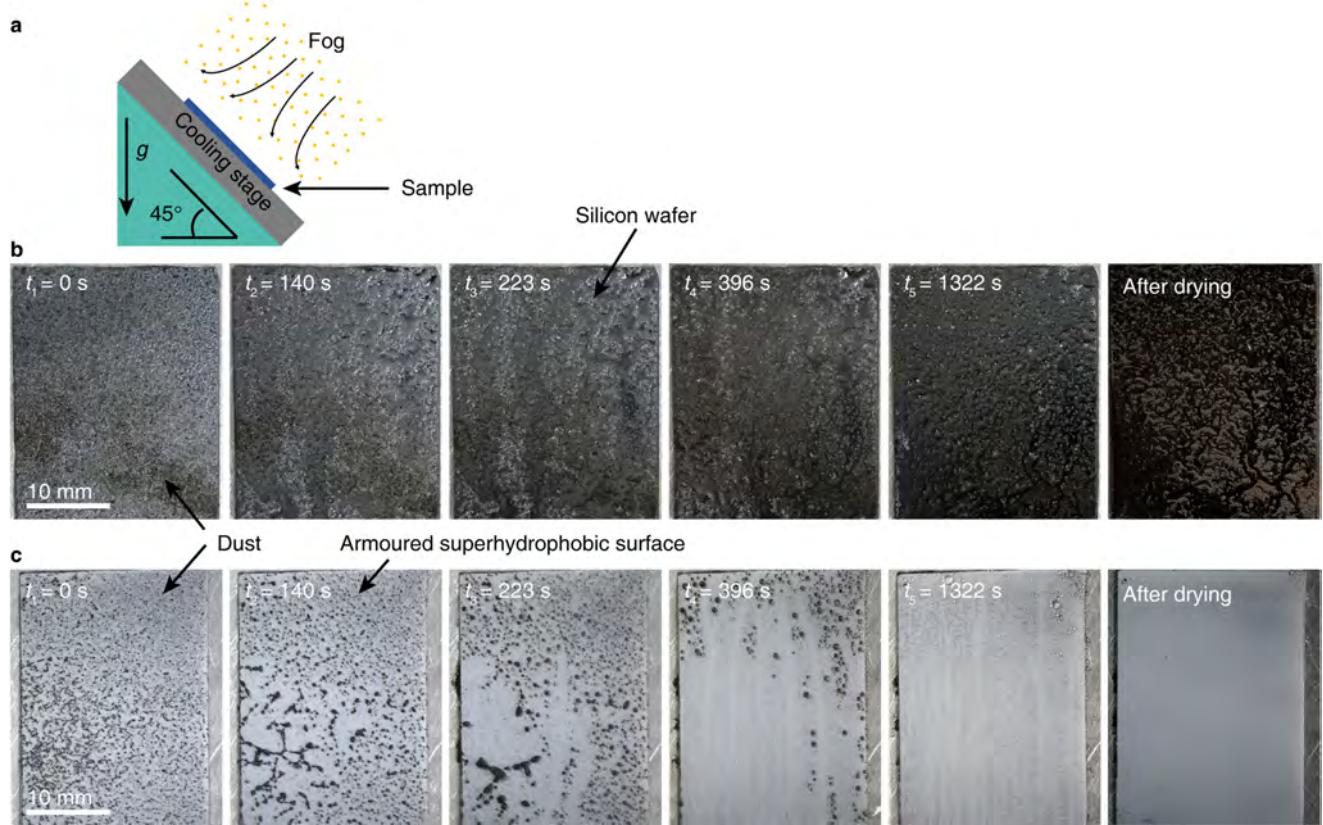
represents the capacity of the corresponding surface to maintain its superhydrophobicity under the above applied conditions. All samples, including those ones received from collaborators, were abraded in the same setup under similar conditions. Data are mean  $\pm$  s.d. from at least nine independent measurements. See refs. <sup>29,32–42</sup>.



**Extended Data Fig. 5 | Self-cleaning of the surface in fog.** The optical image sequences show the self-cleaning of the armoured superhydrophobic surface in fog. The red arrows indicate the dust particles (collected from air in Chengdu, China) on the surface. The yellow circle highlights how the droplet captures the dust particle and rolls off of the surface. The images were recorded with a Photron SA5 high-speed camera fitted with a macro lens at a

frame rate of 5,000 frames per second. The surface was placed on an aluminium alloy cold plate with a tilt angle of approximately 45°. The cold plate was controlled at around 2 °C, and the fog was at an ambient temperature with a high relative humidity of around 95%. The fog consisted of a cloud of water droplets (average radius 3.5 μm) suspended in air, generated using an ultrasonic humidifier.

## Article



**Extended Data Fig. 6 | Dust removal by self-cleaning.** **a**, Illustration showing the self-cleaning of the armoured superhydrophobic surface in fog. **b, c**, Sequences of photographs show the condensation of fog on a

dust-polluted planar silicon wafer (**b**) and on the dust-polluted armoured superhydrophobic surface (**c**). The dust was collected from air in Chengdu, China.

# Selective prebiotic formation of RNA pyrimidine and DNA purine nucleosides

<https://doi.org/10.1038/s41586-020-2330-9>

Received: 12 December 2019

Accepted: 16 April 2020

Published online: 3 June 2020

 Check for updates

Jianfeng Xu<sup>1,6</sup>, Václav Chmela<sup>1,6</sup>, Nicholas J. Green<sup>1</sup>, David A. Russell<sup>1</sup>, Mikołaj J. Janicki<sup>2</sup>, Robert W. Góra<sup>2</sup>, Rafał Szabla<sup>3,4</sup>, Andrew D. Bond<sup>5</sup> & John D. Sutherland<sup>1</sup>✉

The nature of the first genetic polymer is the subject of major debate<sup>1</sup>. Although the ‘RNA world’ theory suggests that RNA was the first replicable information carrier of the prebiotic era—that is, prior to the dawn of life<sup>2,3</sup>—other evidence implies that life may have started with a heterogeneous nucleic acid genetic system that included both RNA and DNA<sup>4</sup>. Such a theory streamlines the eventual ‘genetic takeover’ of homogeneous DNA from RNA as the principal information-storage molecule, but requires a selective abiotic synthesis of both RNA and DNA building blocks in the same local primordial geochemical scenario. Here we demonstrate a high-yielding, completely stereo-, regio- and furanosyl-selective prebiotic synthesis of the purine deoxyribonucleosides: deoxyadenosine and deoxyinosine. Our synthesis uses key intermediates in the prebiotic synthesis of the canonical pyrimidine ribonucleosides (cytidine and uridine), and we show that, once generated, the pyrimidines persist throughout the synthesis of the purine deoxyribonucleosides, leading to a mixture of deoxyadenosine, deoxyinosine, cytidine and uridine. These results support the notion that purine deoxyribonucleosides and pyrimidine ribonucleosides may have coexisted before the emergence of life<sup>5</sup>.

Considerable progress in the prebiotic synthesis of the pyrimidine ribonucleosides of RNA: cytidine (C) **1** and uridine (U) **2**, and their 2-thio derivatives: **3** and **4**<sup>6,7</sup>, together with recent advances in non-enzymatic RNA replication<sup>8–10</sup> have given credence to the ‘RNA world’ theory. Progress towards the abiotic synthesis of purine nucleosides has been made, but only using routes that use chemically and enantiomerically pure sugars as starting materials<sup>11–15</sup>, which were probably not found on the primordial earth. Additionally, no prebiotically plausible route has been shown to provide a mixture containing a competent set of nucleosides for information storage at the polymeric level.

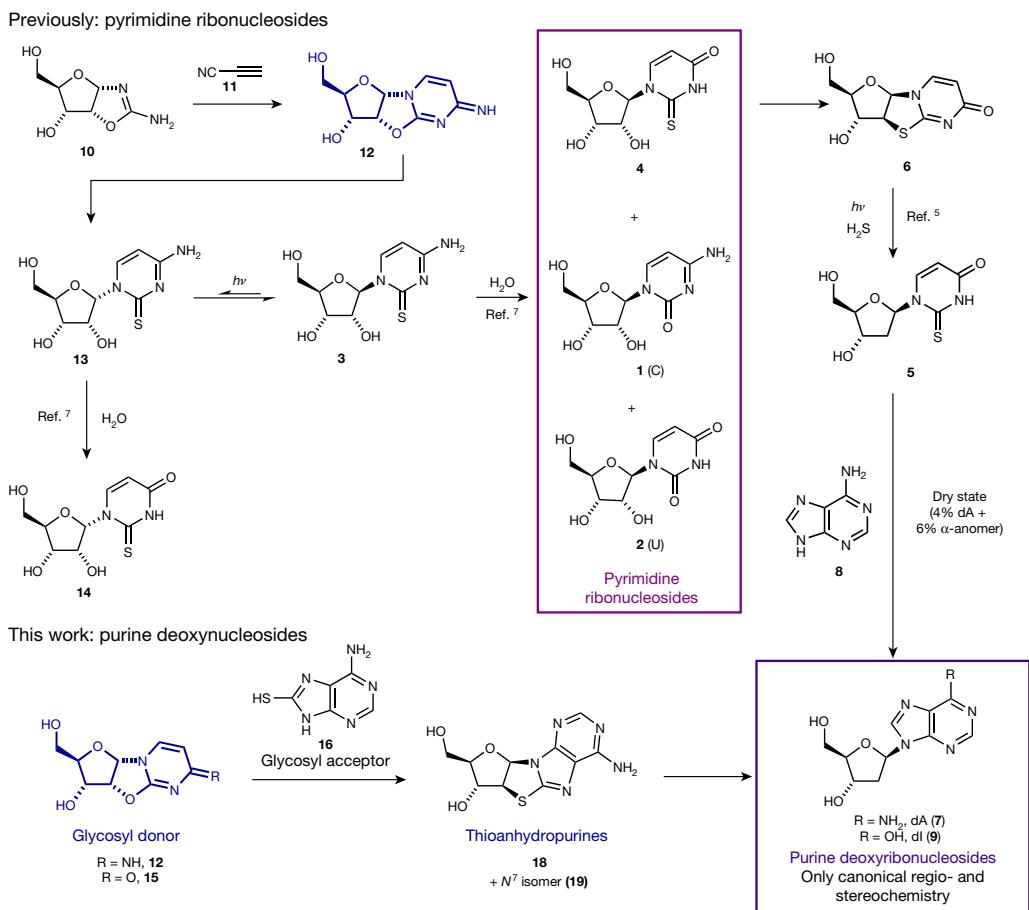
Extant biology, in contrast to the proposed RNA world theory, features DNA as the central information-carrying molecule. This discrepancy between the RNA world and modern biology requires a ‘genetic takeover’ that invokes the power of primitive biosynthetic machinery and natural selection operating over millions of years, ultimately resulting in an ancestral biosynthetic route to DNA<sup>16</sup>. The superior hydrolytic stability and replication fidelity<sup>17</sup> of DNA could have resulted in selection of primitive organisms capable of synthesizing DNA, and thus its rise to prominence, but the feasibility of this evolutionary process in a pre-DNA world is debated<sup>1</sup>. To circumvent this potentially problematic transition, an RNA/DNA world has been proposed, in which nascent biology had access to both RNA and DNA building blocks from the outset, without requiring elaborate biosynthesis<sup>18–20</sup>. In such a world, heterogeneous polymers would have initially been most common, but polymers with increased homogeneity—and hence properties closer to that of either RNA or DNA—would have been selected for over their mixed counterparts<sup>4</sup>.

For the RNA/DNA world to be plausible, an efficient prebiotic synthesis of DNA building blocks is required, and one that provides building blocks for both RNA and DNA in the same localized geochemical scenario is preferable. We recently demonstrated proof of this principle by showing that 2'-deoxy-2-thiouridine (**5**)—a non-canonical deoxynucleoside—can be synthesized from thioanhydrouridine (**6**)—an RNA derivative—by way of a prebiotically plausible hydrogen sulfide-mediated photoreduction<sup>5</sup>. Although this finding provides an important prebiotic link between RNA and DNA building blocks, the lability of **5** to hydrolysis may limit its phosphorylation and subsequent oligomerization<sup>21,22</sup>. Additionally, the synthesis of canonical deoxyadenosine (dA; **7**) from **5** and adenine (**8**) was low-yielding (4%), and generated a more abundant, undesired side product: the α-anomer of **7** (6%). Using guidance from a geochemical scenario<sup>23</sup>, we now demonstrate a synthesis of purine deoxynucleosides that is based on prebiotically plausible reactions and substrates. We then evaluate our route at a systems level by enacting the synthesis on mixtures of materials that might have arisen in a primordial environment, culminating in the demonstration of multiple reaction sequences able to selectively furnish a mixture of U (**1**), C (**2**), dA (**7**) and deoxyinosine (dI, **9**) (Extended Data Fig. 1).

## Prebiotic route to purine deoxyribonucleosides

A route to purine nucleosides that diverges from a prebiotic RNA synthesis is attractive because it implies that the constituents of a set of nucleosides capable of storing information—pyrimidines and

<sup>1</sup>MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Cambridge, UK. <sup>2</sup>Department of Physical and Quantum Chemistry, Faculty of Chemistry, Wrocław University of Science and Technology, Wrocław, Poland. <sup>3</sup>EastCHEM, School of Chemistry, University of Edinburgh, Edinburgh, UK. <sup>4</sup>Institute of Physics, Polish Academy of Sciences, Warsaw, Poland. <sup>5</sup>Department of Chemistry, University of Cambridge, Cambridge, UK. <sup>6</sup>These authors contributed equally: Jianfeng Xu, Václav Chmela. ✉e-mail: johns@mrc-lmb.cam.ac.uk



**Fig. 1 | Previous synthesis of RNA pyrimidine nucleosides C (1), U (2) and a deoxypyrimidine nucleoside (5), and the present work.** RAO (10) is a starting point in the network because it crystallizes in enantiopure form from minimally enantio-enriched solutions<sup>25,26</sup>. It can be elaborated via 12 and 3 to the pyrimidine nucleosides<sup>7</sup>. Although we have previously developed a low-yielding route to deoxyadenosine 7 (dA) from 6 via 5<sup>5</sup>, we recognized that 12 and 15 are ideal candidates for tethered glycosylation with 16. The products,

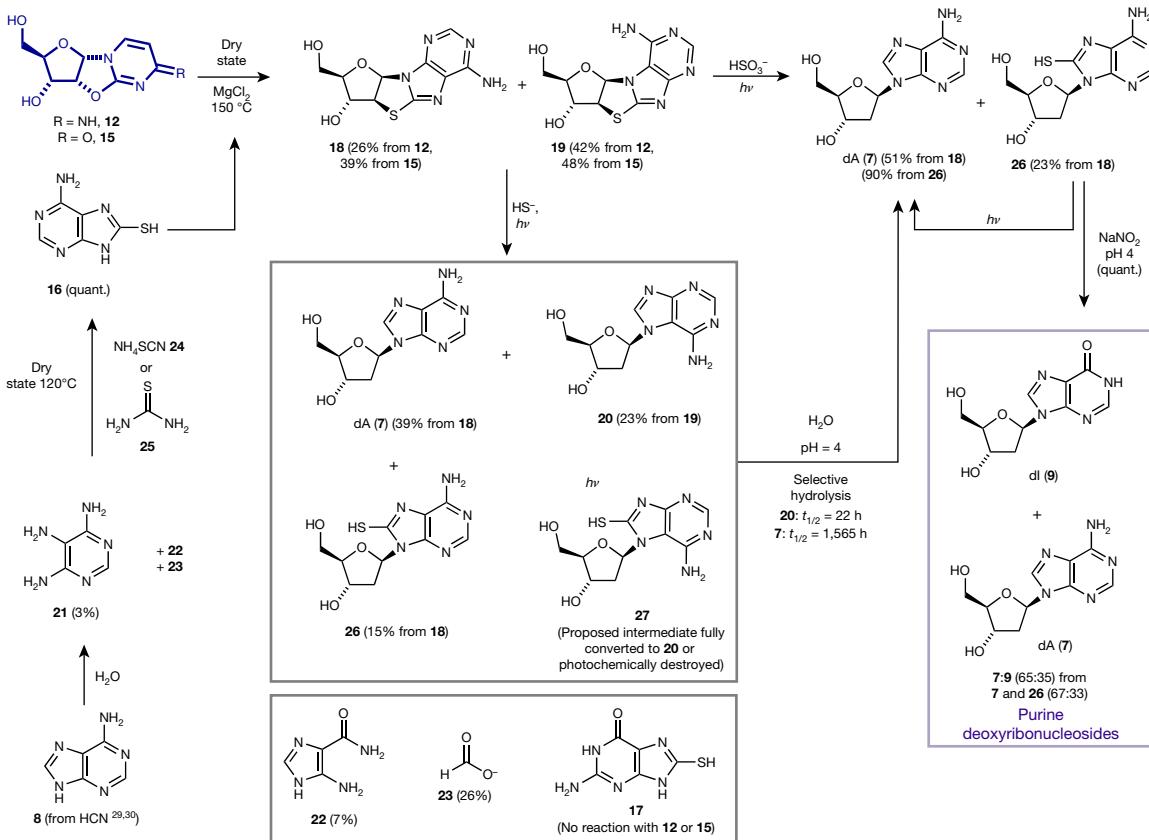
thioanhydropurines 18 and 19, are reduced photochemically in a similar way to 6, providing an efficient route to deoxynucleosides. Critically, once produced, pyrimidines 1(C) and 2(U) survive the sequence that produces purines 7 (dA) and 9 (dl), and we show that the four nucleosides 1(C), 2(U), 7 (dA) and 9 (dl) can be produced alongside one another. v, frequency of UV irradiation, centred at 254 nm.

purines—may have formed in the same location on a primordial Earth, and not necessarily brought together by environmental processes after their separate formation. To develop such a route, we evaluated intermediates in the prebiotic RNA pyrimidine nucleoside synthesis<sup>6,7</sup> as ribosyl donors (Fig. 1). The RNA synthesis proceeds from ribo-aminoazoline (RAO, 10), which reacts with cyanoacetylene (11) to provide α-anhydrocytidine (12). Thiolysis of 12 in formamide produces α-2-thiocytidine (13), which undergoes efficient ultraviolet (UV)-mediated photoanomerisation to 2-thiocytidine (3), which hydrolyses to the canonical pyrimidines cytidine (1) and uridine (2), and the biologically important non-canonical pyrimidine (4). Alternatively, in the dark, 13 is hydrolysed to α-2-thiouridine (14)<sup>7</sup>. Although 14 was initially viewed as only a by-product that would be produced in the dark on the early Earth, it is readily cyclized to anhydrouridine (15) at 80 °C (63% yield in water or 89% yield in formamide; Extended Data Fig. 2). We recognized α-anhydropyrimidines (12 and 15) as ideal glycosyl donors for 1',2'-cis tethered glycosylation<sup>24</sup>. The sugar of 12 and 15 is fixed in its furanosyl form, and so the formation of pyranosyl nucleosides—one of the weaknesses of previous strategies—should be excluded. Additionally, the α-stereochemistries of C1' and C2' of 12 and 15 led us to expect transglycosylation to provide only β-anomers, the correct stereochemistry at C1' for all natural (deoxy)ribonucleosides. Finally, because 12 and 15 are ultimately derived from RAO

(10)—which crystallizes enantiopure from solutions of minimally enantioenriched carbohydrates or amino acids<sup>25,26</sup>—this route offered the as-yet-unmet potential to deliver enantio- and diastereomerically pure furanosyl-nucleosides by glycosylation.

Accordingly, we evaluated 8-mercaptopadenine (16) and 8-mercaptopguanine (17) as potential nucleophiles to participate in trans-glycosylation with 12 and 15 (Fig. 2). Although 17 proved unreactive, 16 reacts with 12 and 15 at 150 °C in the dry state (Fig. 2) to provide two new β-configured nucleoside products in moderate yields (respectively, 14% and 16% from 15, trace amounts from 12). The minor product was determined to be N<sup>7</sup>-8,2'-anhydro-thioadenosine (18) by X-ray crystallography and <sup>1</sup>H NMR spiking experiments with a synthetic standard. The major product was inferred to be N<sup>7</sup>-8,2'-anhydro-thioadenosine (19)—the regioisomer of 18—by its subsequent conversion to 2'-deoxy-N<sup>7</sup>-adenosine (20). The presence of magnesium chloride in the reaction—presumably acting as a Lewis acid<sup>27</sup>—dramatically improved the yield of 18 and 19 to 39% and 48% respectively from 15 (combined yield 87%), and 26% and 42% respectively from 12 (combined yield 68%). Thus, in a prebiotic environment where 12 or 15, and 16 are brought together, perhaps by converging streams that then undergo evaporation, 18 and 19 could be readily generated, especially in the presence of magnesium ions<sup>28</sup>.

Any prebiotic synthesis requires a viable route to all reagents from plausible early-Earth feedstocks. We considered adenine (8) as a



**Fig. 2 | Prebiotic route to purine deoxyribonucleosides, 7 (dA) and 9 (dl).**

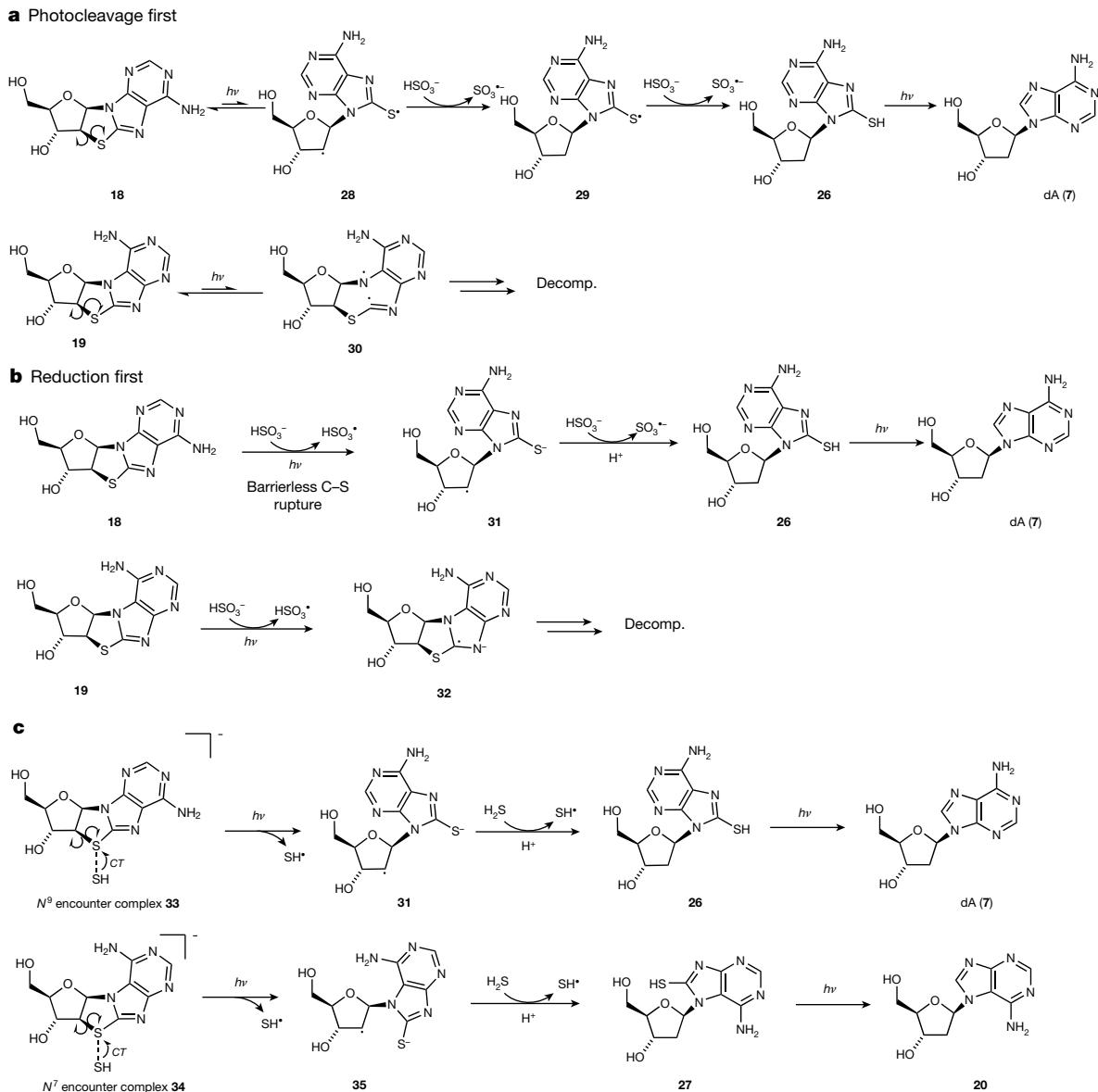
The route starts with  $\alpha$ -anhydropurines **12** and **15**, which are intermediates in the RNA pyrimidine synthesis, and 8-mercaptopadenine (**16**), which is available from adenine (**8**) via hydrolysis and reaction with ammonium thiocyanate or thiourea. Dry-state tethered glycosylation of **16** and **12** or **15** provides thioanhydropurines **18** and **19**, which can be photochemically reduced by two routes. If bisulfite is the reductant, only  $N^{\circ}$ -configured

products **7** (dA) and **26** are formed. **26** can be converted to **7** (dA) by further irradiation or by nitrosation. If hydrosulfide is used as the reductant, both  $N^{\circ}$ -configured **7** (dA) and **26** as well as  $N^7$ -configured **20** is formed. **20** has a half-time of hydrolysis nearly two orders of magnitude lower than that of **7** (dA) and so is selectively degraded. To generate **9** (dl) alongside **7** (dA), the products of either photoreduction are treated with nitrous acid at pH 4.

starting point for the provision of 8-mercaptopadenine (**16**) owing to its widely accepted prebiotic plausibility as a relatively stable pentamer of hydrogen cyanide<sup>29,30</sup>. Notably, despite the reactivity of related purines<sup>31</sup>, adenine did not react with elemental sulfur at temperatures up to 300 °C. However, adenine does undergo (slow) hydrolysis in aqueous media. Previous reports provided a half-time for hydrolysis of adenine of about one year at 100 °C, and identified (but did not quantify) 4,5,6-triaminopyrimidine (TAP, **21**) among the products of hydrolysis<sup>32</sup>. We reinvestigated this hydrolysis of **8** under conditions more suited to a laboratory timescale (138 °C, phosphate buffer pH 8), and at partial conversion after 10–12 d we confirmed the presence of TAP in yields of 2–3% (8–9% based on recovered adenine) (Fig. 2). Owing to the differential solubilities of adenine and TAP, the supernatants of adenine hydrolysis reactions are enriched in TAP after cooling. A typical supernatant contains 5-aminoimidazole-4-carboxamide (**22**), TAP (**21**) and **8** in a 4:2:1 ratio, and formate (**23**) as the only other major component (see Supplementary Figs. 1–5 for full details). We found that TAP (either commercially supplied or in the crude adenine hydrolysate) is converted to 8-mercaptopadenine (**16**) by heating in the dry state with either ammonium thiocyanate (**24**) or thiourea (**25**). **24** is an inevitable by-product of the photochemistry of hydrogen cyanide and hydrogen sulfide<sup>33</sup>, two precursors that were probably abundant on the primordial earth, and that are heavily implicated in the origin of life by our cyanosulfidic chemical network<sup>23</sup>. **25** has also been

widely invoked as a prebiotically plausible reagent<sup>34</sup>. Thus, we envision that a primordial environment supplied with adenine and water would continuously generate TAP, which can be enriched in aqueous solution by moving down a thermal gradient. **24** can be mixed with TAP at any stage, and eventual evaporation and dry-state reaction leads to **16**. This method of accumulation of TAP also improves the plausibility of some intermediates of other prebiotic syntheses<sup>12</sup>.

With thioanhydropurine nucleosides (**18** and **19**) in hand, we moved on to evaluate their photoreduction chemistry to see if we might directly generate deoxyadenosine. Our previous synthesis of a deoxypyrimidine via a thioanhydropurine (**6**) (Fig. 1) proceeded by the reduction of a C–S to a C–H bond mediated by a hydrated electron, generated by UV irradiation of hydrosulfide<sup>5,33</sup>. **18** and **19** were separately subjected to UV irradiation at 254 nm in water with hydrogen sulfide (H<sub>2</sub>S) as the reductant (Fig. 2). In the photoreduction of **18**, the natural regio-isomer dA (**7**) was detected in 39% yield, along with 15% of 8-mercaptop-deoxyadenosine (**26**). **26** was demonstrated to be a competent intermediate in the reaction by desulfurization to give dA (**7**) either by UV irradiation<sup>35</sup> or treatment with nitrous acid, which is produced from common atmospheric gases, nitrogen and carbon dioxide<sup>36</sup>. Nucleobase loss was also apparent (**16** in 10% yield and **8** in 17% yield). The same reaction starting with **19** gave  $N^7$ -deoxyadenosine (**20**) in 23% yield with no other nucleoside products. Our proposed intermediate in this process, 8-mercato- $N^7$ -deoxyadenosine (**27**), is either fully converted to **20** or photochemically destroyed.



**Fig. 3 | Proposed mechanism of photoreduction of N<sup>7</sup>-8,2'-anhydro-thioadenosine (18) and N<sup>9</sup>-8,2'-anhydro-thioadenosine (19) nucleosides.** **a**, A potential mechanism involving bisulfite, proceeding with initial photoexcitation of the thioanhydronucleosides to **28**, followed by reduction of C2', sulfur and C8. Photoexcitation of the N<sup>7</sup> isomer **19** to **30** leads to decomposition. **b**, A potential mechanism involving bisulfite, proceeding via intial reduction of ground-state thioanhydronucleosides, followed by

desulfurisation of **26**. Reduction of **19** gives **32**, which leads to decomposition. **c**, A distinct mechanism involving reduction of thioanhydronucleoside–hydrosulfide encounter complexes, **33** and **34**, which both undergo charge transfer and concomitant C–S bond cleavage to produce **31** and **35**. **31** and **35** undergo reduction at C2' and desulfurisation to furnish **7** (dA) and **20**. CT, charge transfer; Decomp., decomposition.

Photoreduction was also carried out on a mixture of **18** and **19** compatible with our synthesis by tethered transglycosylation. The ratio of N<sup>9</sup>:N<sup>7</sup> regioisomers was increased from 38:62 of **18:19** in the starting mixture to 56:44 of **7:20** after photoreduction (31% yield for **7**, 17% yield for **20**), indicating, compared to N<sup>7</sup> isomers, an enhanced stability of intermediates or products bearing the natural N<sup>9</sup> glycosidic linkage. Replacing hydrosulfide as the electron donor with bisulfite (HSO<sub>3</sub><sup>-</sup>, pH 7)<sup>37</sup>, which is readily formed by the dissolution of atmospheric SO<sub>2</sub> in water<sup>38</sup>, improved both the yield and selectivity of the photoreduction. Photoreduction with bisulfite of **18** alone provided dA (**7**) in 51% yield and **26** in 23% yield, whereas a similar reaction with the N<sup>7</sup>-regio-isomer (**19**) led only to its photochemical destruction. Photoreduction of a mixture of **18** and **19** with bisulfite led only to N<sup>9</sup>-linked products: dA (**7**) and **26** in 44% and 18% yield, respectively (Extended Data Fig. 3). Separate experiments probing the stability of starting materials and

products under the reaction conditions indicated that the relative stabilities of intermediates are the cause of this selectivity. This strikingly selective destruction is highly suggestive of a potential mechanism by which primordial nucleosides may have been restricted to a near-canonical set<sup>39,40</sup>. We found further evidence for such restriction in the hydrolysis rates of the N<sup>9</sup> and N<sup>7</sup> isomers of deoxyadenosine. In acetate buffer (pH 4, 25 °C), the natural isomer dA (**7**) is more than 70 times more stable than **20** (half-lives of 1,565 and 22 h, respectively), which is consistent with the reported difference in stabilities towards acid hydrolysis between the corresponding isomers of adenosine<sup>41,42</sup>.

## Photoreduction mechanism

To provide a mechanistic rationale for the observed photochemical selectivity, we performed quantum chemical calculations using density

functional theory and algebraic diagrammatic construction to the second order (ADC(2)) methods<sup>43,44</sup>. These calculations revealed, in the case of bisulfite, two possible competing mechanisms that explain the difference in reactivity of the two regioisomers. **18** and **19** can both undergo photoexcitation, but generate dissimilar biradical species (Fig. 3a). Photoexcitation of **18** leads to rupture of the C2'-S bond on the surface of the lowest excited singlet ( $S_1$ ) state, generating biradical **28** (Fig. 3a,  $N^7$ ; Extended Data Fig. 4a). Reduction of this species by intermolecular hydrogen atom transfer (HAT) or proton-coupled electron transfer (PCET) probably leads to the C2'-reduced species **29**, and ultimately—via a second HAT or PCET and subsequent photolysis of the C8-S bond of **26**<sup>35</sup>—to dA (7) (Fig. 3a,  $N^9$ ). By contrast, photoexcitation of **19** leads to N7-C8 bond rupture through the  $S_1/S_0$  state crossing (Fig. 3a,  $N^7$ ; Extended Data Fig. 4b), generating **30**, which probably undergoes decomposition without C2'-S reduction. Bisulfite is well-known to provide a hydrated electron upon irradiation<sup>45</sup>, and so a second possibility is the reduction by hydrated electrons of **18** and **19** in the ground state. Again, calculations suggest different fates of **18** and **19** upon reduction. Reduction of **18** is predicted to proceed with concomitant barrierless C2'-S bond rupture to give the radical anion intermediate **31** (Fig. 3b,  $N^9$ ; Extended Data Fig. 5), whereas reduction of **19** is predicted to lead to formation of a C8, N9 radical anion (**32**) which also probably undergoes decomposition instead of C2' reduction (Fig. 3b  $N^7$ ; Extended Data Fig. 5). In the absence of any reducing agent, both **18** and **19** undergo (equally) slow photochemical decomposition, presumably via the calculated biradical structures **28** and **30**, but in the presence of bisulfite, reduction of the ground state or photochemically generated intermediates results in remarkably different fates.

The successful reduction of **19** alongside **18** when using hydrosulfide as the reducing agent is explained by a distinct mechanism. Calculations located stable encounter complexes, **33** and **34**, between HS<sup>-</sup> and thioanhydronucleosides **18** and **19**, respectively (Fig. 3c, Extended Data Fig. 4c, d). This interaction is predominantly stabilized by electrostatic and dispersion interactions and our interaction energy decomposition demonstrates its stability in aqueous solution (see Supplementary Information for details). Similar S···S interactions were recently identified in intramolecular complexes and classified as chalcogen bonds<sup>46</sup>. Such an encounter complex facilitates charge transfer from the hydrosulfide anion to the thioanhydropurine fragment almost immediately after UV absorption by the complex to the  $S_1$  state. Subsequent relaxation on the  $S_1$  surface enables practically barrierless C2'-S bond breaking completed by a peaked  $S_1/S_0$  state crossing for both intermediates **31** and **35**, thus facilitating C2'-S reduction of both **18** and **19** (Extended Data Fig. 4c, d). The products of this photochemical transformation, **26** and **27**, may further undergo photochemical sulfur cleavage through the mechanism described previously<sup>35</sup> (Fig. 3c). Thus, a HS<sup>-</sup> thioanhydropurine encounter complex facilitates C-S bond cleavage and partially protects  $N^7$  isomer **19** from the photodestruction observed in the presence of bisulfite. This finding not only explains the distinctive outcomes of photoreduction between the two reducing agents, but also points towards a potentially important stabilizing role for hydrosulfide in prebiotic chemistry and photochemistry in general.

### Prebiotic route to a purine/pyrimidine genetic system

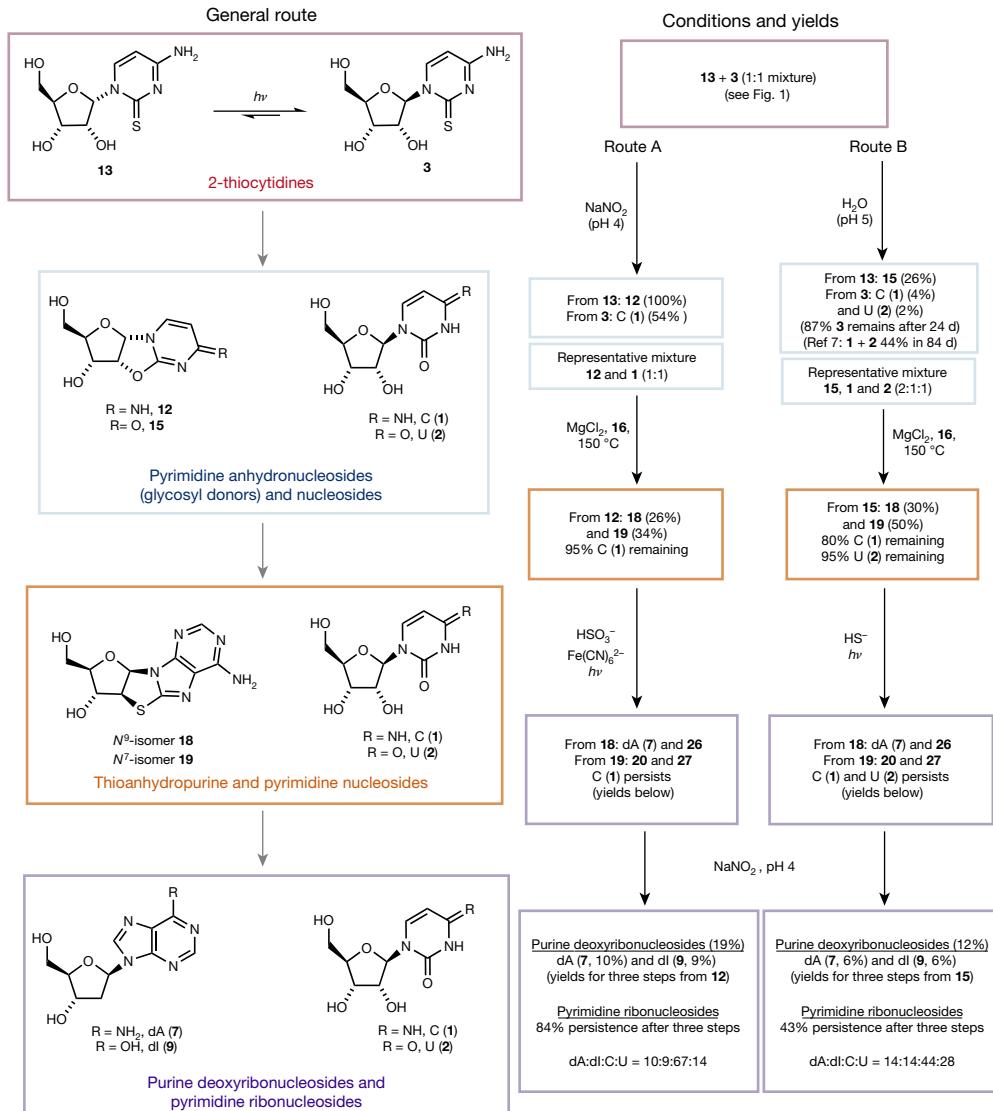
Although our attempts to glycosylate 8-mercaptoguanine (**17**) to provide thioanhydroguanosine (and ultimately deoxyguanosine) failed, the triple selectivity and high yield of our route to deoxyadenosine combined with recent results<sup>47</sup> suggest a possible alternative genetic alphabet that does not include (deoxy)guanosine. Guanosine is yet to be accounted for in a plausible prebiotic synthesis, but it has recently been shown<sup>47</sup> that inosine (**I**), which is capable of base-pairing with cytosine, can replace guanosine in non-enzymatic RNA replication systems with no loss of rate or fidelity. **7** (dA) is readily converted to **9** (dI) (Fig. 2) by deaminative hydrolysis, which spontaneously occurs

very slowly in nucleic acid polymers<sup>48</sup>, and is greatly accelerated by the presence of nitrous acid<sup>49</sup>. To demonstrate that this conversion can occur under mild conditions consistent with our primordial geochemical scenario<sup>50</sup>, we treated **7** (dA) with nitrous acid at pH 4 (the same conditions in which we were able to effect desulfurization of **26**). After four days at room temperature, approximately 40% of **7** (dA) had been converted to **9** (dI), providing a 60:40 mixture of **7** (dA) and **9** (dI) (Fig. 2). A control experiment monitoring the decomposition of deoxyadenosine **7** (dA) at pH 4 without nitrous acid showed only a trace of depurination (half-life,  $t_{1/2} = 1,600$  h). When a 67:33 mixture of **7** and **26** (representative of the outcome of photoreduction) was submitted to the reaction conditions, **26** underwent relatively rapid desulfurization first, with deoxyadenosine **7** (dA) undergoing slower deaminative hydrolysis to ultimately provide a 65:35 mixture of **7** (dA) and **9** (dI). Thus, mixtures of **7** (dA) and **9** (dI) are readily obtainable from the partial deaminative hydrolysis of **7** (dA) or its precursor **26**, thereby supplying half of a potential primordial alphabet. Despite the potential for a mismatch in reactivity between deoxypurines and pyrimidines, a 47:53 mixture of **7** (dA) and cytidine (**1**; C) underwent nitrous acid-promoted deamination to provide all four (deoxy)nucleosides **7** (dA), **9** (dI), **1** (C) and uridine (**2**, U) in a 30:17:42:11 ratio (Extended Data Fig. 6). A similar primordial mixture may have been a starting point for genetic information storage. Furthermore, in the absence of geochemically plausible sources of pyrimidine deoxynucleotides and purine ribonucleotides, heteropolymers made from a mixture of purine deoxyribonucleotides and pyrimidine ribonucleotides should possess heritable backbone heterogeneity and thus a 1:1 phenotype-to-genotype correspondence, which is potentially advantageous in the evolution of catalytic activity<sup>18</sup>.

### Systems-level prebiotic plausibility

Having demonstrated the potential of a divergent route to yield a local mixture of **7** (dA), **9** (dI), **1** (C) and **2** (U), we sought to evaluate the key question of whether all four nucleosides could persist after divergence in the sequence. We chose a 1:1 mixture of  $\alpha$ - and  $\beta$ -2-thiouridines (**13** and **3**) as our starting point, obtained from the partial photoanomerisation of **13**, and evaluated two particular combinations of reactions as representative permutations of a primordial geochemical process (Fig. 4, routes A and B). In route A, exposure of the mixture to nitrous acid (pH 4) generates a mixture of **12** and **1** (100% yield for **12** from **13**, 54% yield for **1** from **3**). **12** is formed from **13** potentially by intramolecular addition of the C2' hydroxyl to C2 of an S-nitrosoyl intermediate, and subsequent elimination of SNO<sup>-</sup>. Dry-state glycosylation of **16** and a 1:1 mixture of **12** and **1** (C) in the presence of MgCl<sub>2</sub> leads to a mixture of **18** and **19** as described in our route development above; however, critically, 95% of **1** persists in this mixture. Subsequent photoreduction in the presence of ferrocyanide and bisulfite generates the expected mixture of purine nucleosides **7** (dA), **26**, **20** and **27** alongside **1** (C). Finally, a second exposure to nitrous acid converts this mixture into the components of a competent genetic system, **7** (dA), **9** (dI) (10% and 9% yield respectively from **12** for three steps), **1** (C) and **2** (U) (84% combined persistence after three steps) with no major nucleoside impurities. Products derived from **19**—with the wrong  $N^7$  regiochemistry—are hydrolysed in the last step. It is noteworthy that this route is only viable from a systems-level approach—for instance, the pyrimidines are fairly rapidly destroyed in the photoreduction step in the absence of the thioanhydropurines (Extended Data Fig. 7).

Route B presents an alternative in which initial hydrolysis of the mixture of **13** and **3** generates glycosyl donor **15** (26% yield) alongside the pyrimidine nucleosides (4% of **1** (C), 2% of **2** (U), 92% of **3** remaining). **3** has previously been shown to hydrolyse to **1** (C) and **2** (U) in greater yields (44%) over longer periods<sup>7</sup>. A representative mixture of **15**, **1** (C) and **2** (U) (2:1:1) was then subjected to tethered glycosylation, resulting in **18** and **19** as above (30% and 50% yield respectively), with 80% and 95% persistence of **1** (C) and **2** (U), respectively. Photoreduction of the mixture—this time with hydrogen sulfide—provides purine



**Fig. 4 | A systems-level approach to a potential primordial genetic alphabet composed of 1(C), 2(U), 7(dA) and 9(dl).** A mixture of the  $\alpha$ - and  $\beta$ -epimers of the 2-thiocytidines 13 and 3 (which interconvert in UV light), can generate a mixture containing 1(C), 2(U), 7(dA) and 9(dl). A general route is shown at left. The thiopyrimidines are initially converted into the canonical pyrimidines (cytidine (1) and uridine (2)) and the  $\alpha$ -anhydropurimidines 12 and 15. The latter

undergo tethered glycosylation and then photoreduction to selectively provide purine deoxyribonucleosides 7 (dA) and 9 (dl) as depicted in Fig. 2. The pyrimidines 1(C) and 2(U) persist through each step of this sequence, ultimately generating a mixture of all four nucleosides. Specific conditions and yields for two possible particular routes (routes A and B) are shown at right.

products 7, 26, 20 and 27 alongside the pyrimidines 1(C) and 2(U). Finally, nitrosation furnished the key mixture of 7 (dA) and 9 (dl) (6% for each from 15 for three steps) alongside pyrimidine nucleosides (43% persistence over three steps; the final ratio of dA:dl:C:U in the mixture is 14:13:45:28, Extended Data Fig. 8). Thus, sequences comprised of various orders of operations and various photoreduction conditions—which might plausibly emulate a terrestrial geochemical scenario—simultaneously generate the components of a mixed genetic system. The exact ratio of 1(C) and 2(U) (ribosylpyrimidines) to 7 (dA) and 9 (dl) (deoxyribosylpurines) in the final mixture will depend on the ratio of  $\alpha$ -(anhydro)pyrimidines (13, 12 and 15) to  $\beta$ -(thio)pyrimidines (1, 2 and 3) earlier in the sequence, which in turn varies with environmental conditions.

In conclusion, a highly efficient synthesis of both deoxyadenosine 7 (dA) and deoxyinosine 9 (dl), requiring only prebiotically plausible reagents and conditions, is reported. In contrast to all previous attempts to synthesize purine nucleosides, our synthesis is both prebiotically plausible and strictly stereo-, regio- and furanosyl-selective for the only

isomer of the deoxypurine nucleosides used in modern biology. The pathway proceeds mostly via simple hydrolysis or dry-state processes, with a key reduction step promoted by UV irradiation supported by distinct mechanisms. The (photo)chemical selection exhibited by this route hints at an explanation for the biological importance of one isomer of nucleic acid from the many that are conceivable. We have demonstrated that co-occurring sequences leading selectively to both RNA pyrimidine and DNA purine nucleosides can result in mixtures that could conceivably complete a genetic alphabet. DNA building blocks can thus be co-produced with the RNA pyrimidine nucleosides—which is consistent with, and perhaps evidence for, the coexistence of RNA and DNA building blocks at the dawn of life.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions

# Article

and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2330-9>.

1. Samanta, B. & Joyce, G. F. A reverse transcriptase ribozyme. *eLife* **6**, e31153 (2017).
2. Gilbert, W. Origin of life: the RNA world. *Nature* **319**, 618 (1986).
3. Joyce, G. F. The antiquity of RNA-based evolution. *Nature* **418**, 214–221 (2002).
4. Bhowmik, S. & Krishnamurthy, R. The role of sugar-backbone heterogeneity and chimeras in the simultaneous emergence of RNA and DNA. *Nat. Chem.* **11**, 1009–1018 (2019).
5. Xu, J., Green, N. J., Gibard, C., Krishnamurthy, R. & Sutherland, J. D. Prebiotic phosphorylation of 2-thiouridine provides either nucleotides or DNA building blocks via photoreduction. *Nat. Chem.* **11**, 457–462 (2019).
6. Powner, M. W., Gerland, B. & Sutherland, J. D. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459**, 239–242 (2009).
7. Xu, J. et al. A prebiotically plausible synthesis of pyrimidine β-ribonucleosides and their phosphate derivatives involving photoanomerization. *Nat. Chem.* **9**, 303–309 (2017).
8. Heuberger, B. D., Pal, A., Del Frate, F., Topkar, V. V. & Szostak, J. W. Replacing uridine with 2-thiouridine enhances the rate and fidelity of nonenzymatic RNA primer extension. *J. Am. Chem. Soc.* **137**, 2769–2775 (2015).
9. Walton, T. & Szostak, J. W. A highly reactive imidazolium-bridged dinucleotide intermediate in nonenzymatic RNA primer extension. *J. Am. Chem. Soc.* **138**, 11996–12002 (2016).
10. Li, L. et al. Enhanced nonenzymatic RNA copying with 2-aminoimidazole activated nucleotides. *J. Am. Chem. Soc.* **139**, 1810–1813 (2017).
11. Fuller, W. D., Orgel, L. E. & Sanchez, R. A. Studies in prebiotic synthesis: VI. Solid-state synthesis of purine nucleosides. *J. Mol. Evol.* **1**, 249–257 (1972).
12. Becker, S. et al. A high-yielding, strictly regioselective prebiotic purine nucleoside formation pathway. *Science* **352**, 833–836 (2016).
13. Kim, H. & Benner, S. A. Prebiotic stereoselective synthesis of purine and noncanonical pyrimidine nucleotides from nucleobases and phosphorylated carbohydrates. *Proc. Natl. Acad. Sci. USA* **114**, 11315–11320 (2017).
14. Becker, S. et al. Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. *Science* **366**, 76–82 (2019).
15. Teichert, J. S., Kruse, F. M. & Trapp, O. Direct prebiotic pathway to DNA nucleosides. *Angew. Chem. Int. Ed.* **58**, 9944–9947 (2019).
16. Reichard, P. From RNA to DNA, why so many ribonucleotide reductases? *Science* **260**, 1773–1777 (1993).
17. Leu, K., Obermayer, B., Rajamani, S., Gerland, U. & Chen, I. A. The prebiotic evolutionary advantage of transferring genetic information from RNA to DNA. *Nucleic Acids Res.* **39**, 8135–8147 (2011).
18. Sutherland, J. D. & Whitfield, J. N. Prebiotic chemistry: a bioorganic perspective. *Tetrahedron* **53**, 11493–11527 (1997).
19. Trevino, S. G., Zhang, N., Elenko, M. P., Lupták, A. & Szostak, J. W. Evolution of functional nucleic acids in the presence of nonheritable backbone heterogeneity. *Proc. Natl. Acad. Sci. USA* **108**, 13492–13497 (2011).
20. Gavette, J. V., Stoop, M., Hud, N. V. & Krishnamurthy, R. RNA–DNA chimeras in the context of an RNA world transition to an RNA/DNA world. *Angew. Chem. Int. Ed.* **55**, 13204–13209 (2016).
21. Schöffstall, A. M. Prebiotic phosphorylation of nucleosides in formamide. *Orig. Life* **7**, 399–412 (1976).
22. Lohrmann, R. & Orgel, L. E. Urea-inorganic phosphate mixtures as prebiotic phosphorylating agents. *Science* **171**, 490–494 (1971).
23. Patel, B. H., Percivalle, C., Ritson, D. J., Duffy, C. D. & Sutherland, J. D. Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nat. Chem.* **7**, 301–307 (2015).
24. Ishiwata, A., Lee, Y. J. & Ito, Y. Recent advances in stereoselective glycosylation through intramolecular aglycon delivery. *Org. Biomol. Chem.* **8**, 3596–3608 (2010).
25. Springsteen, G. & Joyce, G. F. Selective derivatization and sequestration of ribose from a prebiotic mix. *J. Am. Chem. Soc.* **126**, 9578–9583 (2004).
26. Anastasi, C., Crowe, M. A., Powner, M. W. & Sutherland, J. D. Direct assembly of nucleoside precursors from two- and three-carbon units. *Angew. Chem. Int. Ed.* **45**, 6176–6179 (2006).
27. Vorbrüggen, H. & Ruh-Pohlenz, C. *Handbook of Nucleoside Synthesis* (Wiley, 2001).
28. Holm, N. G., Oze, C., Mousis, O., Waite, J. H. & Guillert-Lepoutre, A. Serpentization and the formation of H<sub>2</sub> and CH<sub>4</sub> on celestial bodies (planets, moons, comets). *Astrobiology* **15**, 587–600 (2015).
29. Sanchez, R. A., Ferris, J. P. & Orgel, L. E. Studies in prebiotic synthesis. II: Synthesis of purine precursors and amino acids from aqueous hydrogen cyanide. *J. Mol. Biol.* **80**, 223–253 (1967).
30. Hudson, J. S. et al. A unified mechanism for abiotic adenine and purine synthesis in formamide. *Angew. Chem. Int. Ed.* **51**, 5134–5137 (2012).
31. Giner-Sorolla, A., Thom, E. & Bendich, A. Studies on the thiation of purines. *J. Org. Chem.* **29**, 3209–3212 (1964).
32. Levy, M. & Miller, S. L. The stability of the RNA bases: implications for the origin of life. *Proc. Natl. Acad. Sci. USA* **95**, 7933–7938 (1998).
33. Ritson, D. J. & Sutherland, J. D. Synthesis of aldehydic ribonucleotide and amino acid precursors by photoredox chemistry. *Angew. Chem. Int. Ed.* **52**, 5845–5847 (2013).
34. Robertson, M. P., Levy, M. & Miller, S. L. Prebiotic synthesis of diaminopyrimidine and thiocytosine. *J. Mol. Evol.* **43**, 543–550 (1996).
35. Roberts, S. J. et al. Selective prebiotic conversion of pyrimidine and purine anhydronucleosides into Watson–Crick base-pairing arabino-furanosyl nucleosides in water. *Nat. Commun.* **9**, 4073–4082 (2018).
36. Ranjan, S., Todd, Z. R., Rimmer, P. B., Sasselov, D. D. & Babbin, A. R. Nitrogen oxide concentrations in natural waters on early Earth. *Geochim. Geophys. Geosyst.* **20**, 2021–2039 (2019).
37. Xu, J. et al. Photochemical reductive homologation of hydrogen cyanide using sulfite and ferrocyanide. *Chem. Commun.* **54**, 5566–5569 (2018).
38. Marion, G. M., Kargel, J. S., Crowley, J. K. & Catling, D. C. Sulfite–sulfide–sulfate–carbonate equilibria with applications to Mars. *Icarus* **225**, 342–351 (2013).
39. Rios, A. C. & Tor, Y. On the origin of the canonical nucleobases: an assessment of selection pressures across chemical and early biological evolution. *Isr. J. Chem.* **53**, 469–483 (2013).
40. Rios, A. C., Yu, H. T. & Tor, Y. Hydrolytic fitness of N-glycosyl bonds: comparing the deglycosylation kinetics of modified, alternative, and native nucleosides. *J. Phys. Org. Chem.* **28**, 173–180 (2014).
41. Panzica, R. P., Rousseau, R. J., Robins, R. K. & Townsend, L. B. Relative stability and a quantitative approach to the reaction mechanism of the acid-catalyzed hydrolysis of certain 7-and 9-β-D-ribofuranosylpurines. *J. Am. Chem. Soc.* **94**, 4708–4714 (1972).
42. Lindahl, T. & Nyberg, B. Rate of depurination of native deoxyribonucleic acid. *Biochemistry* **11**, 3610–3618 (1972).
43. Hättig, C. Structure optimizations for excited states with correlated second-order methods: CC2 and ADC(2). *Adv. Quantum Chem.* **50**, 37–60 (2005).
44. Dreuw, A. & Wormit, M. The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **5**, 82–95 (2015).
45. Sauer, M. C., Crowell, R. A. & Shkrob, I. A. Electron photodetachment from aqueous anions. I. Quantum yields for generation of hydrated electron by 193 and 248 nm laser photoexcitation of miscellaneous inorganic anions. *J. Phys. Chem. A* **108**, 5490–5502 (2004).
46. Pascoe, D. J., Ling, K. B. & Cockroft, S. L. The origin of chalcogen-bonding interactions. *J. Am. Chem. Soc.* **139**, 15160–15167 (2017).
47. Kim, S. C., O’Flaherty, D. K., Zhou, L., Lelyveld, V. S. & Szostak, J. W. Inosine, but none of the 8-oxo-purines, is a plausible component of a primordial version of RNA. *Proc. Natl. Acad. Sci. USA* **115**, 13318–13323 (2018).
48. Karran, P. & Lindahl, T. Hypoxanthine in deoxyribonucleic acid: generation by heat-induced hydrolysis of adenine residues and release in free form by a deoxyribonucleic acid glycosylase from calf thymus. *Biochemistry* **19**, 6005–6011 (1980).
49. Shapiro, R. & Pohl, S. H. Reaction of ribonucleosides with nitrous acid. Side products and kinetics. *Biochemistry* **7**, 448–455 (1968).
50. Mariani, A. D., Russell, A., Javelle, T. & Sutherland, J. D. A light-releasable potentially prebiotic nucleotide activating agent. *J. Am. Chem. Soc.* **140**, 8657–8661 (2018).

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

© Crown 2020

## Data and materials availability

The Supplementary Information available for this Article contains all procedures, characterization data, NMR spectra, HPLC traces, X-ray data and Cambridge Crystallographic Data Centre (CCDC) numbers, plus theoretical methods and data. Any additional data are available from the corresponding author upon reasonable request.

## Code availability

All custom code used to generate the data in this study is available upon reasonable request.

**Acknowledgements** The authors thank all JDS group members for discussions. This research was supported by the Medical Research Council (MC\_UP\_A024\_1009), the Simons Foundation

(290362 to J.D.S., 494188 to R.S.), and a grant from the National Science Centre Poland (2016/23/B/ST4/01048 to R.W.G.). M.J.J. acknowledges the support of the 'Diamond Grant' (0144/DIA/2017/46) from the Polish Ministry of Science and Higher Education and a computational grant from Wrocław Centre of Networking and Supercomputing (WCSS). R.S. thanks the Foundation for Polish Science for support from the START Fellowship.

**Author contributions** Experimental contributions by J.X., V.C., N.J.G., D.A.R. and A.D.B. Theoretical contributions by M.J.J., R.W.G. and R.S. Crystallography by A.D.B. This work was supervised by J.D.S. All authors co-wrote the manuscript.

**Competing interests** The authors declare no competing interests.

### Additional information

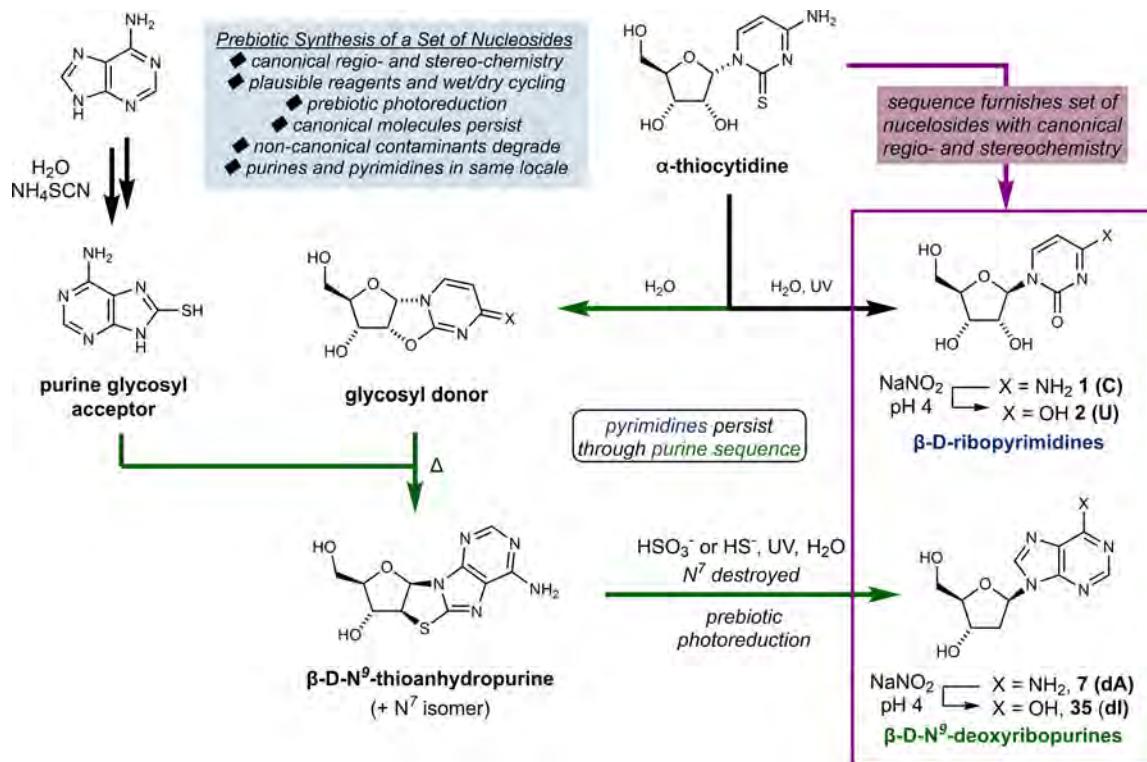
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2330-9>.

**Correspondence and requests for materials** should be addressed to J.S.

**Peer review information** *Nature* thanks Hannes Mutschler and Yitzhak Tor for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

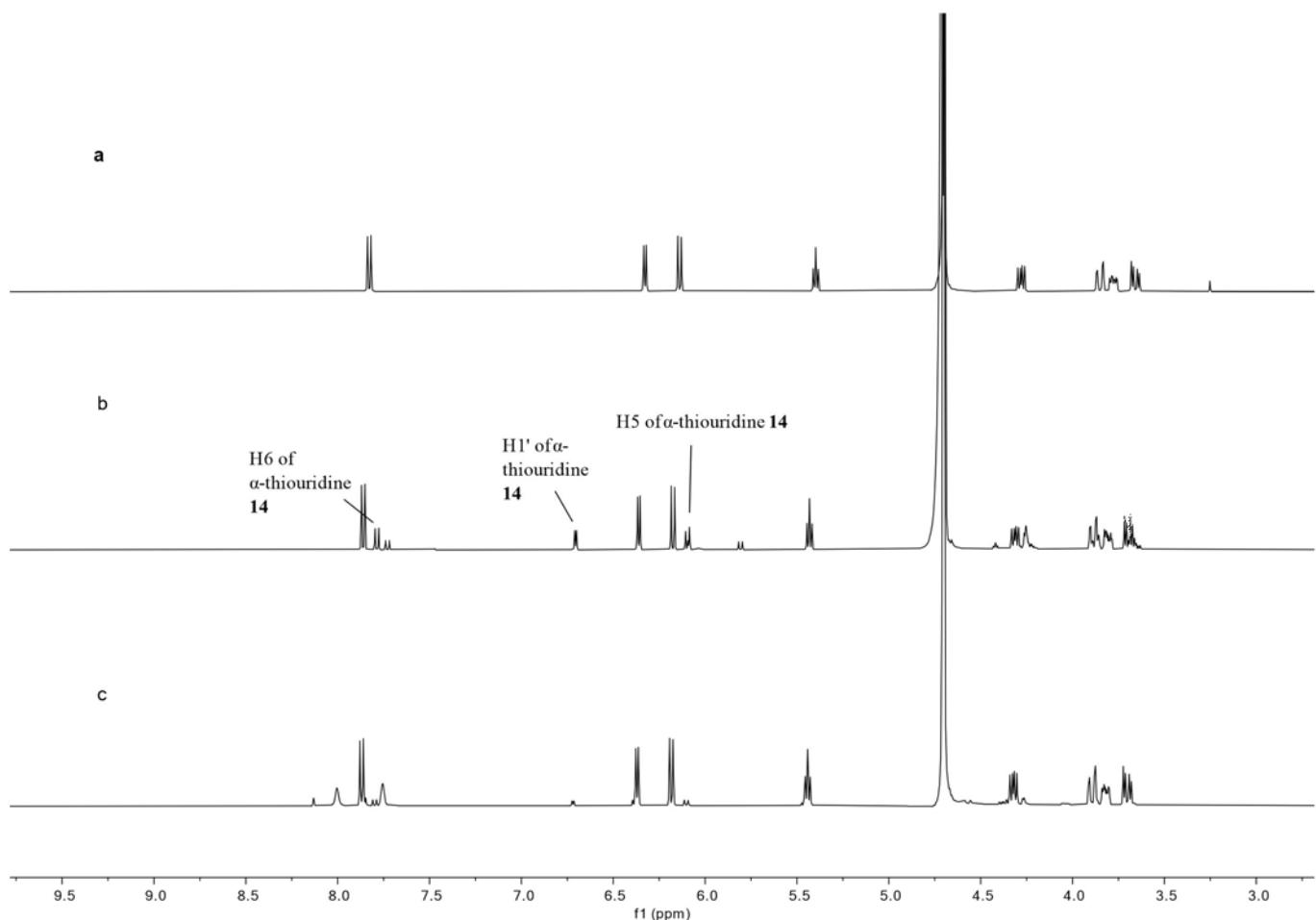
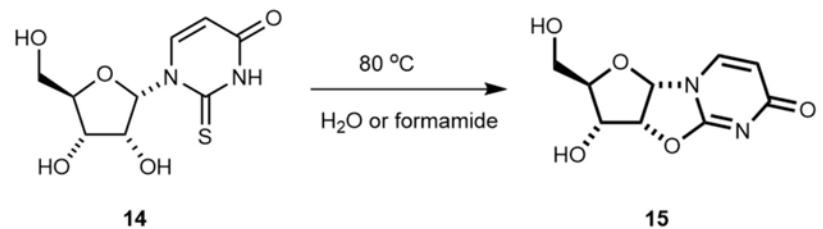
## Article



## **Extended Data Fig. 1 | A summary of the main findings of the work.**

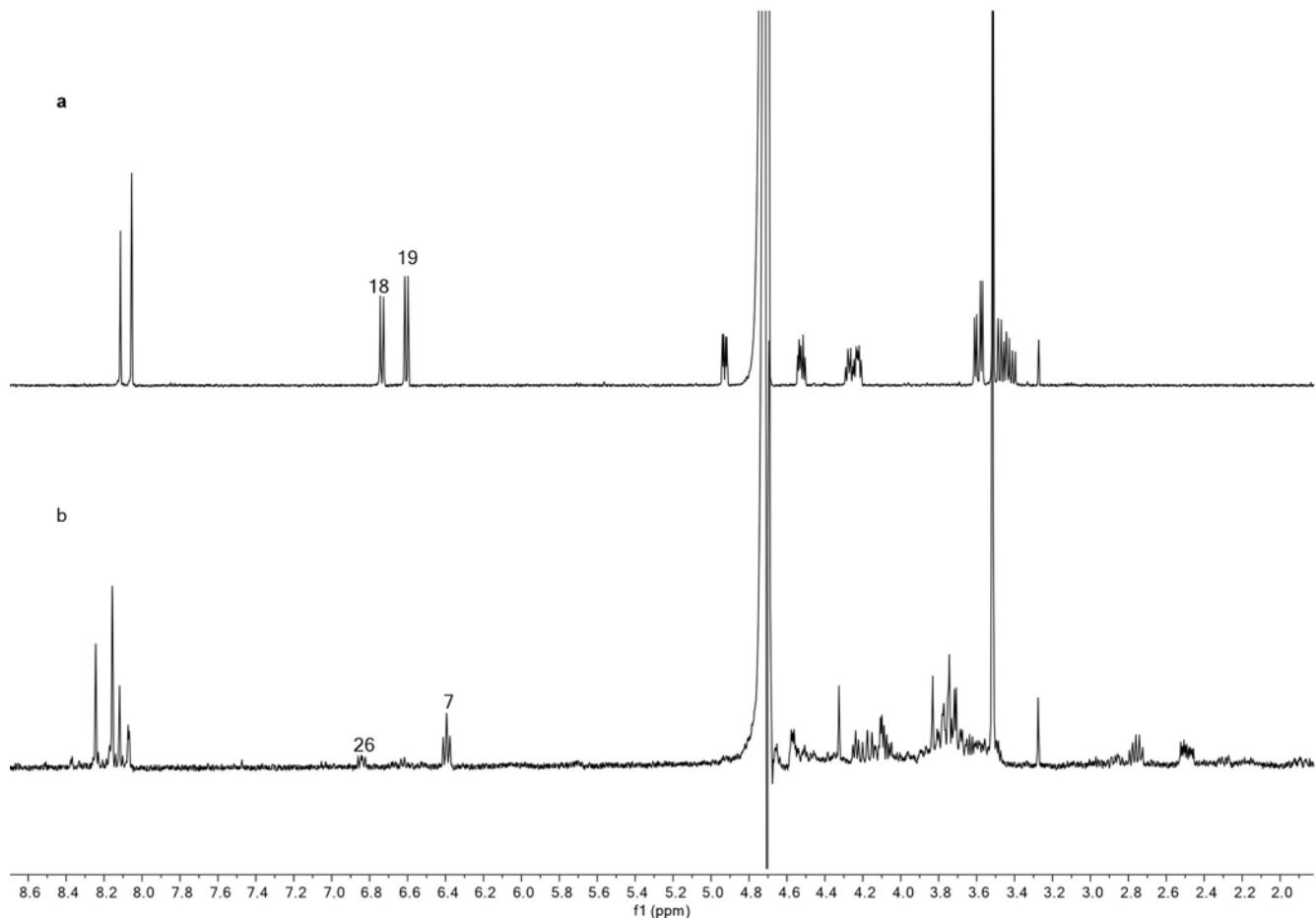
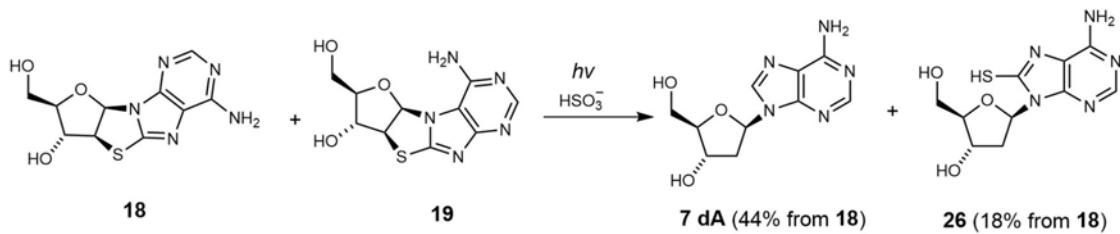
Previously, a prebiotically plausible synthesis of  $\beta$ -ribopyrimidines C and U has been identified using  $\alpha$ -thiocytidine. Herein, we demonstrate that the same intermediate can undergo a distinct prebiotically plausible process that could have happened in a similar—or the same—environment. This process furnishes  $\beta$ -D- $N^9$ -deoxyribopurine nucleosides dA and dI alongside the pyrimidines.

Remarkable selectivity enforced by UV irradiation and hydrolysis operates throughout the reported ribosylpyrimidine synthesis and the discovered deoxyribosylpurine synthesis, resulting in a set of nucleosides with only the canonical regio- and stereochemistry. The coexistence in one location of a set of nucleosides similar to this is thought to be a precondition for the spontaneous emergence of life on Earth<sup>6,47</sup>.



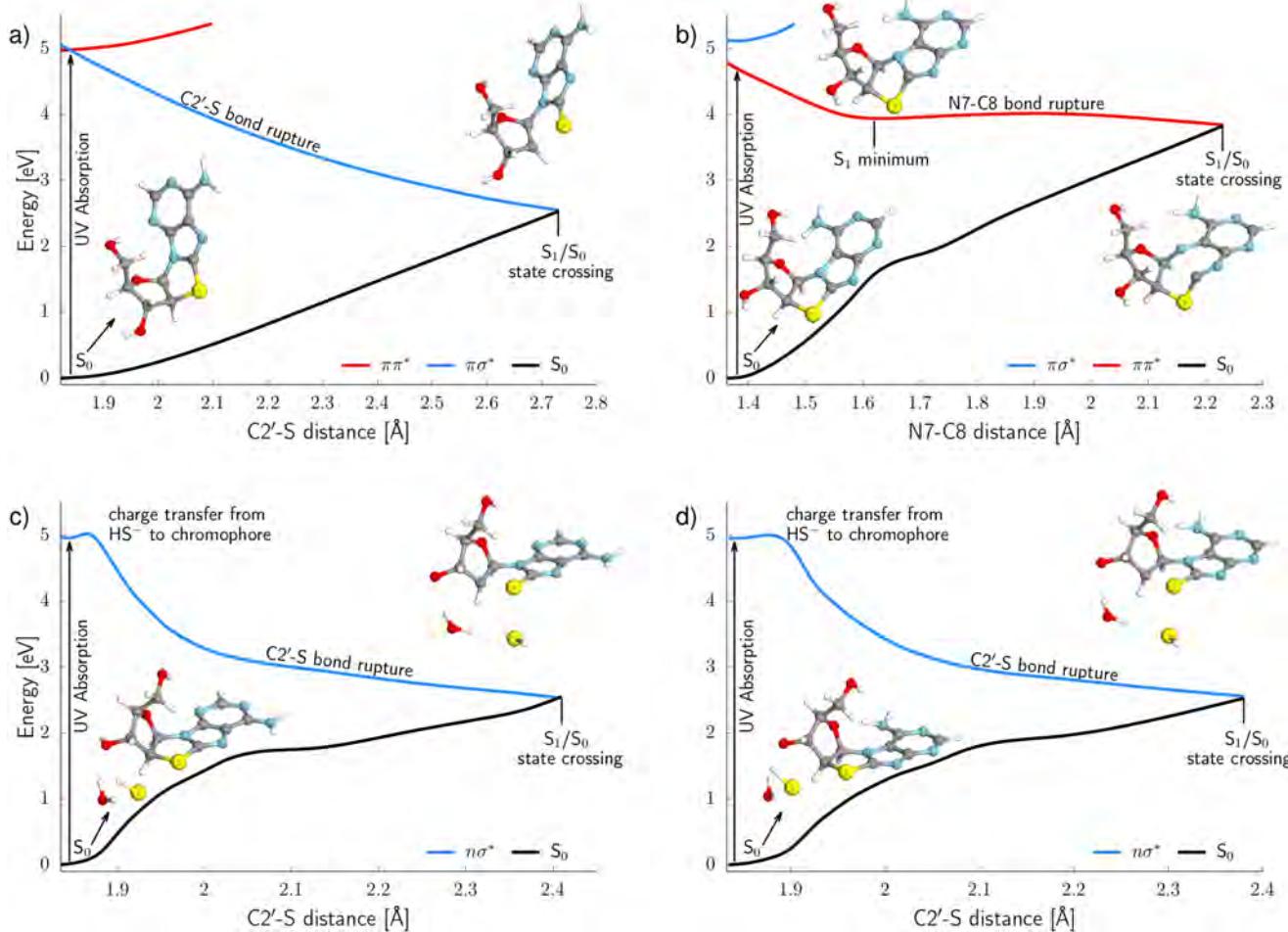
**Extended Data Fig. 2 |  $^1\text{H}$  NMR spectra of conversion of  $\alpha$ -anhydrouridine (**15**) from  $\alpha$ -thiouridine (**14**). a.**  $^1\text{H}$  NMR spectrum of **15**. **b.**  $^1\text{H}$  NMR spectrum of the reaction mixture after heating **14** in  $\text{H}_2\text{O}$ . **c.**  $^1\text{H}$  NMR spectrum of the reaction mixture after heating **14** in formamide. f1, chemical shift ( $\delta$ ).

## Article



**Extended Data Fig. 3 |  $^1\text{H}$  NMR spectra of photoreduction of  $N^7$ -8,2'-anhydro-thioadenosine (**18**) and  $N^9$ -8,2'-anhydro-thioadenosine (**19**) mixture with bisulfite. a,**  $^1\text{H}$  NMR spectrum of the crude mixture before

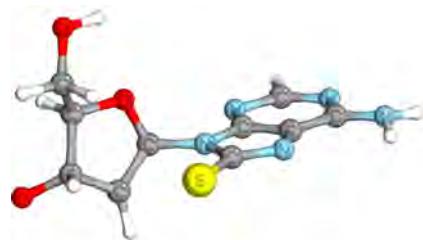
irradiation; the ratio of  $N^7:N^9$  isomer was 4:5. **b,**  $^1\text{H}$  NMR spectrum of the mixture after irradiation for 7 h; the  $N^9$  isomers dA (**7**) and **26** are the only detectable products. f1, chemical shift ( $\delta$ ).



**Extended Data Fig. 4 | Potential energy surfaces and  $S_1/S_0$  state crossings of the key photochemical steps in deoxyadenosine synthesis calculated using ADC(2) and the ma-def2-TZVP basis set.** See Supplementary Information for details. **a**, Potential energy profile of UV-induced C-S bond scission of **18**. C-S bond opening may spontaneously occur in **18**, leading to a peaked  $S_1/S_0$  state crossing; however, a reducing agent is necessary to maintain that geometry after reaching the  $S_0$  state. **b**, Potential energy profile of

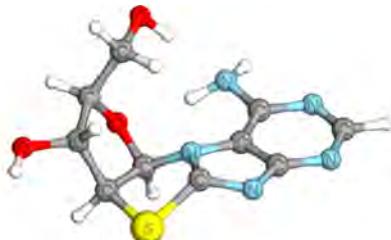
UV-induced N7-C8 bond scission of **19**. N7-C8 bond rupture is the lowest-energy photochemical process in **19** and results in destruction of the purine ring. **c, d**, Potential energy profiles of the UV-induced C-S bond scission of encounter complexes **18** (**c**) and **19** (**d**) with HS<sup>-</sup>. Photochemical C-S bond rupture induced by charge transfer from HS<sup>-</sup> to a chromophore and is a barrierless process.

## Article



**31 radical anion**

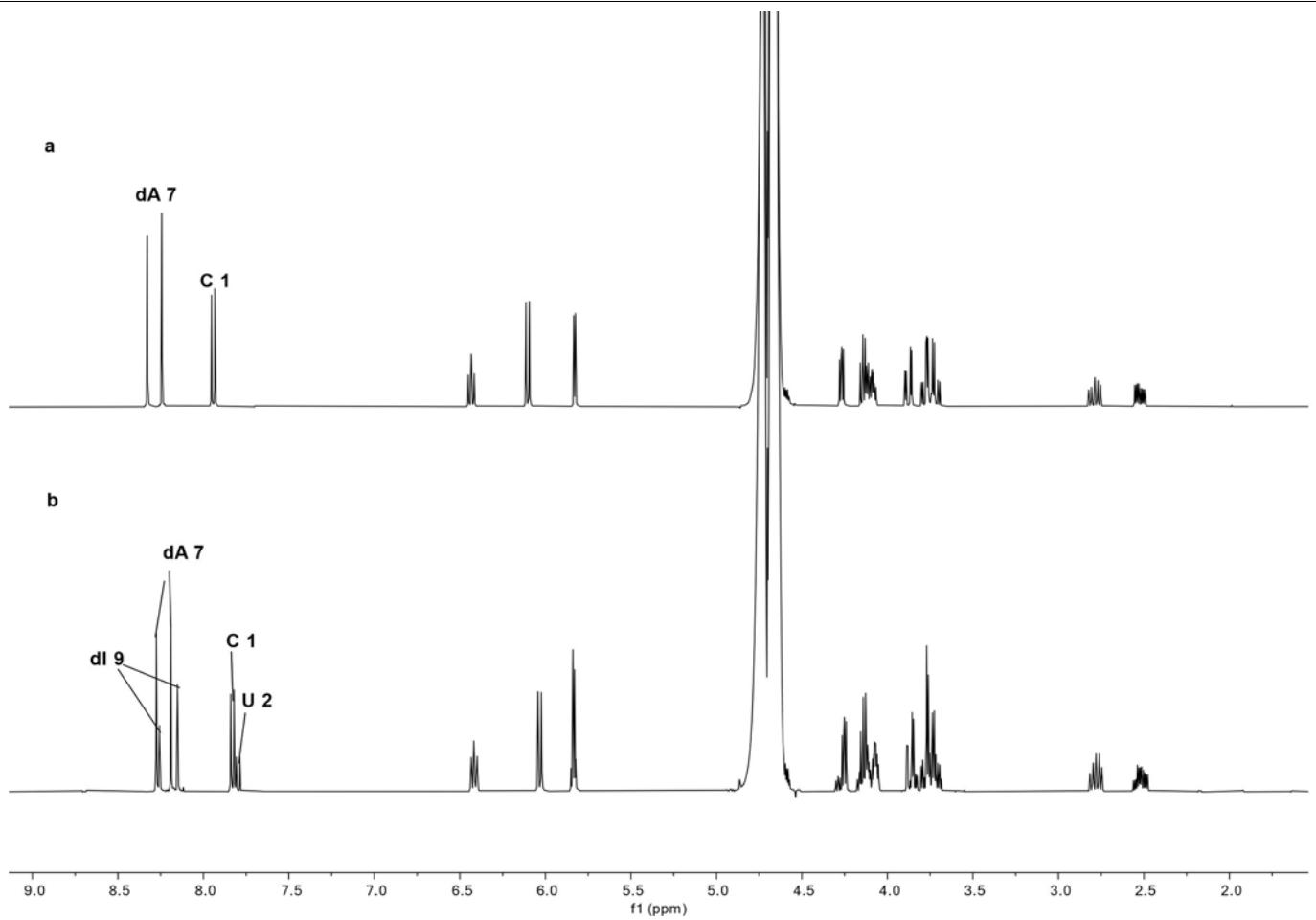
Adiabatic electron affinity: 2.50 eV



**32 radical anion**

Adiabatic electron affinity: 1.83 eV

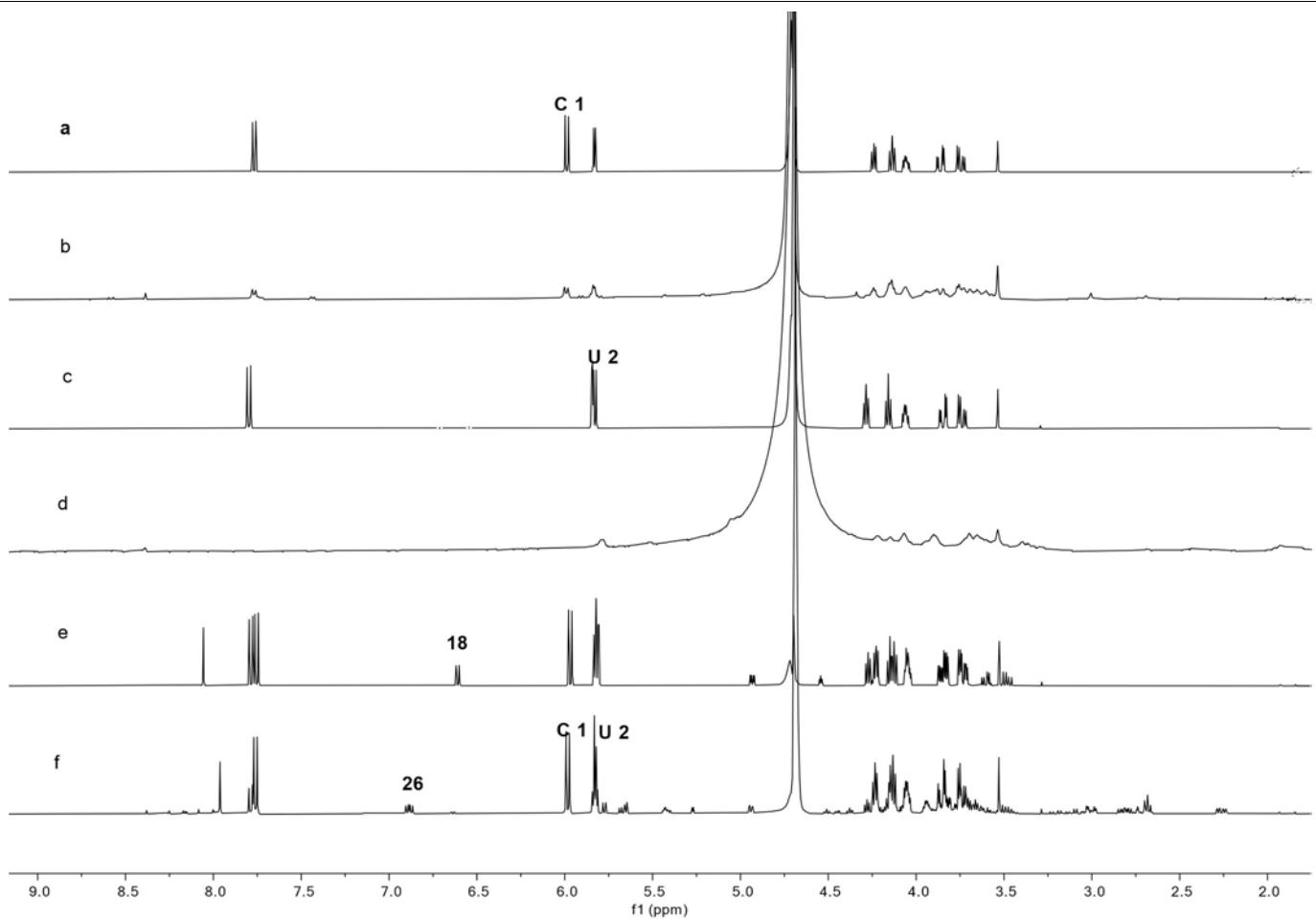
**Extended Data Fig. 5 | Equilibrium geometries of C2, S8 radical anion (31) and C8, N9 radical anion (32).** Radical anions may be formed after accepting a hydrated electron from the environment. The adiabatic electron affinities are calculated using ωB97X-D/IEFPCM and the ma-def2-TZVP basis set.



**Extended Data Fig. 6 | <sup>1</sup>H NMR spectra for the reactions of deoxyadenosine (dA, **7**) and cytidine (C, **1**) with nitrous acid. **a**,** <sup>1</sup>H NMR spectrum of the mixture of dA (**7**) and C (**1**). **b**, <sup>1</sup>H NMR spectrum of the reaction mixture after

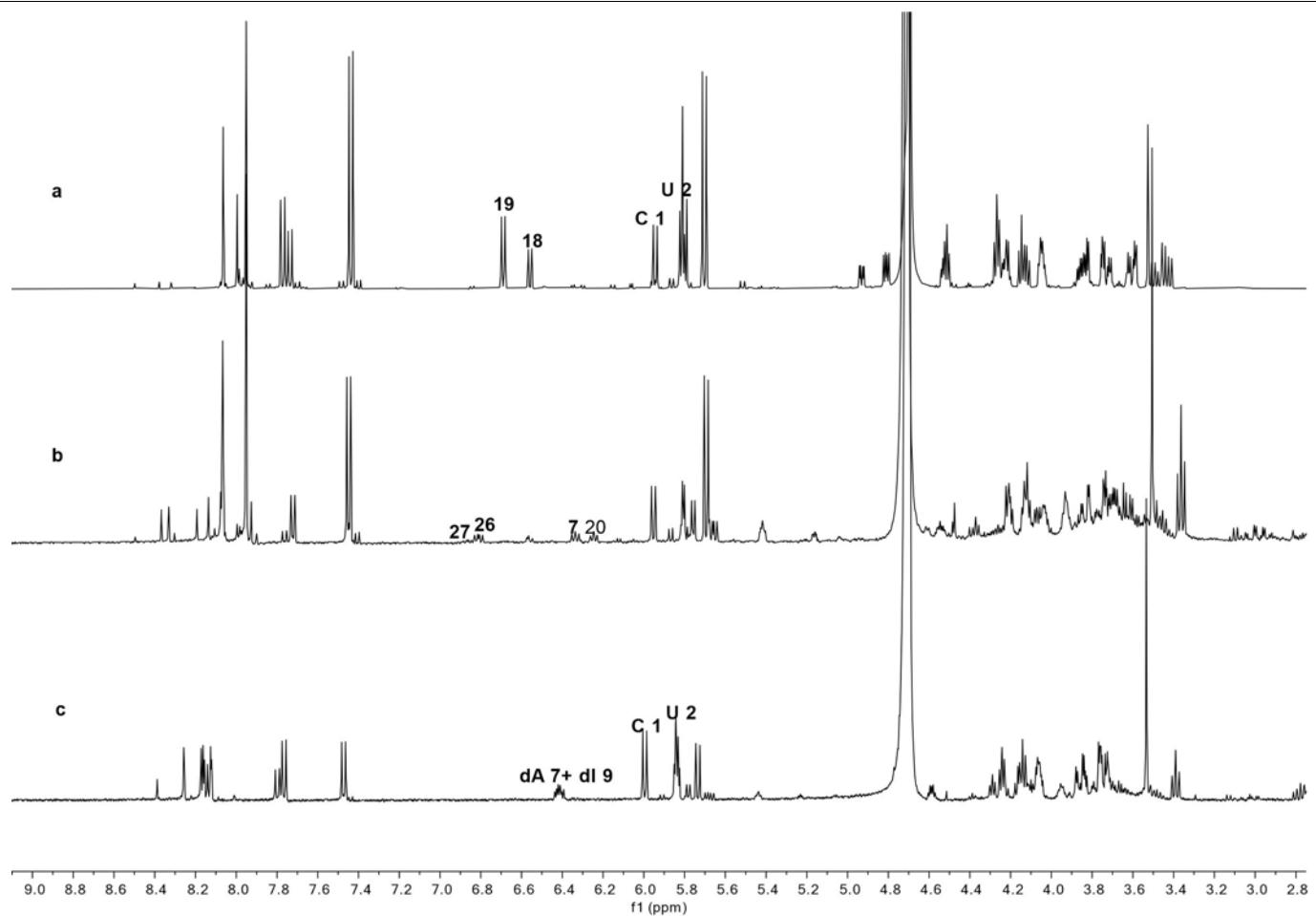
4 d, showing that the ratio of the four (deoxy)nucleosides dA (**7**), deoxyinosine (dl, **9**), C (**1**), and uridine (U, **2**) is 30:17:42:11. f1, chemical shift ( $\delta$ ).

## Article



**Extended Data Fig. 7 | <sup>1</sup>H NMR spectra for stability study of cytidine (C; 1) and uridine (U; 2) at 254 nm irradiation with bisulfite.** **a**, <sup>1</sup>H NMR spectrum of the mixture of C (1), bisulfite and K<sub>4</sub>Fe(CN)<sub>6</sub> in the dark. **b**, As in **a**, after 10 h of irradiation. **c**, <sup>1</sup>H NMR spectrum of the mixture of U (2), bisulfite and K<sub>4</sub>Fe(CN)<sub>6</sub>

in the dark. **d**, As in **c**, after 10 h of irradiation. **e**, <sup>1</sup>H NMR spectrum of the mixture of C (1), U (2), *N*<sup>9</sup>-thioanhydroadenosine (18), bisulfite and K<sub>4</sub>Fe(CN)<sub>6</sub> in the dark. **f**, As in **e**, after 10 h of irradiation. f1, chemical shift ( $\delta$ ).



**Extended Data Fig. 8 |  $^1\text{H}$  NMR spectra for sequential reactions with the mixture of  $\alpha$ -anhydrouridine (**15**), **C** (**1**) and **U** (**2**).** **a.**  $^1\text{H}$  NMR spectrum of the mixture after heating with 8-mercaptopadenine (**16**) and magnesium chloride at

150 °C for 1.5 d. **b.** As in **a**, after irradiation with hydrogen sulfide at 254 nm. **c.** As in **a**, after reacting with nitrous acid for 2 d; dA (**7**):dl (**9**):C (**1**):U (**2**) = 14:14:44:28. f1, chemical shift ( $\delta$ ).

## Article

# Displaced cratonic mantle concentrates deep carbon during continental rifting

<https://doi.org/10.1038/s41586-020-2328-3>

Received: 19 August 2019

Accepted: 20 March 2020

Published online: 3 June 2020

 Check for updates

James D. Muirhead<sup>1,13</sup>, Tobias P. Fischer<sup>2</sup>, Sarah J. Oliva<sup>3</sup>, Amani Laizer<sup>4</sup>, Jolante van Wijk<sup>5</sup>, Claire A. Currie<sup>6</sup>, Hyunwoo Lee<sup>7</sup>, Emily J. Judd<sup>1</sup>, Emmanuel Kazimoto<sup>4</sup>, Yuji Sano<sup>8</sup>, Naoto Takahata<sup>8</sup>, Christel Tiberi<sup>9</sup>, Stephen F. Foley<sup>10</sup>, Josef Dufek<sup>11</sup>, Miriam C. Reiss<sup>12</sup> & Cynthia J. Ebinger<sup>3</sup>

Continental rifts are important sources of mantle carbon dioxide ( $\text{CO}_2$ ) emission into Earth's atmosphere<sup>1–3</sup>. Because deep carbon is stored for long periods in the lithospheric mantle<sup>4–6</sup>, rift  $\text{CO}_2$  flux depends on lithospheric processes that control melt and volatile transport<sup>1,3,7</sup>. The influence of compositional and thickness differences between Archaean and Proterozoic lithosphere on deep-carbon fluxes remains untested. Here we propose that displacement of carbon-enriched Tanzanian cratonic mantle concentrates deep carbon below parts of the East African Rift System. Sources and fluxes of  $\text{CO}_2$  and helium are examined over a 350-kilometre-long transect crossing the boundary between orogenic (Natron and Magadi basins) and cratonic (Balangida and Manyara basins) lithosphere from north to south. Areas of diffuse  $\text{CO}_2$  degassing exhibit increasing mantle  $\text{CO}_2$  flux and  $^3\text{He}/^4\text{He}$  ratios as the rift transitions from Archaean (cratonic) to Proterozoic (orogenic) lithosphere. Active carbonatite magmatism also occurs near the craton edge. These data indicate that advection of the root of thick Archaean lithosphere laterally to the base of the much thinner adjacent Proterozoic lithosphere creates a zone of highly concentrated deep carbon. This mode of deep-carbon extraction may increase  $\text{CO}_2$  fluxes in some continental rifts, helping to control the production and location of carbonate-rich magmas.

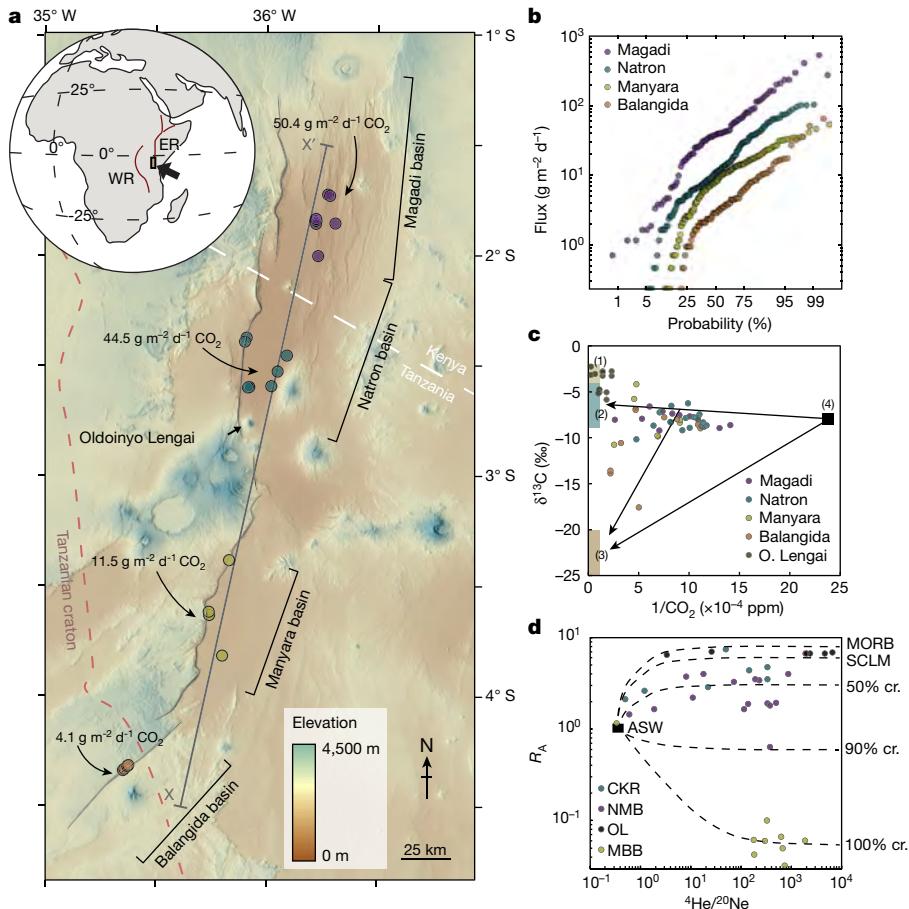
Abundant carbon is sequestered in mantle lithosphere during the infiltration of plume and carbon-rich silicate melts that are generated during mantle convection and periods of subduction<sup>5,6,8</sup>, enriching mantle carbon contents by up to 100 times their original values<sup>6</sup>. Deep carbon accumulated by these processes is released during continental rifting<sup>8,9</sup>. Rift-related  $\text{CO}_2$  degassing thus has the potential to modulate Earth's climate on geological timescales<sup>1,2</sup>; however, the total volume of mantle  $\text{CO}_2$  emitted at rift settings is poorly constrained, as are the deep lithospheric processes that control variations in mantle  $\text{CO}_2$  flux.

The Eastern Rift of the East African Rift System (EARS; Fig. 1) is an ideal location to investigate the factors controlling rift  $\text{CO}_2$  degassing, through measurements of active mantle  $\text{CO}_2$  release<sup>1,10–12</sup>. Given the large aerial extent, pervasive faulting and widespread magmatism in the EARS, quantifying the volume of  $\text{CO}_2$  release requires observations from a wide variety of locations along the rift system. Results of diffuse soil degassing surveys have thus far been reported from the northern and central Main Ethiopian Rift<sup>7</sup> and the Magadi and Natron basins<sup>1</sup>, with estimated  $\text{CO}_2$  fluxes of  $0.52\text{--}4.36 \text{ Mt yr}^{-1}$  and  $2.15\text{--}5.95 \text{ Mt yr}^{-1}$ , respectively. Extrapolation of these estimates to the entire Eastern Rift

indicates potential  $\text{CO}_2$  fluxes of the order of  $10\text{--}100 \text{ Mt yr}^{-1}$ . Current estimates of  $\text{CO}_2$  degassing, however, do not consider spatial variations in mantle  $\text{CO}_2$  release associated with diachronous rifting and variations in lithospheric thickness and composition.

The flux of  $\text{CO}_2$  within any rift basin depends on a number of variables, such as the concentration of stored carbon in the lithosphere, the rates of melt production, the volume of the lithospheric mantle involved in melting and the available transport pathways<sup>6,7</sup>. Geophysical data and numerical modelling show that lithospheric thinning, and the related magmatism, can occur in zones up to 10 times wider than the surface expression of a rift<sup>13,14</sup>. Mantle convection and plume deflection near the base of the lithosphere will result in extensive lateral transport of hot mantle material that preferentially ponds and melts in zones of thinner lithosphere<sup>15</sup>. Lithosphere composition and structure also play a critical role in magma petrogenesis<sup>16</sup>. In the Eastern Rift, K/Na ratios increase from off- to on-craton localities, and carbonatite magmatism (Fig. 2) is spatially associated with potassium-enriched melts near the Tanzanian craton edge<sup>16–18</sup>. Despite clear variations in the style and nature of magmatism in cratonic versus orogenic lithosphere, no

<sup>1</sup>Department of Earth Sciences, Syracuse University, Syracuse, NY, USA. <sup>2</sup>Department of Earth and Planetary Sciences, University of New Mexico, Albuquerque, NM, USA. <sup>3</sup>Department of Earth and Environmental Sciences, Tulane University, New Orleans, LA, USA. <sup>4</sup>Department of Geology, University of Dar es Salaam, Dar es Salaam, Tanzania. <sup>5</sup>New Mexico Institute of Mining and Technology, Socorro, NM, USA. <sup>6</sup>Department of Physics, University of Alberta, Edmonton, Alberta, Canada. <sup>7</sup>School of Earth and Environmental Sciences, Seoul National University, Seoul, South Korea. <sup>8</sup>Atmosphere and Ocean Research Institute, University of Tokyo, Chiba, Japan. <sup>9</sup>Géosciences Montpellier, CNRS, Université de Montpellier, Montpellier, France. <sup>10</sup>Department of Earth and Environmental Sciences, Macquarie University, North Ryde, New South Wales, Australia. <sup>11</sup>Department of Earth Sciences, University of Oregon, Eugene, OR, USA. <sup>12</sup>Institute of Geosciences, Goethe University Frankfurt, Frankfurt am Main, Germany. <sup>13</sup>Present address: School of Environment, University of Auckland, Auckland, New Zealand.  
✉e-mail: james.muirhead@fulbrightmail.org; fischer@unm.edu



**Fig. 1 | CO<sub>2</sub> flux and carbon and helium isotope data in the study region.**

**a.** Annotated Shuttle Radar Topography Mission (SRTM) digital elevation model of the study region. Filled circles show the locations of the sampling regions, and CO<sub>2</sub> flux values represent mean values for high-flux populations in each basin (the helium isotope data in **d** support predominantly crustal volatile contributions in the Balangida and Manyara basins and mantle contribution in the Natron and Magadi basins). Border faults defining each basin are transparent grey. The dashed brown line delineates the surface expression of the eastward-dipping boundary for the Tanzanian craton. The cross-section line X–X' relates to the cross-sections presented in Fig. 3. The top left location inset also shows the Eastern Rift (ER) and Western Rift (WR) of the EARS.

**b.** Probability plots of diffuse CO<sub>2</sub> flux in each basin ( $n=547$ ). **c.** Carbon isotope content ( $\delta^{13}\text{C}-\text{CO}_2$ ) versus the reciprocal of the CO<sub>2</sub> concentration<sup>48</sup> for diffuse CO<sub>2</sub> samples ( $n=62$ ). Shown are end-member carbon isotope values<sup>1</sup> for fumaroles (1), mantle (2), biogenic CO<sub>2</sub> (3) and air (4). **d.**  $^3\text{He}/^4\text{He}$  data showing  $R_A$  versus  $^4\text{He}/^{20}\text{Ne}$  for springs and fumaroles in the Eastern Rift of the EARS ( $n=38$ ), indicating mixing lines between air-saturated water (ASW) and endmembers (depleted mid-ocean ridge basalts, MORB, and sub-continental lithospheric mantle, SLCM) with varying proportions of crustal (cr.) helium. OL, Oldoinyo Lengai; CKR, Central Kenya Rift; NMB, Natron and Magadi basins; MBB, Manyara and Balangida basins. Analytical errors are discussed in Methods and do not exceed the size of the presented data points.

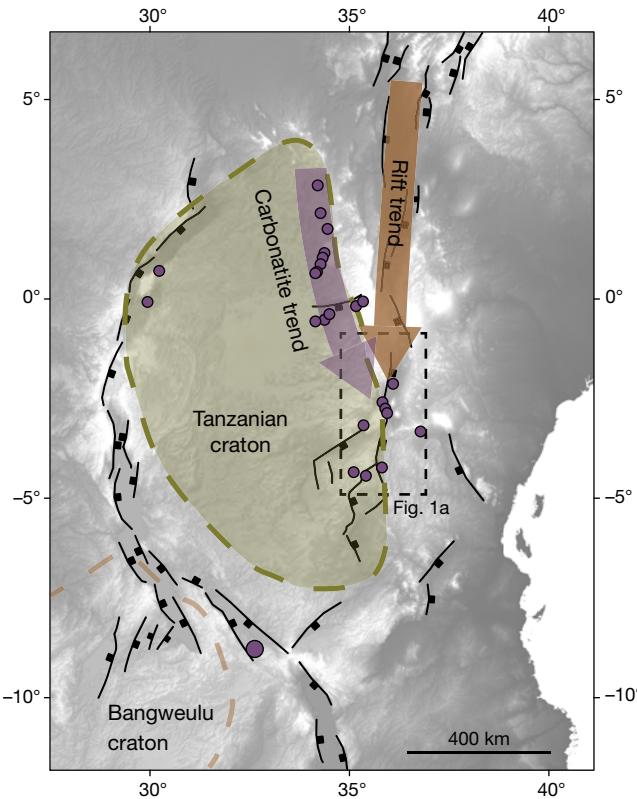
study has so far examined how deep-carbon fluxes in the Eastern Rift vary across cratonic–orogenic boundaries, where gradients in the depth of the lithosphere–asthenosphere boundary can drive lateral advection of buoyant cratonic mantle into zones of lithospheric thinning<sup>19</sup>. We address this critical question by integrating recent and existing gas geochemical data with geophysical imaging of the EARS over an approximately 350-km-long transect extending southwards from orogenic to cratonic lithosphere<sup>20,21</sup> (Fig. 1). By considering the results in the context of recent numerical models of the stability of cratonic lithosphere<sup>19</sup>, we propose a model of the lithosphere displacement of Tanzanian craton mantle that explains distinct variations observed in mantle CO<sub>2</sub> flux.

### Rift degassing across a craton boundary

The Magadi, Natron, Manyara and Balangida basins (south Kenya and north Tanzania) form a series of kinematically linked rift segments representing an early expression (<7 million years ago; Ma) of continental rifting<sup>22</sup> (Fig. 1a). These basins occur near the boundary between two major lithospheric terranes, the Archaean Tanzanian craton and Proterozoic Mozambique Belt of the Pan-African orogeny. The surface

geological expression of the Tanzanian craton is observed only in the westernmost Balangida basin (Fig. 1a). However, xenolith data<sup>23–25</sup> combined with geophysical constraints<sup>21</sup> suggest that the Balangida and Manyara basins, and potentially parts of the southern Natron basin, lie predominantly above the Archaean Tanzanian craton<sup>16</sup> rather than the Proterozoic lithosphere, which is present under the Natron and Magadi basins farther north. This interpretation is supported by the velocity and density structure of the lithosphere derived from a joint analysis of gravity and teleseismic data<sup>21</sup>, which reveals a distinct geophysical boundary with relatively low density and seismic velocities at depths of approximately 30–200 km below the Natron and Magadi basins (Fig. 3; see also Extended Data Fig. 1).

By investigating spring systems and areas of diffuse degassing situated along deeply penetrating faults of the Magadi, Natron, Manyara and Balangida basins, we examine how the nature of deep volatile transport varies across this craton boundary. Geophysical and gas geochemical data reveal the Natron and Magadi basins as key examples of active mantle CO<sub>2</sub> degassing<sup>1,20,26</sup>, whereas mantle volatile discharges in the Manyara and Balangida basins are yet to be examined. Given the inherent variability of diffuse CO<sub>2</sub> flux within individual rift basins<sup>7</sup>, it is important to compare data from sites with similar structure, substrates

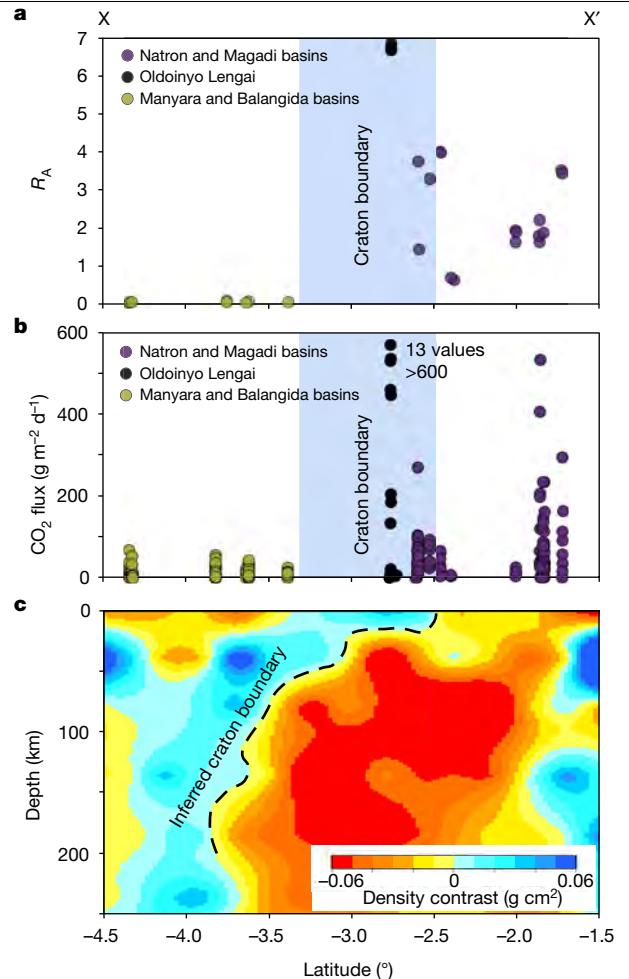


**Fig. 2 | Distribution of <45-Myr-old carbonatite systems associated with the EARS and the Archaean Tanzanian craton.** Carbonatites (purple circles) preferentially follow (purple arrow) the boundary of the Tanzanian craton (dashed line; craton boundary from Foley et al.<sup>16</sup>) rather than the dominant rift trend (brown arrow). The simplified distribution of major faults is presented as black solid lines. All features are annotated on the 90-m SRTM digital elevation model. The presented carbonatite systems are summarized in Supplementary Table 4.

and hydrology when comparing CO<sub>2</sub> discharges between basins. Therefore, we focus specifically on diffuse CO<sub>2</sub> flux data (1) from rift-graben sediments, (2) in the vicinity of faults and (3) in areas within 100 m of observed springs. Additionally, we present helium isotope data from springs, as well as diffuse gas flux and carbon and helium isotope data from Oldoinyo Lengai.

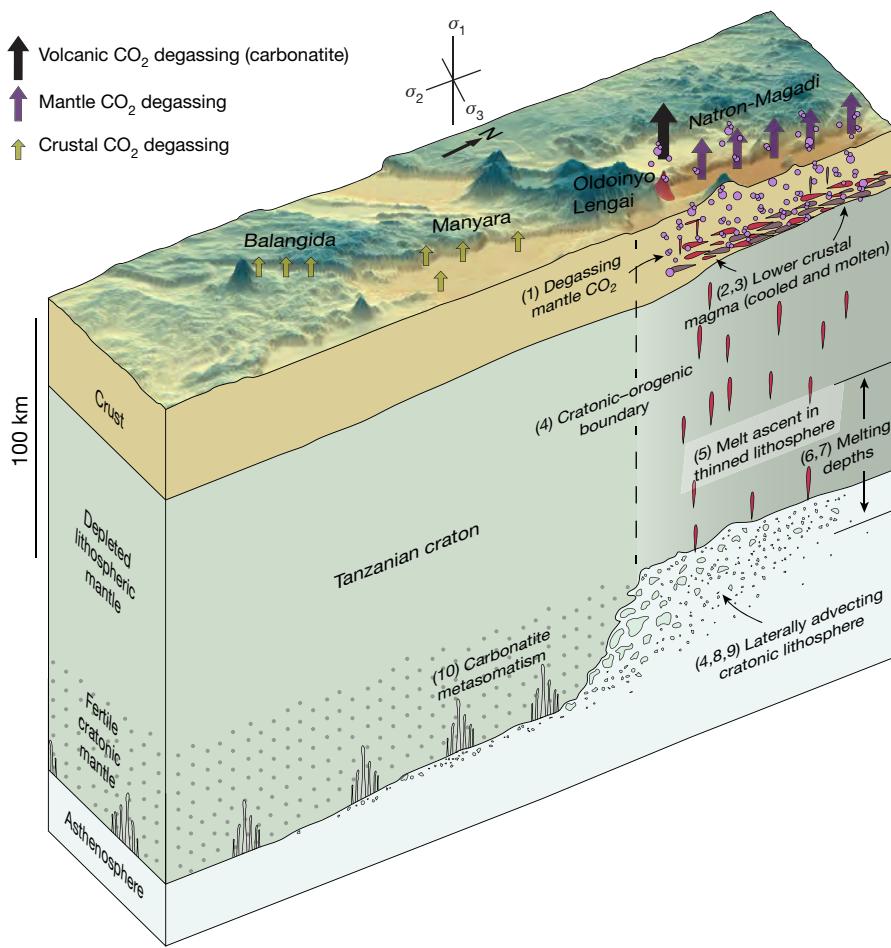
Volatile data exhibit a systematic north-to-south transition in the study region. Helium isotope data (<sup>3</sup>He/<sup>4</sup>He ratios) exhibit values ranging from 0.03R<sub>A</sub> to 4.0R<sub>A</sub> (R<sub>A</sub> = <sup>3</sup>He/<sup>4</sup>He ratio of air), with the highest values in the Natron and Magadi basins (0.7R<sub>A</sub>–4.0R<sub>A</sub>, average of (2.3±1.1)R<sub>A</sub>) compared to the Manyara ((0.05±0.02)R<sub>A</sub>) and Balangida ((0.05±0.01)R<sub>A</sub>) basins (Supplementary Table 3). These helium isotope ratios indicate that dissolved spring volatiles in the Manyara and Balangida basins have a crustal source (typically around 0.05R<sub>A</sub>)<sup>27</sup>. By contrast, <sup>3</sup>He/<sup>4</sup>He ratios measured farther north in the Natron and Magadi basins reach up to 4.0R<sub>A</sub> (ref. <sup>27</sup>), approaching values almost as high as those inferred for the subcontinental lithospheric mantle (6.1R<sub>A</sub>)<sup>28</sup>. A subcontinental lithospheric mantle source for magma and volatiles in the Natron and Magadi basins is further supported by rare-earth and helium isotope data of lavas and xenoliths<sup>29,30</sup>; by contrast, geochemical data<sup>30–33</sup> suggest that plume melts are more prevalent in regions much farther north (for example, Ethiopia). Farther to the south, that is, in the Rungwe Volcanic Province, some helium isotope data of xenoliths suggest plume melt sources<sup>34</sup>, whereas radiogenic isotopes display characteristics inconsistent with primitive mafic lavas of the Afar Depression, Ethiopia<sup>30</sup>.

Carbon isotope data of diffusely degassing CO<sub>2</sub> ( $\delta^{13}\text{C}$ -CO<sub>2</sub>; see Methods) indicate predominantly mantle carbon sources in the Natron and



**Fig. 3 | Latitudinal variations in CO<sub>2</sub> flux and R<sub>A</sub> values for <sup>3</sup>He/<sup>4</sup>He with respect to the modelled lithosphere structure.** **a**, R<sub>A</sub> versus latitude along the Eastern Rift from -1.5° to -4.5° (n = 25). Air-contaminated samples (TZ18-C16) were removed from the analysis. The Tanzanian craton boundary (light grey) is interpreted on the basis of sharp density (c) and velocity contrasts near the northern Manyara and southern Natron basins (see also Extended Data Fig. 1). It is marked as a broad region to account for location uncertainties and the overall three-dimensional nature of the boundary in the region (see also Methods). **b**, Diffuse CO<sub>2</sub> flux versus latitude along the Eastern Rift from -1.5° to -4.5° for the sample sites presented in Fig. 1a and Oldoinyo Lengai volcano (n = 610). **c**, Lithosphere density model from Tiberi et al.<sup>21</sup>. Contrasts are relative to the IASP91 model<sup>49</sup>. The position of the cross-section is shown in Fig. 1a.

Magadi basins<sup>1,27</sup>, with progressively more mixing between biogenic and perhaps crustal CO<sub>2</sub> southwards into the Manyara and Balangida basins (Fig. 1). Additionally, we observe a north-to-south decrease in the flux of diffusely degassing CO<sub>2</sub> (Figs. 1, 3), where mean values for the high-CO<sub>2</sub>-flux populations in the Natron and Magadi basins drop from 44.5–50.4 g m<sup>-2</sup> d<sup>-1</sup> to 4.1–11.5 g m<sup>-2</sup> d<sup>-1</sup> in the Balangida and Manyara basins (Extended Data Fig. 2). This comparison also does not consider substantial degassing from the Oldoinyo Lengai volcano in the Natron basin<sup>10</sup>, which exhibits the highest diffuse CO<sub>2</sub> flux measurements in the region (up to 7,376 g m<sup>-2</sup> d<sup>-1</sup>), suggesting that the Natron basin is probably the major carbon emitter in this sector of the EARS (Fig. 3). The Balangida and Manyara basins exhibit a considerably lower overall CO<sub>2</sub> flux, and <sup>3</sup>He/<sup>4</sup>He ratios (0.03R<sub>A</sub> to 0.07R<sub>A</sub>) support the release of predominantly crustal volatiles. These two observations suggest that the lower diffuse CO<sub>2</sub> fluxes here probably represent a mixed biogenic and minor crustal carbon source, such as that released by dissolution of limestone<sup>27</sup>. In summary, data collected in the Manyara and Balangida



**Fig. 4 | Proposed model for deep-carbon transport along a cratonic boundary in the EARS study region.** Cratonic mantle lithosphere is advected laterally and vertically into a region of melting near the cratonic boundary. CO<sub>2</sub>-rich melts rise through the thinning mobile belt, intrude the lower crust and degass, or erupt at the surface as highly alkaline or carbonatite lavas. The

tectonic stress state is represented as the highest ( $\sigma_1$ ), intermediate ( $\sigma_2$ ) and lowest ( $\sigma_3$ ) compressive stresses. Data from (1) Lee et al.<sup>1</sup>; (2) Roecker et al.<sup>20</sup>; (3) Ebinger et al.<sup>22</sup>; (4) Tiberi et al.<sup>21</sup>; (5) Selway<sup>35</sup>; (6) Mana et al.<sup>29</sup>; (7) Dawson et al.<sup>50</sup>; (8) Currie & van Wijk<sup>19</sup>; (9) Huismans and Beaumont<sup>39</sup>; (10) Rudnick et al.<sup>38</sup>.

basins clearly show that there is negligible mantle volatile degassing locally through the Archaean Tanzania craton. Furthermore, helium data indicate predominantly crustal volatile discharges and, combined with the low CO<sub>2</sub> fluxes, suggest that the craton here is either not a substantial source of mantle carbon and/or represents an impermeable barrier to the ascent of mantle carbon<sup>35</sup>. Given the probable accumulation of carbon in the lower cratonic lithosphere over time<sup>6,9</sup>, the latter explanation is more likely.

### Lithospheric transport of cratonic carbon

Continental rifts are rife with dilational fault systems that act as permeable conduits for the ascent of deep mantle volatiles released from cooling magmas<sup>1,3</sup>, which are extracted from volatile-rich, sub-continental lithospheric mantle<sup>6</sup>. Reconstructions of continental rift lengths over the last 200 million years suggest a potential positive correlation between rift length and partial pressure of atmospheric CO<sub>2</sub> (ref. <sup>2</sup>). However, estimates of CO<sub>2</sub> fluxes from continental rift settings in Earth's geological past are informed by present-day measurements, with the EARS being a typical example of voluminous rift degassing<sup>1,7</sup>. Calculated global fluxes of CO<sub>2</sub> from continental rifts<sup>2</sup> thus consider mean fluxes from relatively sparse measurements<sup>1,7</sup>, neglecting process-driven spatial variations in mantle carbon release both in the present and in the geological past. Indeed, our results indicate dramatic spatial variations in mantle CO<sub>2</sub> flux between adjacent rift

basins in the Eastern Rift and associations with carbonate-rich melt generation near craton–orogen boundaries (Figs. 1–3). It is therefore likely that there are deep processes at play that act to concentrate mantle carbon below specific rift sectors, driving these observed CO<sub>2</sub> flux variations.

Which processes promote higher fluxes and concentrations of mantle CO<sub>2</sub> in a continental rift? One potential answer could be that the cratonic mantle lithosphere is carbon-poor compared to neighbouring lithospheric terranes. However, given the age of the Archaean terrane, it is expected that there are deep-carbon stores in the lowermost lithospheric mantle from infiltration related to previous plume magmatism, mantle convection and subduction–collision that occurred over approximately the past 3 Gyr (refs. <sup>6,36</sup>). Geophysical and petrological data indicate that the Archaean lithosphere is hydrated relative to the neighbouring Proterozoic lithosphere<sup>24,33,36</sup>, in part from a kimberlite intrusion phase during initial plume impingement. The 53–40-Myr-old Mwadui kimberlites in the eastern Tanzanian craton contain inclusions in diamonds that attest to the presence of mantle metasomatism around 10–25 Myr before EARS rifting<sup>37</sup>; these include a high concentration of TiO<sub>2</sub> in garnets and K<sub>2</sub>O in clinopyroxene, as well as a high ratio of lherzolitic to harzburgitic inclusions, indicating re-fertilization of the cratonic lithosphere at depths of >180 km. In northern Tanzania (Fig. 1a), trace element relationships of spinel peridotite xenoliths enriched in light rare-earth elements support mantle metasomatism through addition of carbonatite melts<sup>38</sup>.

Given the fertile and carbon-enriched lithospheric mantle in the eastern Tanzanian craton, the negligible CO<sub>2</sub> fluxes and crustal <sup>3</sup>He/<sup>4</sup>He ratios in the Manyara and Balangida basins could represent purely lower rates of magma production. However, comparison of geochemical data with existing geophysical results<sup>21</sup> suggests a potentially deep (>200 km) lithospheric control on melt and CO<sub>2</sub> transport in this section of the EARS. The modelled seismic velocity and density increase rapidly southwards of the Natron basin (Fig. 3), and the velocity changes coincide with the suture between the Tanzanian craton and Proterozoic Mozambique Belt<sup>21</sup>. Numerical models of buoyant and relatively hydrated cratonic mantle, such as the Tanzanian craton<sup>24,35,38</sup>, predict that edge convection does not develop at such a craton boundary<sup>19</sup>. Instead, the weak cratonic material is laterally advected for distances of up to 125 km towards mobile belts and then transported upwards near craton–mobile belt boundaries<sup>19,39</sup>. Mantle xenolith analyses provide evidence for thinning of the Tanzanian craton and the presence of Archaean lithosphere below the Natron Basin<sup>40</sup>. This mechanical dislodgment of deep-carbon-rich lithosphere increases the probability of carbonate melt production, through entrainment in warmer asthenosphere and related decompression as the material rises to the base of the mobile belt. If very low degrees of melting occur near the base of the craton at depths of around 200 km in even mildly oxidizing conditions, then melts will be strongly enriched in carbonate<sup>41</sup>. However, it is uncertain whether lateral migration of melts is efficient enough to concentrate so much carbon beneath the Magadi and Natron basins in the comparatively carbon-poor Proterozoic lithosphere. As such, we propose that displaced, carbon-enriched cratonic mantle lithosphere is present below the Magadi and Natron basins. This model best explains the high CO<sub>2</sub> flux and the cause of carbon-rich melt production at this location, and is favoured over edge-driven convection, which is suppressed in the presence of metasomatized cratonic mantle<sup>19</sup>.

In this model (Fig. 4), stores of fertile lithospheric mantle are removed from the base of the Archaean craton and transported laterally (guided by the slow northeast movement of Africa in a no-net rotation reference frame<sup>42</sup>). They also rise upwards into thinned lithosphere in the Natron and Magadi basins to drive melting of carbon-rich mantle. These deeper sections of cratonic mantle contain abundant diamonds trapped by redox freezing or carbonate in dykes of phlogopite pyroxenite<sup>9</sup>, which melt preferentially during upward movement, explaining the association of carbonate with potassium-rich magmatism. The sheared and thinned Proterozoic lithosphere then serves to focus and enhance melt and volatile migration through the plate<sup>15,35</sup>. Additionally, the zone of melting below the rift along the craton edge is consistently replenished in carbon from the lateral migration of fertile Archaean lithosphere (Fig. 4). This effect results in greater mantle CO<sub>2</sub> and melt fluxes at the surface in the Natron and Magadi basins. Active carbonatite magmatism occurs near this rheologic boundary (Fig. 2), where volatile sources abruptly transition south to north from the crust to the mantle. Our proposed model therefore explains why the Natron and Magadi basins represents the region of highest estimated tectonic degassing of mantle CO<sub>2</sub> in the EARS<sup>1</sup> (2.15–5.95 Mt yr<sup>-1</sup>) and has remained a region of persistent carbon-rich magmatism for approximately the past 4 Myr (for example, carbonatite volcanism at the Mosonik, Satiman, Shompole, Kerimasi and then Oldoinyo Lengai volcanoes; Fig. 2, Supplementary Table 4) and why some carbonatite melts are generated below the comparatively carbon-poor orogenic lithosphere near the craton border (for example, the Shompole volcano; Supplementary Table 4).

The mode of deep-carbon extraction and transport presented here has implications for constraining global outputs of CO<sub>2</sub> from extensional tectonic settings and the generation of carbon-rich magmas. Rare occurrences of carbonatite magmatism, such as that observed at Oldoinyo Lengai, do not depend exclusively on anomalously carbon-enriched mantle sources<sup>43</sup>, but may manifest owing to lithospheric conditions that allow the focusing of deep carbon. Occurrences of within-rift carbonatite volcanoes in northern Tanzania are

restricted to zones where the EARS intersects the thick Tanzanian craton (Fig. 2). Further north in Kenya and Uganda, carbonatite volcanism has manifested off-axis and continues to follow the north-northwest trending boundary of the Archaean Tanzanian and Kibalian cratons<sup>17</sup>, rather than the EARS, which trends north-northeast to south-southwest in the Kenyan and Main Ethiopian rifts. On the western side of the craton, similar trends are observed: large quantities of SO<sub>2</sub> (ref. 44) and CO<sub>2</sub> are emitted from Nyamuragira<sup>45</sup>; recent carbonatite volcanism at Katwe-Kikorongo and the Fort Portal volcanic field; and Miocene carbonatites in Rukwa between the Tanzanian and Bangweulu cratons<sup>16,46,47</sup> (Fig. 2; Supplementary Table 4). Together, these observations are consistent with the hypothesis that mantle carbon is transported towards, and concentrated along, the edges of Archaean cratons, where melt more efficiently ascends upwards along pressure gradients beneath thinner lithosphere through extending and more permeable lithosphere<sup>9</sup>. From here, crustal-scale, permeable faults<sup>1</sup> and carbonatite volcanism<sup>10</sup> facilitate release of sizable carbon stores in regions such as the Natron and Magadi basins.

Our proposed model of advective removal of fertile cratonic lithosphere (Fig. 4) can increase and sustain CO<sub>2</sub> fluxes in continental rifts situated adjacent to cratonic edges. This implies that global estimates of mantle CO<sub>2</sub> flux from continental rifts must account for both across- and along-strike variations in lithospheric thickness, composition and rheology in these systems.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2328-3>.

- Lee, H. et al. Massive and prolonged deep carbon emissions associated with continental rifting. *Nat. Geosci.* **9**, 145–149 (2016).
- Brune, S., Williams, S. E. & Müller, R. D. Potential links between continental rifting, CO<sub>2</sub> degassing and climate change through time. *Nat. Geosci.* **10**, 941–946 (2017).
- Tamburello, G., Pondrelli, S., Chioldini, G. & Rouwet, D. Global-scale control of extensional tectonics on CO<sub>2</sub> earth degassing. *Nat. Commun.* **9**, 4608 (2018).
- Dasgupta, R. & Hirschmann, M. M. The deep carbon cycle and melting in Earth's interior. *Earth Planet. Sci. Lett.* **298**, 1–13 (2010).
- Kellemen, P. B. & Manning, C. E. Reevaluating carbon fluxes in subduction zones, what goes down, mostly comes up. *Proc. Natl. Acad. Sci. USA* **112**, E3997–E4006 (2015).
- Foley, S. F. & Fischer, T. P. An essential role for continental rifts and lithosphere in the deep carbon cycle. *Nat. Geosci.* **10**, 897–902 (2017).
- Hunt, J. A., Zafu, A., Mather, T. A., Pyle, D. M. & Barry, P. H. Spatially variable CO<sub>2</sub> degassing in the main Ethiopian rift: implications for magma storage, volatile transport, and rift-related emissions. *Geochem. Geophys. Geosyst.* **18**, 3714–3737 (2017).
- Malusà, M. G. et al. Active carbon sequestration in the Alpine mantle wedge and implications for long-term climate trends. *Sci. Rep.* **8**, 4740–4748 (2018).
- Foley, S. F. Rejuvenation and erosion of the cratonic lithosphere. *Nat. Geosci.* **1**, 503–510 (2008).
- Brantley, S. L. & Koepenick, K. W. Measured carbon-dioxide emissions from Oldoinyo Lengai and the skewed distribution of passive volcanic fluxes. *Geology* **23**, 933–936 (1995).
- Sawyer, G. M., Carn, S. A., Tsanev, V. I., Oppenheimer, C. & Burton, M. Investigation into magma degassing at Nyiragongo volcano, Democratic Republic of the Congo. *Geochem. Geophys. Geosyst.* **9**, Q02017 (2008).
- Hutchison, W. et al. Structural controls on fluid pathways in an active rift system: a case study of the Aluto volcanic complex. *Geosphere* **11**, 542–562 (2015).
- van Wijk, J., van Hunen, J. & Goes, S. Small-scale convection during continental rifting: evidence from the Rio Grande rift. *Geology* **36**, 575–578 (2008).
- Bastow, I. D., Keir, D. & Daly, E. The Ethiopia Afar Geoscientific Lithospheric Experiment (EAGLE): probing the transition from continental rifting to incipient seafloor spreading. *Spec. Pap. Geol. Soc. Am.* **478**, 1–26 (2011).
- Ebbing, C. J. & Sleep, N. H. Cenozoic magmatism throughout east Africa resulting from impact of a single plume. *Nature* **395**, 788–791 (1998).
- Foley, S. F., Link, K., Tiberindwa, J. V. & Barifajio, E. Patterns and origin of igneous activity around the Tanzanian craton. *J. Afr. Earth Sci.* **62**, 1–18 (2012).
- Kalt, A., Hegner, E. & Satir, M. Nd, Sr, and Pb isotopic evidence for diverse lithospheric mantle sources of East African Rift carbonatites. *Tectonophysics* **278**, 31–45 (1997).
- Rosenthal, A., Foley, S. F., Pearson, D. G., Nowell, G. M. & Tappe, S. Petrogenesis of strongly alkaline primitive volcanic rocks at the propagating tip of the western branch of the East African Rift. *Earth Planet. Sci. Lett.* **284**, 236–248 (2009).
- Currie, C. A. & van Wijk, J. How craton margins are preserved: insights from geodynamic models. *J. Geodyn.* **100**, 144–158 (2016).

# Article

20. Roecker, S. et al. Subsurface images of the Eastern Rift, Africa, from the joint inversion of body waves, surface waves and gravity: investigating the role of fluids in early-stage continental rifting. *Geophys. J. Int.* **210**, 931–950 (2017).
21. Tiberi, C. et al. Lithospheric modification by extension and magmatism at the craton-ogenetic boundary: north Tanzania divergence, East Africa. *Geophys. J. Int.* **216**, 1693–1710 (2019).
22. Ebinger, C. J. et al. Crustal structure of active deformation zones in Africa: implications for global crustal processes. *Tectonics* **36**, 3298–3332 (2017).
23. Lee, C. T. & Rudnick, R. L. Compositionally stratified cratonic lithosphere: petrology and geochemistry of peridotite xenoliths from the Labait tuff cone, Tanzania. In *Proc. of the 7th International Kimberlite Conference* (eds Gurney, J. J. & Richardson, S. R.) 503–509 (National Book Printers, 1999).
24. Vauchez, A., Dineur, F. & Rudnick, R. Microstructure, texture and seismic anisotropy of the lithospheric mantle above a mantle plume: insights from the Labait volcano xenoliths (Tanzania). *Earth Planet. Sci. Lett.* **232**, 295–314 (2005).
25. Aulbach, S., Rudnick, R. L. & McDonough, W. F. Evolution of the lithospheric mantle beneath the East African Rift in Tanzania and its potential signatures in rift magmas. *Spec. Pap. Geol. Soc. Am.* **478**, 105–125 (2011).
26. Weinstein, A. et al. Fault-magma interactions during early continental rifting: seismicity of the Magadi-Natron-Manyara basins, Africa. *Geochem. Geophys. Geosyst.* **18**, 3662–3686 (2017).
27. Lee, H. et al. Incipient rifting accompanied by the release of subcontinental lithospheric mantle volatiles in the Magadi and Natron basin, East Africa. *J. Volcanol. Geotherm. Res.* **346**, 118–133 (2017).
28. Gautheron, C. & Moreira, M. Helium signature of the subcontinental lithospheric mantle. *Earth Planet. Sci. Lett.* **199**, 39–47 (2002).
29. Mana, S., Furman, T., Turrin, B. D., Feigenson, M. D. & Swisher, C. C. Magmatic activity across the East African North Tanzanian Divergence Zone. *J. Geol. Soc. Lond.* **172**, 368–389 (2015).
30. Halldórrsson, S. A., Hilton, D. R., Scarsi, P., Abebe, T. & Hopp, J. A common mantle plume source beneath the entire East African Rift System revealed by coupled helium–neon systematics. *Geophys. Res. Lett.* **41**, 2304–2311 (2014).
31. Furman, T. et al. Heads and tails: 30 million years of the Afar plume. *Geol. Soc. Lond. Spec. Publ.* **259**, 95–119 (2006).
32. Furman, T. Geochemistry of East African Rift basalts: an overview. *J. Afr. Earth Sci.* **48**, 147–160 (2007).
33. Pik, R., Marty, B. & Hilton, D. R. How many mantle plumes in Africa? The geochemical point of view. *Chem. Geol.* **226**, 100–114 (2006).
34. Hilton, D. R. et al. Helium isotopes at Rungwe Volcanic Province, Tanzania, and the origin of East African Plateaux. *Geophys. Res. Lett.* **38**, L21304 (2011).
35. Selway, K. Negligible effect of hydrogen content on plate strength in East Africa. *Nat. Geosci.* **8**, 543–546 (2015).
36. Koornneef, J. M. et al. Nature and timing of multiple metasomatic events in the sub-cratonic lithosphere beneath Labait, Tanzania. *Lithos* **112**, 896–912 (2009).
37. Stachel, T. & Brey, G. P. Rare and unusual mineral inclusions in diamonds from Mwadui, Tanzania. *Contrib. Mineral. Petrol.* **132**, 34–47 (1998).
38. Rudnick, R. L., McDonough, W. F. & Chappell, B. W. Carbonatite metasomatism in the northern Tanzanian mantle: petrographic and geochemical characteristics. *Earth Planet. Sci. Lett.* **114**, 463–475 (1993).
39. Huismans, R. & Beaumont, C. Depth-dependent extension, two-stage breakup and cratonic underplating at rifted margins. *Nature* **473**, 74–78 (2011).
40. Chesley, J. T., Rudnick, R. L. & Lee, C.-T. Re-Os systematics of mantle xenoliths from the East African Rift: age, structure, and history of the Tanzanian craton. *Geochim. Cosmochim. Acta* **63**, 1203–1217 (1999).
41. Foley, S. F. et al. The composition of near-solidus melts of peridotite in the presence of CO<sub>2</sub> and H<sub>2</sub>O between 40 and 60 kbar. *Lithos* **112**, 274–283 (2009).
42. Conrad, C. P. & Behn, M. D. Constraints on lithosphere net rotation and asthenospheric viscosity from global mantle flow models and seismic anisotropy. *Geochem. Geophys. Geosyst.* **11**, Q05W05 (2010).
43. Fischer, T. P. et al. Upper-mantle volatile chemistry at Oldoinyo Lengai volcano and the origin of carbonatites. *Nature* **459**, 77–80 (2009).
44. Carn, S. A., Fioletov, V. E., McLinden, C. A., Li, C. & Krotkov, N. A. A decade of global volcanic SO<sub>2</sub> emissions measured from space. *Sci. Rep.* **7**, 44095 (2017).
45. Sawyer, G. M., Carn, S. A., Tsanev, V. I., Oppenheimer, C. & Burton, M. Investigation into magma degassing at Nyiragongo volcano, Democratic Republic of the Congo. *Geochem. Geophys. Geosyst.* **9**, Q02017 (2008).
46. Eby, G. N., Lloyd, F. E. & Woolley, A. R. Geochemistry and petrogenesis of the Fort Portal, Uganda, extrusive carbonatite. *Lithos* **113**, 785–800 (2009).
47. Roberts, E. M. et al. Initiation of the western branch of the East African Rift coeval with the eastern branch. *Nat. Geosci.* **5**, 289–294 (2012).
48. Parks, M. M. et al. Distinguishing contributions to diffuse CO<sub>2</sub> emissions in volcanic areas from magmatic degassing and thermal decarbonation using soil gas <sup>222</sup>Rn–δ<sup>13</sup>C systematics: application to Santorini volcano, Greece. *Earth Planet. Sci. Lett.* **377**, 180–190 (2013).
49. Kennett, B. L. N. & Engdahl, E. R. Traveltimes for global earthquake location and phase identification. *Geophys. J. Int.* **105**, 429–465 (1991).
50. Dawson, J. B., James, D., Paslick, C. & Halliday, A. M. Ultrabasic potassic low-volume magmatism and continental rifting in north-central Tanzania: association with enhanced heat flow. *Russ. Geol. Geophys.* **38**, 69–81 (1997).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Methods

We use existing CO<sub>2</sub> flux, helium isotope and carbon isotope data from the Central Kenya Rift and the Natron, Magadi and Manyara basins, and recent data collected in the Natron, Manyara and Balangida basins (Supplementary Data Tables 1–3). To better interpret these data within the context of existing models of mantle volatile transport, we compare our results with a geophysically defined cratonic boundary from Tiberi et al.<sup>21</sup> and the distribution of carbonatite volcanoes, discussed below.

### Defining the craton boundary and uncertainty in the seismic velocity model

The boundary of the Tanzanian craton in the study region is broadly defined using the seismic velocity and density structure of the lithosphere derived from a joint analysis of gravity and teleseismic data from Tiberi et al.<sup>21</sup>, which shows a sharp velocity gradient, perhaps due to the presence of displaced low-density lower cratonic mantle and/or the presence of melt and other fluids<sup>20,21,26</sup>. The boundary in Fig. 3 (also shown in Extended Data Fig. 1) is presented as a broad zone to account for the complex three-dimensional nature of the boundary (compared to the two-dimensional cross-section line shown) and the uncertainty in the model resolution. A slightly modified version of Fig. 3 is presented here to facilitate comparison of the possible location of the craton boundary from the seismic velocity model (Extended Data Fig. 1). Teleseismic tomography suffers from a lack of vertical resolution, particularly within the upper 30 km. However, its lateral resolution is high and mostly depends on the station interspacing. Synthetic tests show that an interstation length of ~20 km enables resolution of lateral boundaries to 15 km in the upper 100 km of the velocity model (see supplementary materials in Tiberi et al.<sup>21</sup>). Moreover, the authors combine seismic delay times with gravity data in the inversion scheme, which increases crustal resolution. Higher resolution crustal imaging at depths of <40 km can be found in Roecker et al.<sup>20</sup> and Ibs-von Seht et al.<sup>51</sup>. Crustal structures as narrow as 6 km can be resolved in the joint inversion of arrival times, ambient noise and gravity in Roecker et al.<sup>20</sup>.

### Areas of carbonatite volcanism associated with the EARS

We set an age threshold of 45 Myr for carbonatite volcanism associated with the development of the EARS, which is consistent with Ernst & Bell<sup>52</sup>. Although the age and initiation of the EARS is often linked with the eruption of Oligocene trap lavas between ~33.9 Ma and 27 Ma (refs. 32,53,54), 45–34-Myr-old flood lavas are identified in the Turkana region (North Kenya) and the Southern Ethiopian plateau with geochemical affinities to 33.9–22-Myr-old EARS lavas, suggesting a potentially longer-lived thermo-chemical anomaly associated with EARS rifting<sup>32,54</sup>.

Supplementary Data Table 4 provides a summary of the location and age of <45-Myr-old carbonatite centres associated with the EARS<sup>29,47,52,55–62</sup>. A few of these volcanoes in Uganda have no reported ages, but form part of a sub-linear belt of carbonatite volcanoes, all with reported ages <45 Myr. Therefore, we considered these carbonatite systems to be associated with magmatism and rifting in the EARS, consistent with the interpretation of Ernst & Bell<sup>52</sup> for these carbonatites. Additionally, the exact location of <45-Myr-old carbonatite volcanoes in the Rukwa/Rungwe region are tentatively assigned, because they are identified only through observations of distal carbonatite tephra deposits<sup>47</sup>, with only Mesozoic ages assigned to mapped carbonatite volcanoes.

### Model of active mantle volatile degassing in the Natron and Magadi basins

Previous geochemical and geophysical studies investigating active seismicity and volatile discharge in the Natron and Magadi basins support mantle volatile transport through fault and fracture networks,

which is further corroborated by recent helium and carbon isotope data presented in this study. Recent seismic station deployments (for example, the CRAFTI network<sup>26</sup>) have detected persistent low-magnitude earthquakes since 2013 that span the full range of the crust. Abundant seismic activity occurs at mid- to lower-crustal depths (for example, 15–40 km), where P- and S-wave velocities ( $V_p$  and  $V_s$ , respectively) suggest the presence of discrete magma chambers at various depths. Lower-crustal earthquakes occur in regions of low  $V_p/V_s$ , indicating the presence of compressible pore fluids<sup>20</sup>, and correspond to deeply penetrating border faults that actively emit mantle volatiles<sup>1,27</sup>. Collectively, these data support the presence of deep crustal magma bodies in the Natron and Magadi basins, which are probably generated from a sub-continental lithospheric melt source<sup>27,29</sup> that actively releases mantle volatiles into the overlying extensional fault system<sup>1,26,58</sup>. On the basis of these observations, a tectonic degassing model<sup>63</sup> has been proposed for the Natron and Magadi basin<sup>1</sup>, where mantle volatiles (including CO<sub>2</sub>) escape along fault and fracture networks at the surface and in regions not directly affected by volcanism.

### Laboratory analyses of thermal springs

Thermal-spring samples were collected in copper tubes and analysed for <sup>3</sup>He/<sup>4</sup>He ratios on noble gas mass spectrometers (Helix-SFT and VG-5400) at the Atmosphere and Ocean Research Institute of the University of Tokyo (AORI). Both He and Ne were purified using hot titanium getters (at 400 °C) and charcoal traps (at liquid-nitrogen temperature), and Ne was removed using a cryogenic trap (at 40 K) after measuring <sup>4</sup>He/<sup>20</sup>Ne ratios using a quadrupole mass spectrometer (Prisma 80 QMS, Pfeiffer). The experimental errors for <sup>3</sup>He/<sup>4</sup>He and <sup>4</sup>He/<sup>20</sup>Ne ratios are about 1% and 5% (1 $\sigma$ ; SFT-Helix) and about 3.5% and 5% (1 $\sigma$ ; VG5400), respectively<sup>64</sup>.

### Flux measurement and analyses of carbon isotopes of diffuse CO<sub>2</sub>

The flux of diffusely degassing CO<sub>2</sub> in each study location was measured using EGM-4 and EGM-5 gas analysers (PP Systems), which have a measurement range of 0–30,000 ppm and <1% error. For each measurement, a cylindrical chamber with a volume of  $1.18 \times 10^{-3} \text{ m}^3$  was placed on the ground forming a tight seal, and the CO<sub>2</sub> flux was calculated using changes in CO<sub>2</sub> concentration over a 120-s timespan. Diffuse gas samples were also collected by diverting gas from the chamber into pre-evacuated glass vials, and CO<sub>2</sub> concentrations were recorded at the time of sampling. Gas samples were analysed using a Finnigan Delta XL Isotope Ratio Mass Spectrometer with a gas bench and auto-sampler at the Center for Stable Isotopes, University of New Mexico. Measured values are reported relative to the Vienna Pee Dee belemnite (VPDB) standard and presented as  $\delta^{13}\text{C}$  values in parts per thousand (‰) in Fig. 1. Fractionation effects between ionized carbon (HCO<sub>3</sub><sup>-</sup> and CO<sub>3</sub><sup>2-</sup>) and CO<sub>2</sub> gases in the Natron and Magadi basin hydrothermal systems were examined by Lee et al.<sup>27</sup>, who found corrected  $\delta^{13}\text{C}$  values ranging from -0‰ to -7‰. These values are consistent with predominantly mantle-derived carbon and are within the range of soil gas samples collected in Natron and Magadi for this study, which range from -2.2‰ to 10.1‰.

### Differentiating CO<sub>2</sub> flux populations

Discrimination of different sources of diffusely degassing CO<sub>2</sub> in volcano-tectonic settings is often achieved through a combination of carbon isotope analyses and statistical analyses of CO<sub>2</sub> flux populations<sup>1,65,66</sup>. This process differentiates two or more CO<sub>2</sub> flux populations exhibiting log-normal distributions, with the highest mean flux often attributed to a magmatic population<sup>10</sup>, which can be corroborated through isotopic analyses<sup>66</sup>. To make these subdivisions, we use the method of Sinclair<sup>67</sup> adopted in Chiodini et al.<sup>65</sup>, where flux populations are delineated on either side of a distinct inflection point on logarithmic probability plots (Extended Data Fig. 2).

# Article

## Data availability

All data generated or analysed during this study are provided with this article and in Supplementary Tables 1–4. The SRTM digital elevation model used to generate maps is publicly available at [http://srtm.cgiar.org/srtmdata](http://srtm.csi.cgiar.org/srtmdata). The recently analysed and previously unpublished CO<sub>2</sub> flux and isotopic data (<https://doi.org/10.26022/IEDA/111520>) from the 2018 Tanzania field campaign can be found at <http://www.earthchem.org>.

51. Ibs-von Seht, M., Blumenstein, S., Wagner, R., Hollnack, D. & Wohlenberg, J. Seismicity, seismotectonics and crustal structure of the southern Kenya Rift—new data from the Lake Magadi area. *Geophys. J. Int.* **146**, 439–453 (2001).
52. Ernst, R. E. & Bell, K. Large igneous provinces (LIPs) and carbonatites. *Mineral. Petrol.* **98**, 55–76 (2010).
53. Hofmann, C. et al. Timing of the Ethiopian flood basalt event and implications for plume birth and global change. *Nature* **389**, 838–841 (1997).
54. Rooney, T. O. The Cenozoic magmatism of East-Africa: Part I – flood basalts and pulsed magmatism. *Lithos* **286–287**, 264–301 (2017).
55. Woolley, A. R. & Kjarsgaard, B. A. *Carbonatite Occurrences of the World* (Geological Survey of Canada, 2008).
56. Fairhead, J. D., Mitchell, J. G. & Williams, L. A. J. New K/Ar determinations on rift volcanics of S. Kenya and their bearing on age of rift faulting. *Science* **238**, 66–69 (1972).
57. Le Gall, B. et al. Rift propagation at craton margin. Distribution of faulting and volcanism in the North Tanzanian Divergence (East Africa) during Neogene times. *Tectonophysics* **448**, 1–19 (2008).
58. Muirhead, J. D. et al. Evolution of upper crustal faulting assisted by magmatic volatile release during early-stage continental rift development in the East African Rift. *Geosphere* **12**, 1670–1700 (2016).
59. Sherrod, D. R., Magigita, M. M. & Kwelwa, S. *Geologic Map of Oldonyo Lengai (Oldoinyo Lengai) and Surroundings, Arusha Region, United Republic of Tanzania*. Report No. 1306 (U.S. Geological Survey, 2013).
60. Muirhead, J. D., Kattenhorn, S. A. & Le Corvec, N. Varying styles of magmatic strain accommodation in the East African Rift. *Geochem. Geophys. Geosyst.* **16**, 2775–2795 (2015).
61. Bagdasaryan, G., Gerasimovskiy, V. I., Polykov, A. I. & Gukasyan, R. K. Age of volcanic rocks in the rift zones of East Africa. *Geokhimiya* **1**, 84–90 (1973).
62. Mollel, G. F. Petrochemistry and Geochronology of Ngorongoro Volcanic Highland Complex (NVHC) and its Relationship to Laetoli and Olduvai Gorge, Tanzania. PhD thesis, Rutgers Univ. (2007).
63. Burton, M. R., Sawyer, G. M. & Granieri, D. in *Carbon in Earth* (eds Hazen, R. M. et al.) 323–354 (De Gruyter, 2013).
64. Sano, Y., Tokutake, T. & Takahata, N. Accurate measurement of atmospheric helium isotopes. *Anal. Sci.* **24**, 521–525 (2008).
65. Chioldini, G., Cioni, R., Guidi, M., Raco, B. & Marini, L. Soil CO<sub>2</sub> flux measurements in volcanic and geothermal areas. *Appl. Geochem.* **13**, 543–552 (1998).
66. Chioldini, G. et al. Carbon isotopic composition of soil CO<sub>2</sub> efflux, a powerful method to discriminate different sources feeding soil CO<sub>2</sub> degassing in volcanic-hydrothermal areas. *Earth Planet. Sci. Lett.* **274**, 372–379 (2008).
67. Sinclair, J. A. Selection of thresholds in geochemical data using probability graphs. *J. Geochem. Explor.* **3**, 129–149 (1974).

**Acknowledgements** This work was funded by the NSF EAR GeoPRISMS Program, grant numbers 1654518 (J.D.M.), 1654433 (T.P.F.) and 1836651 (J.D.), Deutsche Forschungsgemeinschaft (DFG) grant RE 4321/1-1 (M.C.R.) and the Marshall-Heape fund at Tulane University (C.J.E.). We are grateful to COSTECH and Tanzania Wildlife Research Institute for permitting us to conduct research in Tanzania. We thank E. Saria and K. Nkembo for assistance during fieldwork in the Lake Natron region in 2018, G. Kianji for assistance during collection of data in 2014, and K. Rahilly for assisting with field planning.

**Author contributions** The initial project was conceived by J.D.M., T.P.F., C.J.E. and J.D., with planning and execution of field data collection by J.D.M., T.P.F., C.J.E., A.L., S.J.O., E.K. and M.C.R. CO<sub>2</sub> flux data were compiled and analysed by J.D.M., T.P.F., E.J.J., S.J.O. and A.L., and laboratory analyses of helium and carbon isotopes were performed by T.P.F., H.L., Y.S. and N.T. Compilation and examination of geophysical and gas chemical data were conducted by C.T., J.D.M., C.J.E., J.v.W. and C.A.C. The final model presented in Fig. 4 was conceived and designed by T.P.F., J.D.M., C.J.E., J.v.W., C.A.C. and S.F.F. The manuscript was written by J.D.M. and T.P.F. with contributions from all co-authors.

**Competing interests** The authors declare no competing interests.

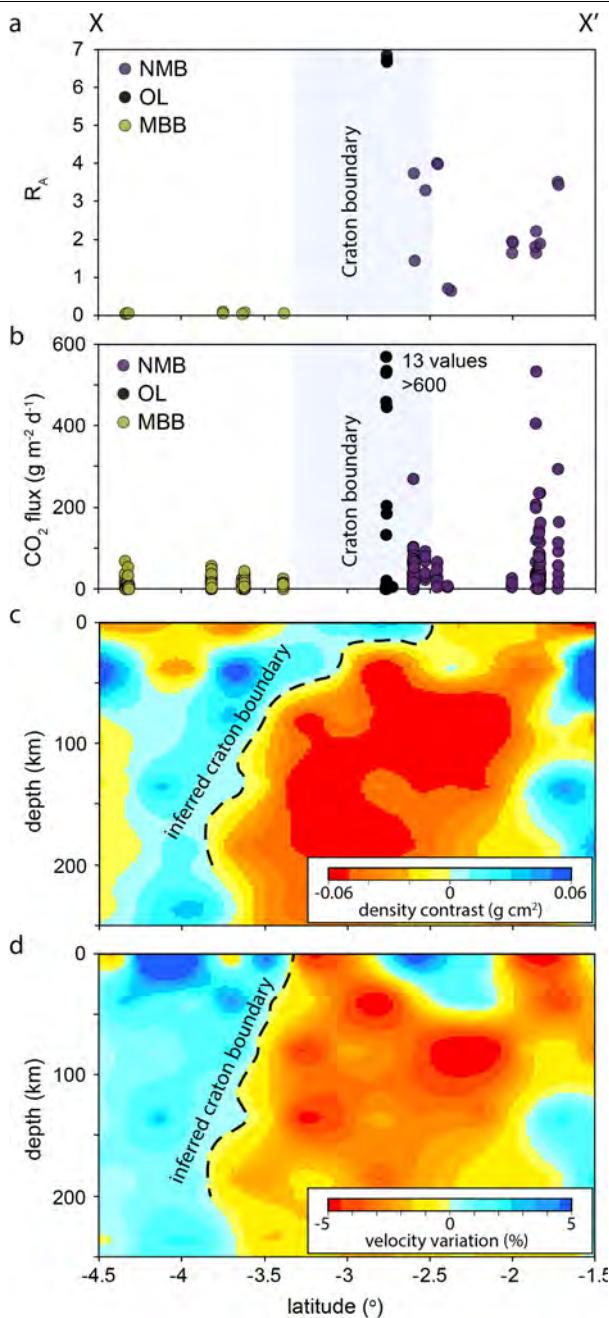
### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2328-3>.

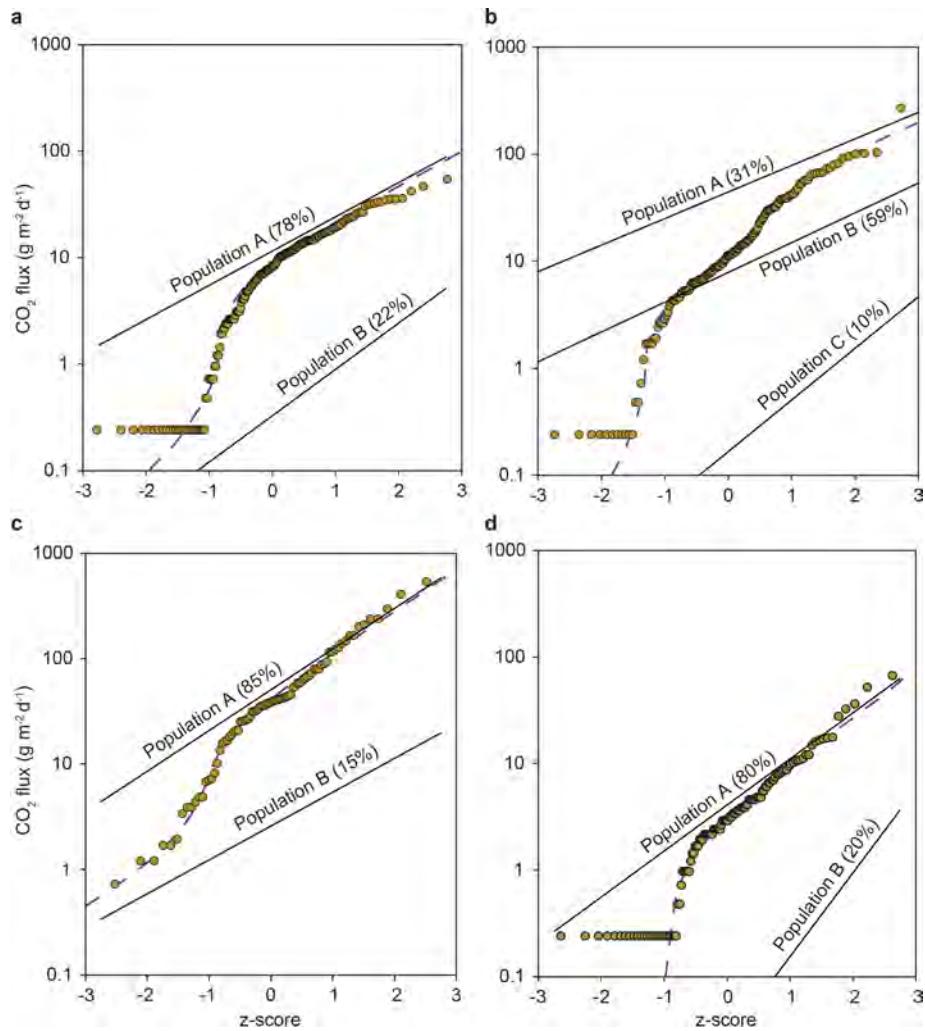
**Correspondence and requests for materials** should be addressed to J.D.M. or T.P.F.

**Peer review information** *Nature* thanks Sascha Brune, Giovanni Chioldini and Tanya Furman for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**Extended Data Fig. 1 | Latitudinal variations in  $\text{CO}_2$  flux and  $R_A$  values**  
 $(^3\text{He}/^4\text{He})$  with respect to the modelled lithosphere structure. **a**,  $R_A$  versus latitude along the Eastern Rift from  $-1.5^{\circ}$  to  $-4.5^{\circ}$ . Air-contaminated samples (TZ18-C16) were removed from the analysis. The Tanzanian craton boundary (light grey) is interpreted using the sharp density (**c**) and velocity (**d**) contrasts near the northern Manyara and southern Natron basins. It is marked as a broad region to account for location uncertainties and the overall three-dimensional nature of the boundary in the region. **b**, Diffuse  $\text{CO}_2$  flux versus latitude along the Eastern Rift from  $-1.5^{\circ}$  to  $-4.5^{\circ}$  for sample sites presented in Fig. 1a and the Oldoinyo Lengai volcano. **c**, Lithosphere density model from Tiberi et al.<sup>21</sup>. **d**, Lithosphere velocity model from Tiberi et al.<sup>21</sup> (model resolution discussed in Methods). The density and velocity contrasts are relative to the IASP91 model<sup>49</sup>. The position of the cross-section is shown in Fig. 1a.



**Extended Data Fig. 2 |  $\text{CO}_2$  flux population analyses.** **a–d,** Shown are results for the Manyara (a), Natron (b), Magadi (c) and Balangida (d) basins. The analyses were performed in line with the method of Sinclair<sup>67</sup> outlined in

Chiodini et al.<sup>65</sup>, with the dashed lines representing a modelled mixed population based on the distributions of the lower, higher and occasionally intermediate flux populations.

## Article

# Repositioning of the global epicentre of non-optimal cholesterol

<https://doi.org/10.1038/s41586-020-2338-1>

Received: 18 October 2019

Accepted: 2 April 2020

Published online: 3 June 2020

Open access

 Check for updates

NCD Risk Factor Collaboration (NCD-RisC)\*

High blood cholesterol is typically considered a feature of wealthy western countries<sup>1,2</sup>. However, dietary and behavioural determinants of blood cholesterol are changing rapidly throughout the world<sup>3</sup> and countries are using lipid-lowering medications at varying rates. These changes can have distinct effects on the levels of high-density lipoprotein (HDL) cholesterol and non-HDL cholesterol, which have different effects on human health<sup>4,5</sup>. However, the trends of HDL and non-HDL cholesterol levels over time have not been previously reported in a global analysis. Here we pooled 1,127 population-based studies that measured blood lipids in 102.6 million individuals aged 18 years and older to estimate trends from 1980 to 2018 in mean total, non-HDL and HDL cholesterol levels for 200 countries. Globally, there was little change in total or non-HDL cholesterol from 1980 to 2018. This was a net effect of increases in low- and middle-income countries, especially in east and southeast Asia, and decreases in high-income western countries, especially those in northwestern Europe, and in central and eastern Europe. As a result, countries with the highest level of non-HDL cholesterol—which is a marker of cardiovascular risk—changed from those in western Europe such as Belgium, Finland, Greenland, Iceland, Norway, Sweden, Switzerland and Malta in 1980 to those in Asia and the Pacific, such as Tokelau, Malaysia, The Philippines and Thailand. In 2017, high non-HDL cholesterol was responsible for an estimated 3.9 million (95% credible interval 3.7 million–4.2 million) worldwide deaths, half of which occurred in east, southeast and south Asia. The global repositioning of lipid-related risk, with non-optimal cholesterol shifting from a distinct feature of high-income countries in northwestern Europe, north America and Australasia to one that affects countries in east and southeast Asia and Oceania should motivate the use of population-based policies and personal interventions to improve nutrition and enhance access to treatment throughout the world.

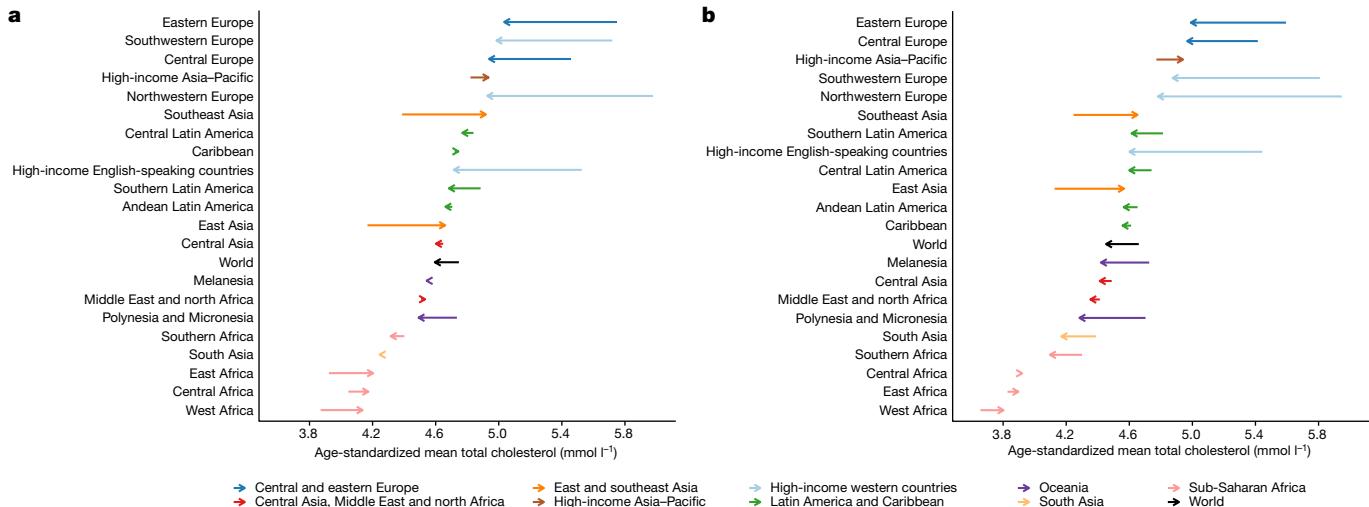
Blood cholesterol is one of the most important risk factors for ischaemic heart disease (IHD) and ischaemic stroke<sup>4–6</sup>. Consistent and comparable information on cholesterol levels and trends in different countries can help to benchmark national performance in addressing non-optimal cholesterol, investigate the reasons behind differential trends and identify countries in which interventions are needed the most.

A previous global analysis<sup>7</sup> reported trends in total cholesterol from 1980 to 2008, but did not analyse important lipid fractions—including HDL and non-HDL cholesterol—that are key to understanding the cardiovascular disease risk associated with non-optimal cholesterol. Dietary and behavioural determinants of cholesterol have changed throughout the world in the past decades, including a worldwide rise in adiposity<sup>8,9</sup>, divergent global trends in alcohol use<sup>10</sup>, a rise in the intake of animal-source foods in middle-income countries (especially in east Asia)<sup>3,11</sup>, and a replacement of saturated fats and trans fats with unsaturated fats in some high-income countries<sup>3,11,12</sup>. There is also considerable variation in how much different

countries have adopted lipid-lowering medications<sup>13</sup>. These changes are likely to have influenced cholesterol levels substantially in the decade since the last estimates were made. Furthermore, HDL and non-HDL cholesterol, which have opposite associations with cardiovascular diseases<sup>4,5</sup>, respond differently to diet and treatment, and may therefore have different geographical patterns and trends over time<sup>14</sup>. Information on these major lipid fractions, which were not included in the previous global estimates, is essential for priority setting and intervention choice.

Here we pooled 1,127 population-based studies that measured blood lipids in 102.6 million individuals aged 18 years and older (Extended Data Figs. 1, 2 and Supplementary Table 1) and used a Bayesian hierarchical model to estimate trends from 1980 to 2018 in mean total, non-HDL and HDL cholesterol levels for 200 countries. We also estimated the number of deaths caused by IHD and ischaemic stroke that were attributable to high levels of non-HDL cholesterol using information on its hazards from epidemiological studies.

\*A list of participants and their affiliations appears in the online version of the paper.



**Fig. 1 | Change in age-standardized mean total cholesterol between 1980 and 2018 by region for women and men.** **a**, Age-standardized mean total cholesterol in women. **b**, Age-standardized mean total cholesterol in men.

The start of the arrow shows the level in 1980 and the head indicates the level in 2018. See Extended Data Fig. 3 for age-standardized mean HDL cholesterol. One  $\text{mmol l}^{-1}$  is equivalent to  $38.61 \text{ mg dl}^{-1}$ .

## Trends in total cholesterol

In 2018, global age-standardized mean total cholesterol was  $4.6 \text{ mmol l}^{-1}$  (95% credible interval,  $4.5\text{--}4.7$ ) for women and  $4.5 \text{ mmol l}^{-1}$  ( $4.3\text{--}4.6$ ) for men. Global age-standardized mean total cholesterol changed little over these nearly four decades, decreasing by  $0.03 \text{ mmol l}^{-1}$  per decade ( $-0.02\text{--}0.08$ ) in women and  $0.05 \text{ mmol l}^{-1}$  per decade ( $0.00\text{--}0.11$ ) in men (posterior probability of the observed declines being true declines =  $0.90$  for women and  $0.98$  for men) (Fig. 1). Regionally, total cholesterol decreased the most in high-income western regions and in central and eastern Europe. The decrease was the largest (around  $0.3 \text{ mmol l}^{-1}$  per decade; posterior probability  $>0.9999$ ) in northwestern Europe, where mean total cholesterol levels had been the highest in 1980. The decrease in total cholesterol in high-income western regions and central and eastern Europe was largely due to a decline in non-HDL cholesterol (Extended Data Fig. 4), which among women was offset partly by an increase in mean HDL cholesterol levels. Mean total cholesterol changed little in most of the other regions, with the notable exception of east and southeast Asia, where it increased by more than  $0.1 \text{ mmol l}^{-1}$  per decade in both women and men (posterior probability  $\geq 0.95$ ). The increase in east and southeast Asia was largely due to an increase in non-HDL cholesterol.

## Trends in non-HDL and HDL cholesterol

In 2018, global age-standardized mean non-HDL cholesterol was  $3.3 \text{ mmol l}^{-1}$  ( $3.2\text{--}3.4$ ) for women and  $3.3 \text{ mmol l}^{-1}$  ( $3.3\text{--}3.4$ ) for men; global age-standardized mean HDL cholesterol was  $1.3 \text{ mmol l}^{-1}$  ( $1.2\text{--}1.3$ ) for women and  $1.1 \text{ mmol l}^{-1}$  ( $1.1\text{--}1.2$ ) for men. Global age-standardized mean non-HDL cholesterol remained almost unchanged from 1980 to 2018, decreasing by only  $0.02 \text{ mmol l}^{-1}$  per decade ( $-0.02\text{--}0.06$ ; posterior probability =  $0.80$ ) in women and  $0.01 \text{ mmol l}^{-1}$  per decade ( $-0.03\text{--}0.06$ ; posterior probability =  $0.72$ ) in men. Global age-standardized mean HDL cholesterol remained unchanged for women and decreased slightly for men (by  $0.02 \text{ mmol l}^{-1}$  per decade, posterior probability =  $0.91$ ).

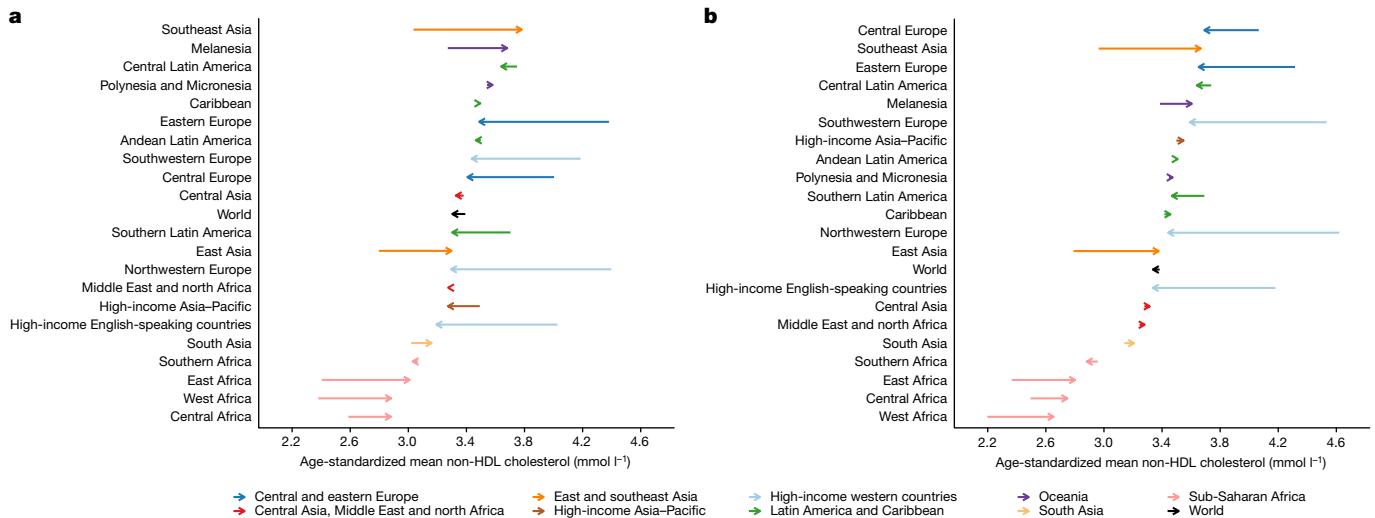
Regionally, non-HDL cholesterol decreased substantially in high-income western regions and central and eastern Europe. The largest decrease occurred in northwestern Europe ( $>0.3 \text{ mmol l}^{-1}$  per decade; posterior probability  $>0.9999$ ) (Fig. 2). By contrast, it increased in east and southeast Asia, parts of sub-Saharan Africa and Melanesia. The increase was the largest in southeast Asia, increasing by

approximately  $0.2 \text{ mmol l}^{-1}$  per decade (posterior probability  $>0.9999$ ). Mean HDL cholesterol increased in the high-income Asia-Pacific region, by as much as  $0.1 \text{ mmol l}^{-1}$  per decade in women (posterior probability  $>0.9999$ ) but decreased in Melanesia, Polynesia and Micronesia (Extended Data Fig. 3).

Belgium, Finland, Greenland, Iceland, Norway, Sweden, Switzerland and Malta had some of the highest non-HDL cholesterol levels in 1980 ( $>4.5 \text{ mmol l}^{-1}$  in women and  $>4.7 \text{ mmol l}^{-1}$  in men) but experienced some of the largest declines (Figs. 3, 4). At the extreme, mean non-HDL cholesterol declined by around  $0.45 \text{ mmol l}^{-1}$  per decade or more in Belgian and Icelandic women and men, changing their ranks from being in the top 10 countries in terms of non-HDL cholesterol in 1980 to being ranked in the lower half of the countries in 2018—below countries in southwestern Europe such as France and Italy. The largest increases were found in east Asian countries (for example, China) and southeast Asian countries (for example, Indonesia, Thailand, Malaysia, Cambodia and Lao PDR). In these countries, age-standardized mean non-HDL cholesterol increased by as much as  $0.23 \text{ mmol l}^{-1}$  per decade. As a result of these opposite trends, countries with the highest age-standardized mean non-HDL cholesterol levels in 2018 were all outside northwestern Europe: Tokelau, Malaysia, The Philippines and Thailand, all of which had mean non-HDL cholesterol around or above  $4 \text{ mmol l}^{-1}$ . China, which had one of the lowest mean non-HDL cholesterol levels in 1980, reached or surpassed non-HDL cholesterol levels of many high-income western countries in 2018. Sub-Saharan African countries had the lowest mean non-HDL cholesterol in 2018, as low as  $2.6 \text{ mmol l}^{-1}$  in some countries, as they had in 1980. Not only did high-income countries benefit from decreasing non-HDL cholesterol levels, they had higher mean HDL cholesterol than low- and middle-income countries (Extended Data Fig. 6).

## Deaths attributable to non-optimal cholesterol

In 2017, high non-HDL cholesterol was responsible for an estimated 3.9 million ( $3.7\text{--}4.2$  million) worldwide deaths from IHD and ischaemic stroke (Fig. 5), accounting for a third of deaths from these causes. From 1990 to 2017, the number of deaths caused by IHD and ischaemic stroke that were attributable to high non-HDL cholesterol increased by around 910,000 globally. This increase was a net result of a large decrease in western countries, from 950,000 (890,000–990,000) to 480,000 (430,000–530,000), and a large increase throughout Asia. In particular, the number of deaths attributable to high non-HDL cholesterol more

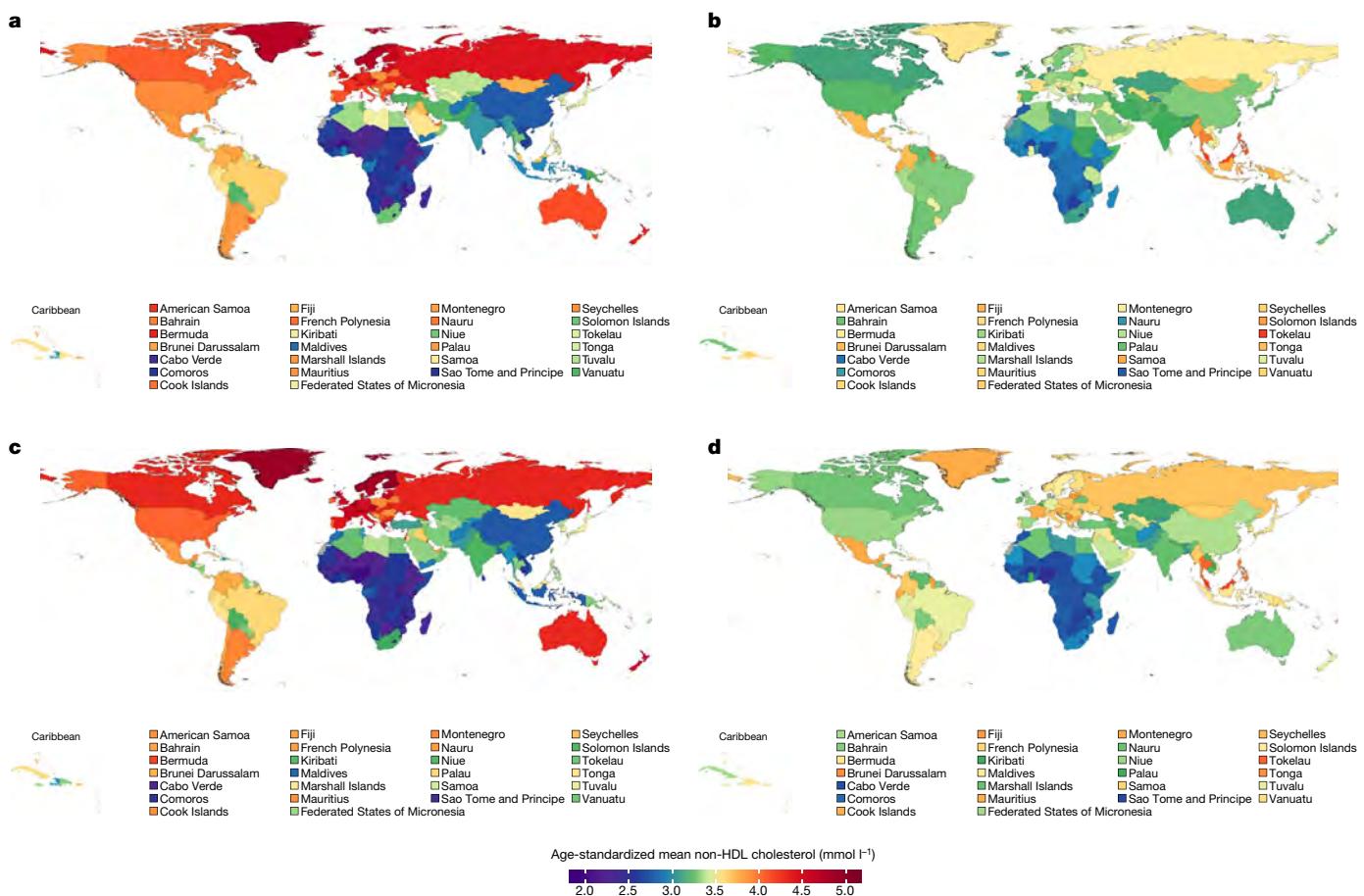


**Fig. 2 | Change in age-standardized mean non-HDL cholesterol between 1980 and 2018 by region for women and men.** **a**, Age-standardized mean non-HDL cholesterol in women. **b**, Age-standardized mean non-HDL

cholesterol in men. The start of the arrow shows the level in 1980 and the head indicates the level in 2018. See Extended Data Fig. 3 for age-standardized mean HDL cholesterol. One mmol l<sup>-1</sup> is equivalent to 38.61 mg dl<sup>-1</sup>.

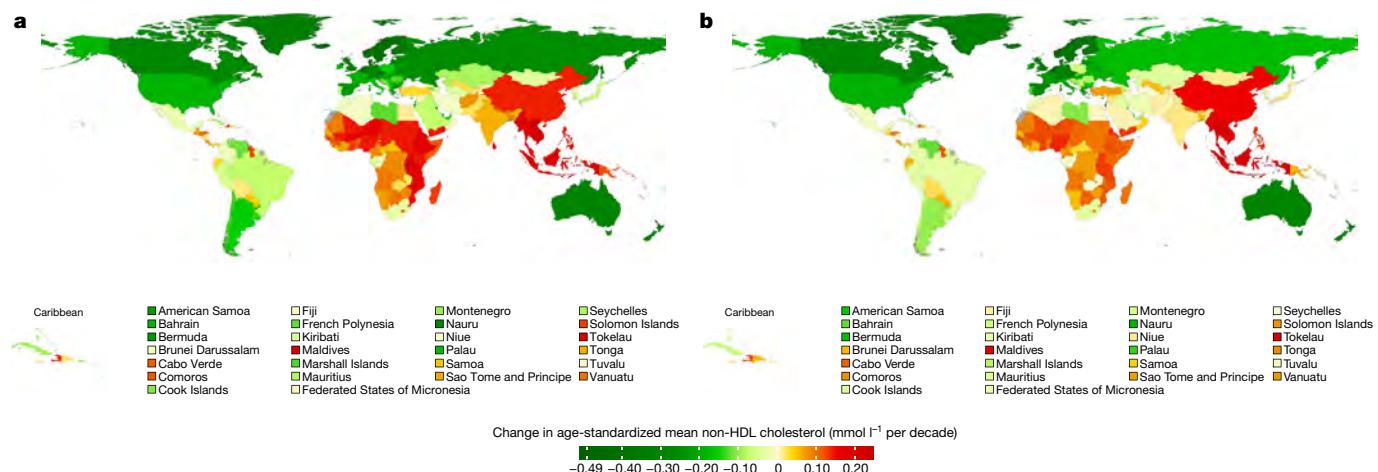
than tripled in east Asia, from 250,000 (230,000–270,000) to 860,000 (770,000–940,000), and more than doubled in southeast Asia, from 110,000 (100,000–120,000) to 310,000 (290,000–330,000). As a

result, by 2017 east, southeast and south Asia accounted for half of all deaths attributable to high non-HDL cholesterol, compared with a quarter in 1990.



**Fig. 3 | Age-standardized mean non-HDL cholesterol by country in 1980 and 2018 for women and men.** **a**, Age-standardized mean non-HDL cholesterol in women in 1980. **b**, Age-standardized mean non-HDL cholesterol in women in 2018. **c**, Age-standardized mean non-HDL cholesterol in men in 1980.

**d**, Age-standardized mean non-HDL cholesterol in men in 2018. See Extended Data Fig. 5 for age-standardized mean total cholesterol and Extended Data Fig. 6 for age-standardized mean HDL cholesterol. One mmol l<sup>-1</sup> is equivalent to 38.61 mg dl<sup>-1</sup>.



for change per decade in age-standardized mean total cholesterol and Extended Data Fig. 8 for change per decade in age-standardized mean HDL cholesterol. One  $\text{mmol L}^{-1}$  is equivalent to  $38.61 \text{ mg dL}^{-1}$ .

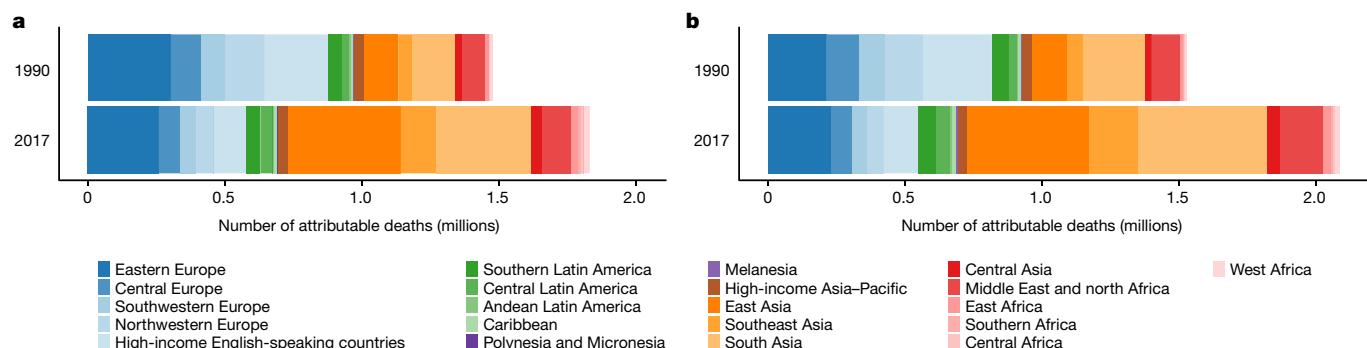
## Implications

Our results show that over the past nearly four decades, there has been a major global repositioning of lipid-related risk, with non-optimal cholesterol patterns shifting from being a distinct feature of high-income countries in northwestern Europe, north America and Australasia to one that affects middle-income countries in east and southeast Asia, as well as some countries in Oceania and central Latin America. This transition is especially noticeable for non-HDL cholesterol, which had not been quantified previously in a global analysis. This global repositioning has occurred as a consequence of opposing trends in high-income western countries and in Asia, which has led to some Asian countries having the highest worldwide non-HDL cholesterol levels in 2018.

The decrease in non-HDL cholesterol in western countries started in the 1980s, before statins were widely used<sup>15,16</sup>. This indicates that changes in diet, especially the replacement of saturated with unsaturated fats<sup>3,17–21</sup> and reduction in trans fats<sup>12,17,22</sup>, are major contributors to this decline. Nonetheless, the increased use of statins from the late 1990s onwards<sup>15,16</sup>, may explain up to one half of the decrease in those countries in which statins are widely used<sup>19,23,24</sup>. In contrast to high-income western countries, the consumption of animal-source foods, refined carbohydrates and palm oil has increased substantially in east and southeast Asia<sup>3,25,26</sup>, where statin use remains low<sup>13,27</sup>. For example, the Pearson correlation coefficient between the change in non-HDL cholesterol and the change in a multi-dimensional score of animal-source foods and sugar<sup>3</sup> was 0.69 for women and 0.67 for

men using data from high-income western countries and countries in east and southeast Asia, the two regions that experienced the largest decrease and increase, respectively, in non-HDL cholesterol levels. Finally, changes in diet, especially a decrease in carbohydrate and an increase in fat intake<sup>28–31</sup>, may have contributed to the large increase in HDL cholesterol observed in the high-income Asia-Pacific region, where there was little increase in overweight and obesity relative to other regions<sup>8,9</sup>. By contrast, the large increase in diabetes<sup>32</sup> and adiposity<sup>8</sup> in Oceania may have contributed to the decrease in HDL cholesterol in this region. The Pearson correlation coefficient between the change in HDL cholesterol and the change in body-mass index<sup>8</sup> was  $-0.87$  for women and  $-0.69$  for men using countries in the high-income Asia-Pacific region and Oceania, the two regions that had the largest increase and decrease, respectively, in HDL cholesterol; the Pearson correlation coefficient for the change in HDL cholesterol and change in diabetes prevalence<sup>32</sup> was  $-0.84$  for women and  $-0.69$  for men. In the same regions, the Pearson correlation coefficient between the change in non-HDL cholesterol and the change in body-mass index<sup>8</sup> was  $0.77$  for women and  $0.62$  for men; for the change in non-HDL cholesterol and the change in diabetes prevalence<sup>32</sup>, the Pearson correlation coefficient was  $0.54$  for women and  $0.40$  for men.

Although it has previously been documented that the prevalence of adiposity<sup>8,9</sup>, diabetes<sup>32</sup> and high blood pressure<sup>33</sup> is now higher in low- and middle-income countries than in high-income countries, higher cholesterol is commonly considered to be a feature of affluent western nations<sup>1,2</sup>. We show that, when focusing on non-HDL cholesterol,



middle-income countries have emerged as the new global epicentre of non-optimal cholesterol as they did for other major cardiovascular disease risk factors, indicating that there is no such a thing as a western risk factor. At the same time, the populations of high-income countries would also benefit from further lowering non-HDL cholesterol. Therefore, population-based policies and personal interventions to improve nutrition and enhance treatment are now needed in all countries, especially as a part of the movement towards universal health coverage.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2338-1>.

1. Danaei, G. et al. The global cardiovascular risk transition: associations of four metabolic risk factors with national income, urbanization, and Western diet in 1980 and 2008. *Circulation* **127**, 1493–1502 (2013).
2. Ezzati, M. et al. Rethinking the “diseases of affluence” paradigm: global patterns of nutritional risks in relation to economic development. *PLoS Med.* **2**, e133 (2005).
3. Bentham, J. et al. Multi-dimensional characterisation of global food supply from 1961 to 2013. *Nat. Food* **1**, 70–75 (2020).
4. Prospective Studies Collaboration. Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. *Lancet* **370**, 1829–1839 (2007).
5. The Emerging Risk Factors Collaboration. Major lipids, apolipoproteins, and risk of vascular disease. *J. Am. Med. Assoc.* **302**, 1993–2000 (2009).
6. Global Burden of Metabolic Risk Factors for Chronic Diseases Collaboration. Cardiovascular disease, chronic kidney disease, and diabetes mortality burden of cardiometabolic risk factors from 1980 to 2010: a comparative risk assessment. *Lancet Diabetes Endocrinol.* **2**, 634–647 (2014).
7. Farzadfar, F. et al. National, regional, and global trends in serum total cholesterol since 1980: systematic analysis of health examination surveys and epidemiological studies with 321 country-years and 3·0 million participants. *Lancet* **377**, 578–586 (2011).
8. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. *Lancet* **390**, 2627–2642 (2017).
9. NCD Risk Factor Collaboration (NCD-RisC). Rising rural body-mass index is the main driver of the global obesity epidemic in adults. *Nature* **569**, 260–264 (2019).
10. Manthey, J. et al. Global alcohol exposure between 1990 and 2017 and forecasts until 2030: a modelling study. *Lancet* **393**, 2493–2502 (2019).
11. Micha, R. et al. Global, regional and national consumption of major food groups in 1990 and 2010: a systematic analysis including 266 country-specific nutrition surveys worldwide. *BMJ Open* **5**, e008705 (2015).
12. Micha, R. et al. Global, regional, and national consumption levels of dietary fats and oils in 1990 and 2010: a systematic analysis including 266 country-specific nutrition surveys. *Br. Med. J.* **348**, g2272 (2014).
13. Roth, G. A. et al. High total serum cholesterol, medication coverage and therapeutic control: an analysis of national health examination survey data from eight countries. *Bull. World Health Organ.* **89**, 92–101 (2011).
14. NCD Risk Factor Collaboration (NCD-RisC). National trends in total cholesterol obscure heterogeneous changes in HDL and non-HDL cholesterol and total-to-HDL cholesterol ratio: a pooled analysis of 458 population-based studies in Asian and Western countries. *Int. J. Epidemiol.* **49**, 173–192 (2020).
15. Carroll, M. D., Kit, B. K., Lacher, D. A., Shero, S. T. & Mussolini, M. E. Trends in lipids and lipoproteins in US adults, 1988–2010. *J. Am. Med. Assoc.* **308**, 1545–1554 (2012).
16. Walley, T., Folino-Gallo, P., Stephens, P. & Van Ganse, E. Trends in prescribing and utilization of statins and other lipid lowering drugs across Europe 1997–2003. *Br. J. Clin. Pharmacol.* **60**, 543–551 (2005).
17. Vartiainen, E., Laatikainen, T., Tapanainen, H. & Puska, P. Changes in serum cholesterol and diet in North Karelia and all Finland. *Glob. Heart* **11**, 179–184 (2016).
18. Miller, J. C. et al. Trends in serum total cholesterol and dietary fat intakes in New Zealand between 1989 and 2009. *Aust. N. Z. J. Public Health* **40**, 263–269 (2016).
19. Eriksson, M. et al. Greater decreases in cholesterol levels among individuals with high cardiovascular risk than among the general population: the northern Sweden MONICA study 1994 to 2014. *Eur. Heart J.* **37**, 1985–1992 (2016).
20. Arnett, D. K. et al. Twenty-year trends in serum cholesterol, hypercholesterolemia, and cholesterol medication use: the Minnesota Heart Survey, 1980–1982 to 2000–2002. *Circulation* **112**, 3884–3891 (2005).
21. Houterman, S., Verschuren, W. M., Oomen, C. M., Boersma-Cobbaert, C. M. & Kromhout, D. Trends in total and high density lipoprotein cholesterol and their determinants in The Netherlands between 1993 and 1997. *Int. J. Epidemiol.* **30**, 1063–1070 (2001).
22. Leth, T., Jensen, H. G., Mikkelsen, A. A. & Bysted, A. The effect of the regulation on trans fatty acid content in Danish food. *Atheroscler. Suppl.* **7**, 53–56 (2006).
23. Kypridemos, C. et al. Quantifying the contribution of statins to the decline in population mean cholesterol by socioeconomic group in England 1991–2012: a modelling study. *PLoS ONE* **10**, e0123112 (2015).
24. Ford, E. S. & Capewell, S. Trends in total and low-density lipoprotein cholesterol among U.S. adults: contributions of changes in dietary fat intake and use of cholesterol-lowering medications. *PLoS ONE* **8**, e65228 (2013).
25. He, Y. et al. The dietary transition and its association with cardiometabolic mortality among Chinese adults, 1982–2012: a cross-sectional population-based study. *Lancet Diabetes Endocrinol.* **7**, 540–548 (2019).
26. Wolmarans, P. Background paper on global trends in food production, intake and composition. *Ann. Nutr. Metab.* **55**, 244–272 (2009).
27. Zhang, M. et al. Prevalence of dyslipidemia and achievement of low-density lipoprotein cholesterol targets in Chinese adults: a nationally representative survey of 163,641 adults. *Int. J. Cardiol.* **260**, 196–203 (2018).
28. Song, Y. & Joung, H. A traditional Korean dietary pattern and metabolic syndrome abnormalities. *Nutr. Metab. Cardiovasc. Dis.* **22**, 456–462 (2012).
29. Yoshiike, N., Matsumura, Y., Iwaya, M., Sugiyama, M. & Yamaguchi, M. National Nutrition Survey in Japan. *J. Epidemiol.* **6**, S189–S200 (1996).
30. Matsumura, Y. Nutrition trends in Japan. *Asia Pac. J. Clin. Nutr.* **10**, S40–S47 (2001).
31. Kim, S., Moon, S. & Popkin, B. M. The nutrition transition in South Korea. *Am. J. Clin. Nutr.* **71**, 44–53 (2000).
32. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4·4 million participants. *Lancet* **387**, 1513–1530 (2016).
33. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with 19·1 million participants. *Lancet* **389**, 37–55 (2017).

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



**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/>.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

# Article

NCD Risk Factor Collaboration (NCD-RisC)

Cristina Taddei<sup>1</sup>, Bin Zhou<sup>1</sup>, Honor Bixby<sup>1</sup>, Rodrigo M. Carrillo-Larco<sup>1</sup>, Goodarz Danaei<sup>2</sup>, Rod T. Jackson<sup>3</sup>, Farshad Farzadfar<sup>4</sup>, Marisa K. Sophia<sup>1</sup>, Mariachiara Di Cesare<sup>5</sup>, Maria Laura Caminia Iurilli<sup>1</sup>, Andrea Rodriguez Martinez<sup>1</sup>, Golaleh Asghari<sup>6</sup>, Klodian Dhana<sup>7</sup>, Pablo Gulayin<sup>8</sup>, Sujay Kakarmath<sup>9</sup>, Marilina Santero<sup>8</sup>, Trudy Voortman<sup>10</sup>, Leanne M. Riley<sup>11</sup>, Melanie J. Cowan<sup>11</sup>, Stefan Savin<sup>11</sup>, James E. Bennett<sup>1</sup>, Gretchen A. Stevens<sup>12</sup>, Christopher J. Paciorek<sup>13</sup>, Wichai Aekplakorn<sup>14</sup>, Renata Cifkova<sup>15,16</sup>, Simona Giampaoli<sup>17</sup>, Andre Pascal Kengne<sup>18</sup>, Young-Ho Khang<sup>19</sup>, Kari Kuulasmaa<sup>20</sup>, Avula Laxmaia<sup>21</sup>, Paula Margozzi<sup>22</sup>, Prashant Mathur<sup>23</sup>, Børge G. Nordestgaard<sup>24</sup>, Dong Zhao<sup>25</sup>, Mette Aadahl<sup>26</sup>, Leandra Abarca-Gómez<sup>27</sup>, Hanan Abdul Rahim<sup>28</sup>, Niveen M. Abu-Rmeileh<sup>29</sup>, Benjamin Acosta-Cazares<sup>30</sup>, Robert J. Adams<sup>31</sup>, Imelda A. Agdeppa<sup>32</sup>, Javad Aghazadeh-Attari<sup>33</sup>, Carlos A. Aguilar-Salinas<sup>34</sup>, Charles Agyemang<sup>35</sup>, Tarunveer S. Ahluwalia<sup>36</sup>, Noor Ani Ahmad<sup>37</sup>, Ali Ahmadi<sup>38</sup>, Naser Ahmadi<sup>39</sup>, Soheir H. Ahmed<sup>39</sup>, Wolfgang Ahrens<sup>40</sup>, Kamel Ajlouni<sup>41</sup>, Monira Alarouj<sup>42</sup>, Fadia AlBuhairan<sup>43</sup>, Shahla AlDhukair<sup>44</sup>, Mohamed M. Ali<sup>11</sup>, Abdullah Alkandari<sup>42</sup>, Ala'a Alkerwi<sup>45</sup>, Eman Aly<sup>46</sup>, Deepak N. Amarapurkar<sup>47</sup>, Philippe Amouyel<sup>48,49</sup>, Lars Bo Andersen<sup>50</sup>, Sigmund A. Anderssen<sup>51</sup>, Ranjit Mohan Anjana<sup>52</sup>, Alireza Ansari-Moghaddam<sup>53</sup>, Hajera Aounallah-Skhiri<sup>54</sup>, Joana Araújo<sup>55</sup>, Inger Ariansen<sup>56</sup>, Tahir Aris<sup>37</sup>, Raphael E. Arku<sup>57</sup>, Nimmethatha Arlapappa<sup>21</sup>, Krishna K. Aryal<sup>58</sup>, Thor Aspelund<sup>59</sup>, Maria Cecilia F. Assunção<sup>60</sup>, Juha Avuinen<sup>61,62</sup>, Mária Ávadicová<sup>63</sup>, Ana Azevedo<sup>64</sup>, Freidoun Azizi<sup>65</sup>, Mehrdad Azmin<sup>4</sup>, Nagalla Balakrishna<sup>21</sup>, Mohamed Bamoshmoosh<sup>66</sup>, Maciej Banach<sup>67</sup>, Piotr Bandosz<sup>68</sup>, José R. Banegas<sup>69</sup>, Carlo M. Barbegalio<sup>70</sup>, Alberto Barceló<sup>71</sup>, Amina Barkat<sup>72</sup>, Iqbal Bata<sup>73</sup>, Anwar M. Batieha<sup>74</sup>, Assembekov Batyrbek<sup>75</sup>, Louise A. Baur<sup>76</sup>, Robert Beaglehole<sup>3</sup>, Antonisamy Belavendra<sup>77</sup>, Habibe Ben Romdhane<sup>78</sup>, Mikhail Benet<sup>79</sup>, Marianne Benn<sup>24</sup>, Salim Berkinbayev<sup>80</sup>, Antonio Bernabe-Ortiz<sup>91</sup>, Gaillete Bernotiene<sup>82</sup>, Heloise Bettio<sup>83</sup>, Santosh K. Bhargava<sup>84</sup>, Yufang Bi<sup>85</sup>, Asako Bieneck<sup>86</sup>, Mukharram Bikbov<sup>87</sup>, Bihungum Bista<sup>88</sup>, Peter Bjerregaard<sup>89</sup>, Espen Bjertness<sup>39</sup>, Marius B. Bjertness<sup>39</sup>, Cecilia Björkelund<sup>90</sup>, Katia V. Bloch<sup>91</sup>, Anneke Blokstra<sup>92</sup>, Simona Bo<sup>93</sup>, Bernhard O. Boehm<sup>94</sup>, Jose G. Boggia<sup>95</sup>, Carlos P. Boissonnet<sup>96</sup>, Marialaura Bonaccio<sup>97</sup>, Vanina Bongard<sup>98</sup>, Rossana Borchini<sup>99</sup>, Herman Borghs<sup>100</sup>, Pascal Bovet<sup>101,102</sup>, Imperia Brajkovich<sup>103</sup>, Juergen Breckenkamp<sup>104</sup>, Hermann Brenner<sup>105</sup>, Lizzy M. Brewster<sup>35</sup>, Graziella Bruno<sup>93</sup>, Anna Bugge<sup>106</sup>, Markus A. Busch<sup>107</sup>, Antonio Cabrera de León<sup>108</sup>, Joseph Cacciottolo<sup>109</sup>, Güney Can<sup>110</sup>, Ana Paula C. Cândido<sup>111</sup>, Mario V. Capanzana<sup>32</sup>, Eduardo Capuano<sup>112</sup>, Vincenzo Capuano<sup>112</sup>, Viviane C. Cardoso<sup>83</sup>, Joana Carvalho<sup>113</sup>, Felipe F. Casanueva<sup>114</sup>, Laura Censi<sup>115</sup>, Charalampos A. Chadjigeorgiou<sup>116</sup>, Snehalatha Chamukuttan<sup>117</sup>, Nish Chaturvedi<sup>118</sup>, Chien-Jen Chen<sup>119</sup>, Fangfang Chen<sup>120</sup>, Shuohua Chen<sup>121</sup>, Ching-Yu Cheng<sup>122</sup>, Bahman Cheraghian<sup>123</sup>, Angela Chetrit<sup>124</sup>, Shu-Ti Chiou<sup>125</sup>, María-Dolores Chirlaque<sup>126</sup>, Belong Cho<sup>127</sup>, Yumi Cho<sup>128</sup>, Jerzy Chudek<sup>129</sup>, Frank ClaesSENS<sup>130</sup>, Janine Clarke<sup>131</sup>, Els Clays<sup>132</sup>, Hans Concin<sup>133</sup>, Susana C. Confortin<sup>134</sup>, Cyrus Cooper<sup>125</sup>, Simona Costanzo<sup>97</sup>, Dominique Cottel<sup>126</sup>, Chris Cowell<sup>76</sup>, Ana B. Crujeiras<sup>137</sup>, Semánová Csilla<sup>138</sup>, Liufu Cui<sup>121</sup>, Felipe V. Cureau<sup>139</sup>, Graziella D'Arrigo<sup>140</sup>, Eleonora d'Orsi<sup>141</sup>, Jean Dallongeville<sup>136</sup>, Albertino Damasceno<sup>142</sup>, Rachel Dankner<sup>124</sup>, Thomas M. Dantonf<sup>26</sup>, Luc Dauchet<sup>48,49</sup>, Kairat Davletov<sup>76</sup>, Guy De Backer<sup>132</sup>, Dirk De Bacquer<sup>132</sup>, Giovanni de Gaetano<sup>97</sup>, Stefaan De Henauw<sup>132</sup>, Paula Duarte de Oliveira<sup>60</sup>, David De Ridder<sup>143</sup>, Delphine De Smedt<sup>132</sup>, Mohan Deepa<sup>52</sup>, Alexander D. Deev<sup>144</sup>, Abbas Dehghan<sup>1</sup>, Hélène Delisle<sup>145</sup>, Elaine Dennison<sup>135</sup>, Valérie Deschamps<sup>146</sup>, Meghnath Dhimal<sup>88</sup>, Augusto F. Di Castelnuovo<sup>147</sup>, Zivka Dika<sup>148</sup>, Shirin Jalaljina<sup>149</sup>, Annette J. Dobson<sup>150</sup>, Chiara Donfrancesco<sup>17</sup>, Silvana P. Donoso<sup>151</sup>, Angela Döring<sup>152</sup>, Maria Dorobantu<sup>153</sup>, Nico Dragano<sup>154</sup>, Wojciech Drygas<sup>67,155</sup>, Yong Du<sup>107</sup>, Charmaine A. Duante<sup>32</sup>, Rosemary B. Duda<sup>156</sup>, Vilnis Dzerve<sup>157</sup>, Elzbieta Dzianowska-Zaborszczyk<sup>67</sup>, Ricky Eddie<sup>158</sup>, Ebrahim Eftekhari<sup>159</sup>, Robert Eggertsen<sup>90</sup>, Sarah Eghtesad<sup>4</sup>, Gabriele Eiben<sup>160</sup>, Ulf Ekelund<sup>51</sup>, Jalila El Atti<sup>161</sup>, Denise Eldemire-Shearer<sup>162</sup>, Marie Eliassen<sup>26</sup>, Roberto Elosua<sup>163</sup>, Rajiv T. Erasmus<sup>164</sup>, Raimund Erbel<sup>165</sup>, Cihangir Erem<sup>166</sup>, Louise Eriksen<sup>89</sup>, Johan G. Eriksson<sup>167</sup>, Jorge Escobedo-de la Peña<sup>30</sup>, Saeid Eslami<sup>168</sup>, Ali Esmaeil<sup>169</sup>, Alun Evans<sup>170</sup>, David Faeh<sup>171</sup>, Caroline H. Fall<sup>126</sup>, Elnaz Faramarzi<sup>172</sup>, Mojtaba Farjam<sup>173</sup>, Mohammad Reza Fattah<sup>174</sup>, Francisco J. Felix-Redondo<sup>175</sup>, Trevor S. Ferguson<sup>162</sup>, Daniel Fernández-Bergés<sup>177</sup>, Daniel Ferrante<sup>177</sup>, Marika Ferrari<sup>118</sup>, Catterina Ferreccio<sup>22</sup>, Jean Ferrieres<sup>98</sup>, Bernhard Föger<sup>133</sup>, Leng Huat Foo<sup>178</sup>, Ann-Sofie Forslund<sup>179</sup>, Maria Forsner<sup>179</sup>, Heba M. Fouad<sup>46</sup>, Damian K. Francis<sup>162</sup>, Maria do Carmo Franco<sup>180</sup>, Oscar H. Franco<sup>10</sup>, Guillermo Frontera<sup>181</sup>, Yuki Fujita<sup>182</sup>, Matsuda Fumihiko<sup>183</sup>, Takuro Furusawa<sup>183</sup>, Zbigniew Gacioglu<sup>184</sup>, Fabio Galvano<sup>185</sup>, Jingli Gao<sup>121</sup>, Manoli Garcia-de-la-Hera<sup>186</sup>, Sarah P. Garnett<sup>76</sup>, Jean-Michel Gaspoz<sup>143</sup>, Magda Gasull<sup>187</sup>, Andrea Gazzinelli<sup>188</sup>, Johanna M. Geleijnse<sup>189</sup>, Ali Ghanbari<sup>4</sup>, Erfan Ghasemi<sup>14</sup>, Oana-Florentina Gheorghe-Fronea<sup>153</sup>, Anup Ghimire<sup>190</sup>, Francesco Gianfagna<sup>147,191</sup>, Tiffany K. Gill<sup>192</sup>, Jonathan Giovannelli<sup>48,49</sup>, Glen Gironella<sup>32</sup>, Aleksander Giwercman<sup>193</sup>, David Goltzman<sup>194</sup>, Helen Gonçalves<sup>60</sup>, David A. Gonzalez-Chica<sup>192</sup>, Marcela Gonzalez-Gross<sup>195</sup>, Juan P. González-Rivas<sup>196</sup>, Clicerio González-Villalpando<sup>197</sup>, María-Elena González-Villalpando<sup>198</sup>, Angel R. Gonzalez<sup>199</sup>, Frederic Gottrand<sup>48</sup>, Sidsel Graff-Iversen<sup>56</sup>, Dušan Gralfnetter<sup>200</sup>, Ronald D. Gregor<sup>7</sup>, Tomasz Grodzicki<sup>201</sup>, Anders Grøntved<sup>202</sup>, Giuseppe Grossi<sup>185</sup>, Gabriella Gruden<sup>93</sup>, Dongfeng Gu<sup>203</sup>, Pilar Guallar-Castillón<sup>6</sup>, Ong Peng Guan<sup>200</sup>, Elias F. Gudmundsson<sup>205</sup>, Vilmundur Gudnason<sup>59</sup>, Ramiro Guerrero<sup>206</sup>, Idris Guessous<sup>143</sup>, Johannina Gunnlaugsdottir<sup>205</sup>, Rajeev Gupta<sup>207</sup>, Laura Gutierrez<sup>8</sup>, Felix Gutzwiller<sup>171</sup>, Seongjun Ha<sup>208</sup>, Farzad Hadaegh<sup>209</sup>, Rosa Haghshenas<sup>4</sup>, Hamid Hakimi<sup>169</sup>, Ian R. Hambleton<sup>210</sup>, Behrooz Hamzeh<sup>211</sup>, Sari Hantunen<sup>212</sup>, Rachakulla Hari Kumar<sup>21</sup>, Seyed Mohammad Hashemi-Shahri<sup>53</sup>, Jun Hata<sup>213</sup>, Teresa Haugsgjerd<sup>214</sup>, Alison J. Hayes<sup>76</sup>, Jiang He<sup>215</sup>, Yuna He<sup>216</sup>, Marleen Elisabeth Hendriks<sup>217</sup>, Ana Henriques<sup>55</sup>, Sauli Herrala<sup>62</sup>, Ramin Heshmat<sup>218</sup>, Allan G. Hill<sup>135</sup>, Sai Yin Ho<sup>219</sup>, Suzanne C. Ho<sup>220</sup>, Michael Hobbs<sup>221</sup>, Albert Hofman<sup>10</sup>, Reza Homayounfar<sup>173</sup>, Wilma M. Hopman<sup>222</sup>, Andrea R. V. R. Horimoto<sup>223</sup>, Claudia M. Hormiga<sup>224</sup>, Bernardo L. Horta<sup>60</sup>, Leila Hout<sup>225</sup>, Christina Howitt<sup>210</sup>, Thein Thein Htay<sup>226</sup>, Aung Soe Htet<sup>227</sup>, Maung Maung Than Htike<sup>227</sup>, José María Huerta<sup>228</sup>, Ilpo Tapani Huhtaniemi<sup>1</sup>, Martijn Huisman<sup>229</sup>, Monica L. Hunsberger<sup>90</sup>, Abdullatif S. Hussein<sup>29</sup>, Inge Huybrechts<sup>230</sup>, Nahla Hwalla<sup>231</sup>, Licia Iacoviello<sup>97,101</sup>, Anna G. Iannone<sup>112</sup>, Mohsen M. Ibrahim<sup>232</sup>, Norazizah Ibrahim Wong<sup>37</sup>, Iris Iglesia<sup>233</sup>, Nayu Ikeda<sup>234</sup>, M. Arfan Ikram<sup>10</sup>, Violeta Iotova<sup>235</sup>, Vilma E. Irazola<sup>8</sup>, Takafumi Ishida<sup>236</sup>, Muhammad Islam<sup>237</sup>, Aziz al-Safi Ismail<sup>178</sup>, Masanori Iwasaki<sup>238</sup>, Jeremy M. Jacobs<sup>239</sup>, Hashem Y. Jaddou<sup>74</sup>, Tazeen Jafar<sup>122</sup>, Kenneth James<sup>162</sup>, Konrad Jamrozik<sup>192,448</sup>, Imre Janszky<sup>240</sup>, Edward Janus<sup>241</sup>, Mario-Riitta Jarvelin<sup>1,61,62</sup>, Grazyna Jasinska<sup>201</sup>, Ana Jelakovic<sup>242</sup>, Bojan Jelakovic<sup>243</sup>, Garry Jennings<sup>244</sup>, Gorm B. Jensen<sup>24</sup>, Seung-yeal Jeong<sup>208</sup>, Anjani Kumar Jha<sup>88</sup>, Chao Qiang Jiang<sup>245</sup>, Ramon O. Jimenez<sup>246</sup>, Karl-Heinz Jöckel<sup>165</sup>, Michel Joffres<sup>247</sup>, Jari J. Jokelainen<sup>62</sup>, Jost B. Jonas<sup>248</sup>, Torben Jørgensen<sup>26</sup>, Pradeep Joshi<sup>249</sup>, Farahnaz Joukar<sup>250</sup>, Jacek Józwik<sup>251</sup>, Anne Juolevi<sup>20</sup>, Anthony Kafatos<sup>552</sup>, Eero O. Kajaste<sup>20</sup>, Ofra Kalter-Leibovici<sup>124</sup>, Nor Azmi Kamaruddin<sup>253</sup>, Piia R. Kamstrup<sup>24</sup>, Khem B. Karki<sup>254</sup>, Joanne Katz<sup>255</sup>, Jussi Kauhanen<sup>212</sup>, Prabhdeep Kaur<sup>256</sup>, Maryam Kavousi<sup>10</sup>, Gyulli Kazakbaeva<sup>87</sup>, Ulrich Keil<sup>257</sup>, Sirkka Keinänen-Kiuakaanniemi<sup>62</sup>, Roya Kelishadi<sup>258</sup>, Maryam Keramati<sup>168</sup>, Alina Kerimkulova<sup>259</sup>, Mathilde Kersting<sup>260</sup>, Yousef Saleh Khader<sup>74</sup>, Davood Khalili<sup>6</sup>, Mohammad Khateeb<sup>41</sup>, Motahareh Kheradmand<sup>661</sup>, Alireza Khosravi<sup>262</sup>, Ursula Kiechl-Kohlendorfer<sup>263</sup>, Stefan Kiechl<sup>263</sup>, Japhet Killewo<sup>264</sup>, Hyeon Chang Kim<sup>266</sup>, Jeongseon Kim<sup>266</sup>, Yeon-Yong Kim<sup>208</sup>, Jurate Klumbiene<sup>62</sup>, Michael Knoflach<sup>263</sup>, Stephanie Ko<sup>66</sup>, Hans-Peter Kohler<sup>267</sup>, Iliana V. Kohler<sup>267</sup>, Elin Kolle<sup>51</sup>, Patrick Kolsteren<sup>132</sup>, Jürgen König<sup>268</sup>, Raija Korpelainen<sup>61,269</sup>, Paul Korrovits<sup>270</sup>, Jelena Kos<sup>242</sup>, Seppo Koskinen<sup>20</sup>, Katsuyasu Kouda<sup>271</sup>, Sudhir Kowlessur<sup>272</sup>, Wolfgang Kratzer<sup>273</sup>, Susi Kriemler<sup>171</sup>, Peter Lund Kristensen<sup>202</sup>, Steiner Krokstad<sup>240</sup>, Daan Kromhout<sup>274</sup>, Urho M. Kujala<sup>275</sup>, Paweł Kurjata<sup>185</sup>, Catherine Kyobutungi<sup>276</sup>, Fatima Zahra Laamiri<sup>277</sup>, Tiina Laatikainen<sup>20</sup>, Carl Lachat<sup>132</sup>, Youcef Laid<sup>278</sup>, Tai Hing Lam<sup>219</sup>, Christina-Paulina Lambrinou<sup>279</sup>, Vera Laska<sup>200</sup>, Georg Lappas<sup>280</sup>, Bagher Larjani<sup>281</sup>, Tint Swe Latt<sup>282</sup>, Lars E. Laugsand<sup>240</sup>, Maria Lazo-Porrás<sup>81</sup>, Jeannette Lee<sup>283</sup>, Jeonghee Lee<sup>266</sup>, Nils Lehmann<sup>165</sup>, Terho Lehtimäki<sup>284,285</sup>, Naomi S. Levitt<sup>286</sup>, Yanping Li<sup>2</sup>, Christa L. Lilly<sup>287</sup>, Wei-Yen Lim<sup>283</sup>, M. Fernanda Lima-Costa<sup>288</sup>, Hsien-Ho Lin<sup>289</sup>, Xu Lin<sup>290</sup>, Yi-Ting Lin<sup>291</sup>, Lars Lind<sup>291</sup>, Allan Linneberg<sup>26</sup>, Lauren Lissner<sup>90</sup>, Jing Liu<sup>25</sup>, Helle-Mai Loit<sup>292</sup>, Esther Lopez-Garcia<sup>69</sup>, Tania Lopez<sup>293</sup>, Paulo A. Lotufo<sup>63</sup>, José Eugenio Lozano<sup>294</sup>, Dalia Luksiene<sup>92</sup>, Annamari Lundquist<sup>20</sup>, Robert Lundqvist<sup>295</sup>, Nuno Lunet<sup>112</sup>, Guansheng Ma<sup>296</sup>, George L. L. Machado-Coelho<sup>297</sup>, Aristides M. Machado-Rodrigues<sup>298</sup>, Suka Machi<sup>299</sup>, Ahmed A. Madar<sup>39</sup>, Stefania Maggi<sup>300</sup>, Dianna J. Magliano<sup>301</sup>, Emmanuela Magripilis<sup>302</sup>, Gowri Mahasampath<sup>77</sup>, Bernard Maire<sup>303</sup>, Marcia Makdisse<sup>304</sup>, Fatemeh Malekzadeh<sup>174</sup>, Reza Malekzadeh<sup>4</sup>, Kadavanti Mallikharjuna Rao<sup>21</sup>, Yannis Manios<sup>279</sup>, Jim I. Mann<sup>305</sup>, Fariborz Mansour-Ghanaei<sup>290</sup>, Enzo Manzato<sup>306</sup>, Pedro Marques-Vidal<sup>307</sup>, Reynaldo Martorell<sup>309</sup>, Luis P. Mascarenhas<sup>309</sup>, Ellisiv B. Mathiesen<sup>311</sup>, Tandi E. Matscha<sup>311</sup>, Christina Mavrogiani<sup>279</sup>, Shelly R. McFarlane<sup>162</sup>, Stephen T. McGarvey<sup>312</sup>, Stela McLachlan<sup>313</sup>, Rachael M. McLean<sup>305</sup>, Scott B. McLean<sup>131</sup>, Breige A. McNulty<sup>314</sup>, Sounnija Medina-Bencherif<sup>225</sup>, Parinaz Mehdipour<sup>4</sup>, Kirsten Mehlig<sup>90</sup>, Amir Houshang Mehrparvar<sup>315</sup>, Aline Meirhaeghe<sup>316</sup>, Christa Meisinger<sup>152</sup>, Ana María B. Menezes<sup>60</sup>, Geetha R. Menon<sup>317</sup>, Shahin Merat<sup>4</sup>, Alibek Mereke<sup>75</sup>, Indrapal I. Meshram<sup>21</sup>, Patricia Metcalf<sup>3</sup>, Haakon E. Meyer<sup>39</sup>, Jie Mi<sup>120</sup>, Nathalie Michels<sup>132</sup>, Jody C. Miller<sup>305</sup>, Cláudia S. Minderico<sup>318</sup>, G. K. Mini<sup>319</sup>, Juan Francisco Miquel<sup>22</sup>, J. Jaime Miranda<sup>81</sup>, Mohammad Reza Mirjalili<sup>315</sup>, Erkin Mirrakhimov<sup>299</sup>, Pietro A. Modesti<sup>320</sup>, Sahar Saeedi Moghaddam<sup>4</sup>, Bahram Mohajer<sup>4</sup>, Mostafa K. Mohamed<sup>331</sup>, Kazem Mohammad<sup>4</sup>, Zahra Mohammad<sup>4</sup>, Noushin Mohammadifard<sup>322</sup>, Reza Mohammadpourhodki<sup>168</sup>, Viswanathan Mohan<sup>52</sup>, Salim Mohanna<sup>81</sup>, Muhammad Fahdli Mohd Yusoff<sup>37</sup>, Iraj Mohebbi<sup>33</sup>, Farnam Mohebbi<sup>4</sup>, Marie Moity<sup>323,324</sup>, Line T. Møllehave<sup>26</sup>, Niels C. Moller<sup>202</sup>, Dénes Molnár<sup>325</sup>, Amirabbas Momenan<sup>6</sup>, Charles K. Mondo<sup>326</sup>, Eric Monterrubio-Flores<sup>197</sup>, Mahmood Moosazadeh<sup>261</sup>, Alain Morejon<sup>327</sup>, Luis A. Moreno<sup>233</sup>, Karen Morgan<sup>328</sup>, Suzanne N. Morin<sup>194</sup>, George Moschonis<sup>329</sup>, Małgorzata Mossakowska<sup>330</sup>, Aya Mostafa<sup>321</sup>, Jorge Mota<sup>133</sup>, Mohammad Esmaeel Motlagh<sup>123</sup>, Jorge Motta<sup>331</sup>, Kelias P. Msyamboza<sup>332</sup>, Maria L. Muiesan<sup>333</sup>, Martina Müller-Nerasyid<sup>162</sup>, Jaakko Mursu<sup>212</sup>, Norlaila Mustafa<sup>253</sup>, Iraj Nabipour<sup>334</sup>, Shohreh Naderimaghram<sup>4</sup>, Gabriele Nagel<sup>335</sup>, Balkish M. Naidu<sup>37</sup>, Farid Najafi<sup>211</sup>, Harunobu Nakamura<sup>336</sup>, Jana Námešná<sup>63</sup>, El Ei K. Nang<sup>283</sup>, Vinay B. Nangia<sup>337</sup>, Matthias Nauck<sup>338</sup>, William A. Neal<sup>287</sup>, Azim Nejatizadeh<sup>159</sup>, Ilona Nenko<sup>201</sup>, Flavio Nervi<sup>22</sup>, Nguyen D. Nguyen<sup>339</sup>, Quang Ngoc Nguyen<sup>340</sup>, Ramfis E. Nieto-Martinez<sup>341</sup>, Thomas Nihal<sup>17</sup>, Teemu J. Niiranen<sup>20,342</sup>, Guang Ning<sup>85</sup>, Toshiharu Ninomiya<sup>213</sup>, Marianna Noale<sup>300</sup>, Oscar A. Noboa<sup>95</sup>, Davide Noto<sup>70</sup>, Mohannad Al Nsour<sup>343</sup>, Irfan Nuhoglu<sup>166</sup>, Terence W. O'Neill<sup>344</sup>, Dermot O'Reilly<sup>170</sup>, Angélica M. Ochoa-Avilés<sup>161</sup>, Kyungwon Oh<sup>128</sup>, Ryutaro Ohtsuka<sup>345</sup>, Örn Olafsson<sup>205</sup>, Valérie Olie<sup>146</sup>, Isabel O. Oliveira<sup>60</sup>, Mohd Azahadi Omar<sup>37</sup>, Altan Onat<sup>246,448</sup>, Sok King Ong<sup>347</sup>, Pedro Ordunez<sup>71</sup>, Rui Ornelas<sup>348</sup>, Pedro J. Ortiz<sup>81</sup>, Clive Osmond<sup>349</sup>, Sergej M. Ostojic<sup>350</sup>, Afshin Ostovar<sup>4</sup>, Johanna A. Otero<sup>224</sup>, Ellis Owusu-Dabo<sup>351</sup>, Fred Michel Paccaud<sup>352</sup>, Elena Pahomova<sup>157</sup>, Andrzej Pajak<sup>201</sup>, Luigi Palmieri<sup>17</sup>, Wen-Harn Pan<sup>19</sup>, Songhomitra Panda-Jonas<sup>248</sup>, Francesco Panza<sup>353</sup>, Winsome R. Parnell<sup>105</sup>, Nikhil D. Patel<sup>354</sup>, Nasheeta Peer<sup>355</sup>, Sergio Viana Peixoto<sup>288</sup>, Markku Peltonen<sup>20</sup>, Alexandre C. Pereira<sup>223</sup>, Annette Peters<sup>352</sup>, Astrid Petersmann<sup>333</sup>, Janina Petkovicie<sup>82</sup>, Niloofer Peykar<sup>149</sup>, Son Thai Pham<sup>356</sup>, Rafael N. Pichardo<sup>357</sup>, Iris Pigeot<sup>358</sup>, Aida Pilav<sup>359</sup>, Lorenza Pilotto<sup>360</sup>, Aleksandra Piwonska<sup>155</sup>, Andreia N. Pizarro<sup>113</sup>, Pedro Plans-Rubio<sup>361</sup>, Silvia Plata<sup>362</sup>, Hermann Pohlbeln<sup>358</sup>, Miquel Porta<sup>163</sup>, Marileen L. P. Portegies<sup>10</sup>, Annal Pouly<sup>188</sup>, Farhad Pourfarzi<sup>363</sup>, Hossein Poustchi<sup>1</sup>, Rajendra Pradeepa<sup>62</sup>, Jacqueline F. Price<sup>315</sup>, Rui Providencia<sup>118</sup>, Jardena J. Puder<sup>307</sup>, Soile E. Puhakka<sup>61,269</sup>, Margus Punab<sup>270</sup>, Mostafa Qorbani<sup>364</sup>, Tran Quoc Bao<sup>365</sup>, Ricardas Radisauskas<sup>82</sup>, Salar Rahimkazerooni<sup>174</sup>, Olli Raitakari<sup>342</sup>, Sudha Ramachandra Rao<sup>256</sup>, Ambady Ramachandran<sup>366</sup>, Elisabette Ramos<sup>64</sup>, Rafael Ramos<sup>367</sup>, Lekhraj Rampal<sup>368</sup>, Sanjay Rampal<sup>369</sup>, Josep Redon<sup>370</sup>, Paul Ferdinand M. Regan<sup>371</sup>, Luis Revilla<sup>293</sup>, Abbas Rezaianzadeh<sup>174</sup>, Robespierre Ribeiro<sup>372,448</sup>, Adrian Richter<sup>338</sup>, Fernando Rigo<sup>373</sup>, Tobias F. Rinke de Wit<sup>374</sup>, Fernando Rodríguez-Artalejo<sup>61</sup>, María del Cristo Rodríguez-Perez<sup>375</sup>, Laura A. Rodríguez-Villamizar<sup>376</sup>, Ulla Roggenbuck<sup>165</sup>, Rosalba Rojas-Martínez<sup>197</sup>, Dora Romaguera<sup>137</sup>, Elisabetta L. Romeo<sup>377</sup>, Annika Rosengren<sup>90,378</sup>, Joel G. R. Roy<sup>151</sup>, Adolfo Rubinstein<sup>9</sup>.

- Jean-Bernard Ruidavets<sup>379</sup>, Blanca Sandra Ruiz-Betancourt<sup>30</sup>, Paola Russo<sup>380</sup>, Petra Rust<sup>268</sup>, Marcin Rutkowski<sup>68</sup>, Charumathi Sabanayagam<sup>204</sup>, Harshpal S. Sachdev<sup>301</sup>, Aireza Sadjadi<sup>4</sup>, Ali Reza Safarpour<sup>174</sup>, Saeid Safiri<sup>172</sup>, Olfa Saidi<sup>382</sup>, Nader Saki<sup>123</sup>, Benoit Salanave<sup>146</sup>, Diego Salmerón<sup>228</sup>, Veikko Salomaa<sup>20</sup>, Jukka T. Salonen<sup>167</sup>, Massimo Salvetti<sup>333</sup>, Jose Sánchez-Abanto<sup>383</sup>, Susana Sans<sup>384</sup>, Alba M. Santaliestra-Pasias<sup>235</sup>, Diana A. Santos<sup>385</sup>, María Paula Santos<sup>113</sup>, Rute Santos<sup>113</sup>, Jouko L. Saramies<sup>386</sup>, Luis B. Sardinha<sup>385</sup>, Nizal Sarrafzadegan<sup>387</sup>, Kai-Uwe Saum<sup>105</sup>, Savvas C. Savva<sup>116</sup>, Norie Sawada<sup>388</sup>, Mariana Sbaraini<sup>139</sup>, Marcia Scauzufca<sup>389</sup>, Beatriz D. Schaan<sup>139</sup>, Herman Schargrodsky<sup>390</sup>, Christa Scheidt-Nave<sup>107</sup>, Anja Schienkiewitz<sup>107</sup>, Sabine Schip<sup>338</sup>, Carsten O. Schmidt<sup>338</sup>, Ben Schöttker<sup>105</sup>, Sara Schramm<sup>165</sup>, Sylvain Sebert<sup>61</sup>, Aye Aye Sein<sup>227</sup>, Abhijit Sen<sup>391</sup>, Sadaf G. Sepanlou<sup>4</sup>, Jennifer Servais<sup>31</sup>, Ramin Shakeri<sup>4</sup>, Svetlana A. Shalnava<sup>144</sup>, Teresa Shamah-Levy<sup>197</sup>, Maryam Sharafkhah<sup>4</sup>, Sanjib K. Sharma<sup>190</sup>, Jonathan E. Shaw<sup>301</sup>, Amaneh Shayanrad<sup>4</sup>, Zumin Shi<sup>28</sup>, Kenji Shibuya<sup>392</sup>, Hana Shimizu-Furusawa<sup>393</sup>, Dong Wook Shin<sup>304</sup>, Youchan Shin<sup>204</sup>, Majid Shirani<sup>38</sup>, Rahman Shiri<sup>395</sup>, Namuna Shrestha<sup>88</sup>, Khairil Si-Ramlee<sup>347</sup>, Alfonso Siani<sup>380</sup>, Rosalynn Siantar<sup>204</sup>, Abla M. Siba<sup>231</sup>, Diego Augusto Santos Silva<sup>141</sup>, Mary Simon<sup>366</sup>, Judith Simons<sup>396</sup>, Leon A. Simons<sup>397</sup>, Michael Sjöström<sup>398</sup>, Tea Skaaby<sup>399</sup>, Jolanta Slowikowska-Hilczer<sup>67</sup>, Przemyslaw Slusarczyk<sup>330</sup>, Liam Smeeth<sup>400</sup>, Marieke B. Snijder<sup>35</sup>, Stefan Söderberg<sup>179</sup>, Agustinus Soemerdit<sup>401</sup>, Reecha Sofat<sup>18</sup>, Vincenzo Solfrizzi<sup>402</sup>, Mohammad Hossein Somi<sup>172</sup>, Emily Sonestedt<sup>193</sup>, Thorkild I. A. Sørensen<sup>403</sup>, Charles Sossa Jérôme<sup>404</sup>, Aïcha Soumaré<sup>405</sup>, Kaan Sozmen<sup>406</sup>, Karen Sparrenberger<sup>139</sup>, Jan A. Staessen<sup>407</sup>, Maria G. Stathopoulou<sup>408</sup>, Bill Stavreski<sup>244</sup>, Jostein Steene-Johannessen<sup>51</sup>, Peter Stehle<sup>409</sup>, Aryeh D. Stein<sup>308</sup>, Jochanan Steissman<sup>339</sup>, Ranko Stevanović<sup>410</sup>, Jutta Stieber<sup>152,448</sup>, Doris Stöckl<sup>152</sup>, Jakub Stokwizzewski<sup>411</sup>, Karien Stronks<sup>35</sup>, Maria Wany Strufaldi<sup>180</sup>, Ramón Suárez-Medina<sup>412</sup>, Chien-An Sun<sup>413</sup>, Johan Sundström<sup>291</sup>, Paibul Suriyawongpaisal<sup>141</sup>, Rody G. Sy<sup>371</sup>, René Charles Sylva<sup>144</sup>, Moyses Szko<sup>355</sup>, E. Shyong Tai<sup>283</sup>, Abdonas Tamosiunas<sup>82</sup>, Eng Joo Tan<sup>76</sup>, Mohammed Rasoul Tarawneh<sup>415</sup>, Carolina B. Tarqui-Mamani<sup>383</sup>, Anne Taylor<sup>192</sup>, Julie Taylor<sup>118</sup>, Grethe S. Tell<sup>214</sup>, Tania Tello<sup>81</sup>, K. R. Thankappan<sup>416</sup>, Lutgarde Thijs<sup>407</sup>, Betina H. Thuesen<sup>26</sup>, Ulla Toft<sup>26</sup>, Hanna K. Tolonen<sup>20</sup>, Janne S. Tolstrup<sup>89</sup>, Murat Topbas<sup>166</sup>, Roman Topór-Madry<sup>201</sup>, María José Tormo<sup>417</sup>, Michael J. Tornaritis<sup>16</sup>, Maties Torrent<sup>418</sup>, Laura Torres-Collado<sup>106</sup>, Pierre Traissac<sup>303</sup>, Oanh T. H. Trinh<sup>339</sup>, Julia Truthmann<sup>107</sup>, Shoichiro Tsugane<sup>388</sup>, Marshall K. Tulloch-Reid<sup>162</sup>, Tomi-Pekka Tuomainen<sup>212</sup>, Jaakko Tuomilehto<sup>20</sup>, Anne Tybjaerg-Hansen<sup>24</sup>, Christophe Tzourio<sup>405</sup>, Peter Ueda<sup>398</sup>, Eunice Ugel<sup>419</sup>, Hanno Ulmer<sup>263</sup>, Belgin Una<sup>420</sup>, Hannu M. T. Uusitalo<sup>421</sup>, Gonzalo Valdivia<sup>22</sup>, Damaskini Valvi<sup>422</sup>, Rob M. van Dam<sup>283</sup>, Yvonne T. van der Schouw<sup>423</sup>, Koen Van Herck<sup>132</sup>, Hoang Van Minh<sup>424</sup>, Lenie van Rossem<sup>425</sup>, Natasja M. Van Schoor<sup>229</sup>, Irene G. M. van Valkengoed<sup>35</sup>, Dirk Vanderschueren<sup>150</sup>, Diego Vanuzzo<sup>360</sup>, Anette Varbo<sup>24</sup>, Patricia Varona-Pérez<sup>412</sup>, Senthil K. Vasan<sup>135</sup>, Lars Vatten<sup>240</sup>, Tomas Vega<sup>294</sup>, Toomas Veidebaum<sup>292</sup>, Gustavo Velasquez-Melendez<sup>188</sup>, Silvia J. Venero-Fernández<sup>412</sup>, Giovanni Veronesi<sup>191</sup>, W. M. Monique Verschuren<sup>92</sup>, Cesar G. Victoria<sup>50</sup>, Dhanasari Vidiawati<sup>426</sup>, Lucie Viet<sup>92</sup>, Salvador Villalpando<sup>197</sup>, Jesus Vioque<sup>427</sup>, Jyrki K. Virtanen<sup>212</sup>, Sophie Visvikis-Siest<sup>408</sup>, Bharathi Viswanathan<sup>101</sup>, Tiina Vlasoff<sup>428</sup>, Peter Vollenweider<sup>307</sup>, Ari Voutilainen<sup>212</sup>, Alisha N. Wade<sup>429</sup>, Aline Wagner<sup>323</sup>, Janette Walton<sup>430</sup>, Wan Mohamad Wan Bebakar<sup>178</sup>, Wan Nazaimoon Wan Mohamad<sup>431</sup>, Ming-Dong Wang<sup>432</sup>, Ningli Wang<sup>433</sup>, Qian Wang<sup>434</sup>, Ya Xing Wang<sup>435</sup>, Ying-Wei Wang<sup>125</sup>, S. Goya Wannamethee<sup>118</sup>, Niels Wedderkopp<sup>202</sup>, Wenbin Wei<sup>435</sup>, Peter H. Whinck<sup>436</sup>, Kurt Widhalm<sup>437</sup>, Indah S. Widhyahening<sup>426</sup>, Andrzej Wiecek<sup>125</sup>, Alet H. Wijga<sup>92</sup>, Rainford J. Wilks<sup>162</sup>, Johann Willert<sup>63</sup>, Peter Willert<sup>263</sup>, Tom Wilsgaard<sup>310</sup>, Bogdan Wojtyniak<sup>411</sup>, Roy A. Wong-McClure<sup>27</sup>, Andrew Wong<sup>118</sup>, Tien Yin Wong<sup>122</sup>, Jean Woo<sup>220</sup>, Mark Woodward<sup>397,438</sup>, Frederick C. Wu<sup>344</sup>, Shouling Wu<sup>121</sup>, Haiquan Xu<sup>439</sup>, Liang Xu<sup>433</sup>, Weili Yan<sup>440</sup>, Xiaoguang Yang<sup>21</sup>, Tabara Yasuharu<sup>183</sup>, Xingwang Ye<sup>290</sup>, Teh Peng Yeow<sup>44</sup>, Panayiotis K. Yiallourou<sup>442</sup>, Moein Yoosefi<sup>4</sup>, Akhiro Yoshihara<sup>238</sup>, San-Lin Yu<sup>413</sup>, Novie O. Younger-Coleman<sup>162</sup>, Ahmad Faizudi Yusof<sup>37</sup>, Ahmad A. Zainuddin<sup>37</sup>, Seyed Rasoul Zakavi<sup>168</sup>, Mohammad Reza Zali<sup>6</sup>, Farhad Zalani<sup>443</sup>, Sabina Zambon<sup>300</sup>, Antonis Zampelas<sup>302</sup>, Ko Ko Zaw<sup>282</sup>, Tomasz Zdrojewski<sup>68</sup>, Tajana Zeljkovic Vrkic<sup>442</sup>, Zhen-Yu Zhang<sup>407</sup>, Wenhua Zhao<sup>216</sup>, Shiqi Zhen<sup>444</sup>, Yingfeng Zheng<sup>445</sup>, Bekbolat Zholdin<sup>446</sup>, Baurzhan Zhussupov<sup>80</sup>, Nada Zoghlaei<sup>54</sup>, Julio Zuñiga Cisneros<sup>331</sup>, Edward W. Gregg<sup>1</sup> & Majid Ezzati<sup>1447</sup>.
- <sup>1</sup>Imperial College London, London, UK. <sup>2</sup>Harvard T. H. Chan School of Public Health, Boston, MA, USA. <sup>3</sup>University of Auckland, Auckland, New Zealand. <sup>4</sup>Tehran University of Medical Sciences, Tehran, Iran. <sup>5</sup>Middlesex University, London, UK. <sup>6</sup>Shahid Beheshti University of Medical Sciences, Tehran, Iran. <sup>7</sup>Rush University Medical Center, Chicago, IL, USA. <sup>8</sup>Institute for Clinical Effectiveness and Health Policy, Buenos Aires, Argentina. <sup>9</sup>Harvard Medical School, Boston, MA, USA. <sup>10</sup>Erasmus Medical Center Rotterdam, Rotterdam, The Netherlands. <sup>11</sup>World Health Organization, Geneva, Switzerland. <sup>12</sup>Independent researcher, Los Angeles, CA, USA. <sup>13</sup>University of California Berkeley, Berkeley, CA, USA. <sup>14</sup>Mahidol University, Nakhon Pathom, Thailand. <sup>15</sup>Charles University in Prague, Prague, Czech Republic. <sup>16</sup>Thomayer Hospital, Prague, Czech Republic. <sup>17</sup>Istituto Superiore di Sanità, Rome, Italy. <sup>18</sup>South African Medical Research Council, Cape Town, South Africa. <sup>19</sup>Seoul National University, Seoul, Republic of Korea. <sup>20</sup>Finnish Institute for Health and Welfare, Helsinki, Finland. <sup>21</sup>ICMR-National Institute of Nutrition, Hyderabad, India. <sup>22</sup>Pontificia Universidad Católica de Chile, Santiago, Chile. <sup>23</sup>ICMR-National Centre for Disease Informatics and Research, Bengaluru, India. <sup>24</sup>Copenhagen University Hospital, Copenhagen, Denmark. <sup>25</sup>Capital Medical University Beijing An Zhen Hospital, Beijing, China. <sup>26</sup>Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark. <sup>27</sup>Caja Costarricense de Seguro Social, San José, Costa Rica. <sup>28</sup>Qatar University, Doha, Qatar. <sup>29</sup>Birzeit University, Birzeit, Palestine. <sup>30</sup>Instituto Mexicano del Seguro Social, Mexico City, Mexico. <sup>31</sup>Flinders University, Adelaide, South Australia, Australia. <sup>32</sup>Food and Nutrition Research Institute, Taguig, The Philippines. <sup>33</sup>Urmia University of Medical Sciences, Urmia, Iran. <sup>34</sup>Instituto Nacional de Ciencias Médicas y Nutricion, Mexico City, Mexico. <sup>35</sup>University of Amsterdam, Amsterdam, The Netherlands. <sup>36</sup>Steno Diabetes Center Copenhagen, Gentofte, Denmark. <sup>37</sup>Ministry of Health Malaysia, Kuala Lumpur, Malaysia. <sup>38</sup>Shahrekord University of Medical Sciences, Shahrekord, Iran. <sup>39</sup>University of Oslo, Oslo, Norway. <sup>40</sup>University of Bremen, Bremen, Germany. <sup>41</sup>National Center for Diabetes, Endocrinology and Genetics, Amman, Jordan. <sup>42</sup>Dasman Diabetes Institute, Kuwait City, Kuwait. <sup>43</sup>Aldara Hospital and Medical Center, Riyadh, Saudi Arabia. <sup>44</sup>King Abdullah International Medical Research Center, Riyadh, Saudi Arabia. <sup>45</sup>Luxembourg Institute of Health, Strassen, Luxembourg. <sup>46</sup>World Health Organization Regional Office for the Eastern Mediterranean, Cairo, Egypt. <sup>47</sup>Bombay Hospital and Medical Research Centre, Mumbai, India. <sup>48</sup>University of Lille, Lille, France. <sup>49</sup>Lille University Hospital, Lille, France. <sup>50</sup>Western Norway University of Applied Sciences, Sogndal, Norway. <sup>51</sup>Norwegian School of Sport Sciences, Oslo, Norway. <sup>52</sup>Madras Diabetes Research Foundation, Chennai, India. <sup>53</sup>Zahedan University of Medical Sciences, Zahedan, Iran. <sup>54</sup>National Institute of Public Health, Tunis, Tunisia. <sup>55</sup>Institute of Public Health of the University of Porto, Porto, Portugal. <sup>56</sup>Norwegian Institute of Public Health, Oslo, Norway. <sup>57</sup>University of Massachusetts, Amherst, MA, USA. <sup>58</sup>Abi Associates, Kathmandu, Nepal. <sup>59</sup>University of Iceland, Reykjavík, Iceland. <sup>60</sup>Federal University of Pelotas, Pelotas, Brazil. <sup>61</sup>University of Oulu, Oulu, Finland. <sup>62</sup>Oulu University Hospital, Oulu, Finland. <sup>63</sup>Regional Authority of Public Health, Banska Bystrica, Slovakia. <sup>64</sup>University of Porto Medical School, Porto, Portugal. <sup>65</sup>Research Institute for Endocrine Sciences, Tehran, Iran. <sup>66</sup>University of Science and Technology, Sana'a, Yemen. <sup>67</sup>Medical University of Lodz, Lodz, Poland. <sup>68</sup>Medical University of Gdańsk, Gdańsk, Poland. <sup>69</sup>Universidad Autónoma de Madrid/CIBERESP, Madrid, Spain. <sup>70</sup>University of Palermo, Palermo, Italy. <sup>71</sup>Pan American Health Organization, Washington, DC, USA. <sup>72</sup>Mohammed V University of Rabat, Rabat, Morocco. <sup>73</sup>Dalhousie University, Halifax, Nova Scotia, Canada. <sup>74</sup>Jordan University of Science and Technology, Irbid, Jordan. <sup>75</sup>Al-Farabi Kazakh National University, Almaty, Kazakhstan. <sup>76</sup>University of Sydney, Sydney, New South Wales, Australia. <sup>77</sup>Christian Medical College, Vellore, India. <sup>78</sup>University Tunis El Manar, Tunis, Tunisia. <sup>79</sup>Cafam University Foundation, Bogota, Colombia. <sup>80</sup>Kazakh National Medical University, Almaty, Kazakhstan. <sup>81</sup>Universidad Peruana Cayetano Heredia, Lima, Peru. <sup>82</sup>Lithuanian University of Health Sciences, Kaunas, Lithuania. <sup>83</sup>University of São Paulo, São Paulo, Brazil. <sup>84</sup>Sunder Lal Jain Hospital, Delhi, India. <sup>85</sup>Shanghai Jiao-Tong University School of Medicine, Shanghai, China. <sup>86</sup>Public Health Agency of Canada, Ottawa, Ontario, Canada. <sup>87</sup>Ufa Eye Research Institute, Ufa, Russia. <sup>88</sup>Nepal Health Research Council, Kathmandu, Nepal. <sup>89</sup>University of Southern Denmark, Copenhagen, Denmark. <sup>90</sup>University of Gothenburg, Gothenburg, Sweden. <sup>91</sup>Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil. <sup>92</sup>National Institute for Public Health and the Environment, Bilthoven, The Netherlands. <sup>93</sup>University of Turin, Turin, Italy. <sup>94</sup>Nanyang Technological University, Singapore, Singapore. <sup>95</sup>Universidad de la República, Montevideo, Uruguay. <sup>96</sup>Centro de Educación Médica e Investigaciones Clínicas, Buenos Aires, Argentina. <sup>97</sup>IRCCS Neuromed, Pozzilli, Italy. <sup>98</sup>Toulouse University School of Medicine, Toulouse, France. <sup>99</sup>University Hospital of Varese, Varese, Italy. <sup>100</sup>University Hospital KU Leuven, Leuven, Belgium. <sup>101</sup>Ministry of Health, Victoria, Seychelles. <sup>102</sup>University of Lausanne, Lausanne, Switzerland. <sup>103</sup>Universidad Central de Venezuela, Caracas, Venezuela. <sup>104</sup>Bielefeld University, Bielefeld, Germany. <sup>105</sup>German Cancer Research Center, Heidelberg, Germany. <sup>106</sup>University College Copenhagen, Copenhagen, Denmark. <sup>107</sup>Robert Koch Institute, Berlin, Germany. <sup>108</sup>Universidad de La Laguna, Tenerife, Spain. <sup>109</sup>University of Malta, Msida, Malta. <sup>110</sup>Istanbul University - Cerrahpasa, Istanbul, Turkey. <sup>111</sup>Universidade Federal de Juiz de Fora, Juiz de Fora, Brazil. <sup>112</sup>Gaetano Fucito Hospital, Mercato San Severino, Italy. <sup>113</sup>University of Porto, Porto, Portugal. <sup>114</sup>Santiago de Compostela University, Santiago, Spain. <sup>115</sup>Council for Agricultural Research and Economics, Rome, Italy. <sup>116</sup>Research and Education Institute of Child Health, Nicosia, Cyprus. <sup>117</sup>Dr. A. Ramachandran's Diabetes Hospital, Chennai, India. <sup>118</sup>University College London, London, UK. <sup>119</sup>Academia Sinica, Taipei, Taiwan. <sup>120</sup>Capital Institute of Pediatrics, Beijing, China. <sup>121</sup>Kailuan General Hospital, Tangshan, China. <sup>122</sup>Duke-NUS Medical School, Singapore, Singapore. <sup>123</sup>Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran. <sup>124</sup>The Gertner Institute for Epidemiology and Health Policy Research, Ramat Gan, Israel. <sup>125</sup>Ministry of Health and Welfare, Taipei, Taiwan. <sup>126</sup>Murcia Health Council, Murcia, Spain. <sup>127</sup>Seoul National University College of Medicine, Seoul, Republic of Korea. <sup>128</sup>Korea Centers for Disease Control and Prevention, Cheongju-si, Republic of Korea. <sup>129</sup>Medical University of Silesia, Katowice, Poland. <sup>130</sup>Katholieke Universiteit Leuven, Leuven, Belgium. <sup>131</sup>Statistics Canada, Ottawa, Ontario, Canada. <sup>132</sup>Ghent University, Ghent, Belgium. <sup>133</sup>Agency for Preventive and Social Medicine, Bregenz, Austria. <sup>134</sup>Federal University of Maranhão, São Luís, Brazil. <sup>135</sup>University of Southampton, Southampton, UK. <sup>136</sup>Institut Pasteur de Lille, Lille, France. <sup>137</sup>CIBERONB, Madrid, Spain. <sup>138</sup>University of Debrecen, Debrecen, Hungary. <sup>139</sup>Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil. <sup>140</sup>National Council of Research, Reggio Calabria, Italy. <sup>141</sup>Federal University of Santa Catarina, Florianópolis, Brazil. <sup>142</sup>Eduardo Mondlane University, Maputo, Mozambique. <sup>143</sup>Geneva University Hospitals, Geneva, Switzerland. <sup>144</sup>National Research Centre for Preventive Medicine, Moscow, Russia. <sup>145</sup>University of Montreal, Montreal, Québec, Canada. <sup>146</sup>French Public Health Agency, St Maurice, France. <sup>147</sup>Mediterranea Cardiocentro, Naples, Italy. <sup>148</sup>University of Zagreb, Zagreb, Croatia. <sup>149</sup>Ministry of Health and Medical Education, Tehran, Iran. <sup>150</sup>University of Queensland, Brisbane, Queensland, Australia. <sup>151</sup>Universidad de Cuenca, Cuenca, Ecuador. <sup>152</sup>Helmholtz Zentrum München, Munich, Germany. <sup>153</sup>Carol Davila University of Medicine and Pharmacy, Bucharest, Romania. <sup>154</sup>University Hospital Düsseldorf, Düsseldorf, Germany. <sup>155</sup>National Institute of Cardiology, Warsaw, Poland. <sup>156</sup>Beth Israel Deaconess Medical Center, Boston, MA, USA. <sup>157</sup>University of Latvia, Riga, Latvia. <sup>158</sup>Ministry of Health and Medical Services, Gizo, Solomon Islands. <sup>159</sup>Hormozgan University of Medical Sciences, Bandar Abbas, Iran. <sup>160</sup>University of Skövde, Skövde, Sweden. <sup>161</sup>National Institute of Nutrition and Food Technology, Tunis, Tunisia. <sup>162</sup>The University of the West Indies, Kingston, Jamaica. <sup>163</sup>Institut

# Article

- Hospital del Mar d'Investigacions Mèdiques, Barcelona, Spain. <sup>164</sup>University of Stellenbosch, Cape Town, South Africa. <sup>165</sup>University of Duisburg-Essen, Duisburg, Germany. <sup>166</sup>Karadeniz Technical University, Trabzon, Turkey. <sup>167</sup>University of Helsinki, Helsinki, Finland. <sup>168</sup>Mashhad University of Medical Sciences, Mashhad, Iran. <sup>169</sup>Rafsanjan University of Medical Sciences, Rafsanjan, Iran. <sup>170</sup>Queen's University of Belfast, Belfast, UK. <sup>171</sup>University of Zurich, Zurich, Switzerland. <sup>172</sup>Tabriz University of Medical Sciences, Tabriz, Iran. <sup>173</sup>Fasa University of Medical Sciences, Fasa, Iran. <sup>174</sup>Shiraz University of Medical Sciences, Shiraz, Iran. <sup>175</sup>Centro de Salud Villanueva Norte, Badajoz, Spain. <sup>176</sup>Servicio Extremeño de Salud, Badajoz, Spain. <sup>177</sup>Ministry of Health, Buenos Aires, Argentina. <sup>178</sup>Universiti Sains Malaysia, Kelantan, Malaysia. <sup>179</sup>Umeå University, Umeå, Sweden. <sup>180</sup>Federal University of São Paulo, São Paulo, Brazil. <sup>181</sup>Hospital Universitario Son Espases, Palma, Spain. <sup>182</sup>Kindai University, Osaka-Sayama, Japan. <sup>183</sup>Kyoto University, Kyoto, Japan. <sup>184</sup>Medical University of Warsaw, Warsaw, Poland. <sup>185</sup>University of Catania, Catania, Italy. <sup>186</sup>CIBER en Epidemiología y Salud Pública, Alicante, Spain. <sup>187</sup>CIBER en Epidemiología y Salud Pública, Barcelona, Spain. <sup>188</sup>Universidade Federal de Minas Gerais, Belo Horizonte, Brazil. <sup>189</sup>Wageningen University, Wageningen, The Netherlands. <sup>190</sup>B. P. Koirala Institute of Health Sciences, Dharan, Nepal. <sup>191</sup>University of Insubria, Varese, Italy. <sup>192</sup>University of Adelaide, Adelaide, South Australia, Australia. <sup>193</sup>Lund University, Lund, Sweden. <sup>194</sup>McGill University, Montreal, Québec, Canada. <sup>195</sup>Universidad Politécnica de Madrid, Madrid, Spain. <sup>196</sup>St Anne's University Hospital, Brno, Czech Republic. <sup>197</sup>National Institute of Public Health, Cuernavaca, Mexico. <sup>198</sup>Centro de Estudios en Diabetes A.C., Mexico City, Mexico. <sup>199</sup>Universidad Autónoma de Santo Domingo, Santo Domingo, Dominican Republic. <sup>200</sup>Institute for Clinical and Experimental Medicine, Prague, Czech Republic. <sup>201</sup>Jagiellonian University Medical College, Kraków, Poland. <sup>202</sup>University of Southern Denmark, Odense, Denmark. <sup>203</sup>National Center of Cardiovascular Diseases, Beijing, China. <sup>204</sup>Singapore Eye Research Institute, Singapore, Singapore. <sup>205</sup>Icelandic Heart Association, Kopavogur, Iceland. <sup>206</sup>Universidad Icesi, Cali, Colombia. <sup>207</sup>Eternal Heart Care Centre and Research Institute, Jaipur, India. <sup>208</sup>National Health Insurance Service, Wonju, Republic of Korea. <sup>209</sup>Prevention of Metabolic Disorders Research Center, Tehran, Iran. <sup>210</sup>The University of the West Indies, Cave Hill, Barbados. <sup>211</sup>Kermanshah University of Medical Sciences, Kermanshah, Iran. <sup>212</sup>University of Eastern Finland, Kuopio, Finland. <sup>213</sup>Kyushu University, Fukuoka, Japan. <sup>214</sup>University of Bergen, Bergen, Norway. <sup>215</sup>Tulane University, New Orleans, LA, USA. <sup>216</sup>Chinese Center for Disease Control and Prevention, Beijing, China. <sup>217</sup>Joep Lange Institute, Amsterdam, The Netherlands. <sup>218</sup>Chronic Diseases Research Center, Tehran, Iran. <sup>219</sup>University of Hong Kong, Hong Kong, China. <sup>220</sup>The Chinese University of Hong Kong, Hong Kong, China. <sup>221</sup>University of Western Australia, Perth, Western Australia, Australia. <sup>222</sup>Kingston Health Sciences Centre, Kingston, Ontario, Canada. <sup>223</sup>Heart Institute, São Paulo, Brazil. <sup>224</sup>Fundación Oftalmológica de Santander, Bucaramanga, Colombia. <sup>225</sup>University Oran 1, Oran, Algeria. <sup>226</sup>Independent Public Health Specialist, Nay Pyi Taw, Myanmar. <sup>227</sup>Ministry of Health and Sports, Nay Pyi Taw, Myanmar. <sup>228</sup>CIBER en Epidemiología y Salud Pública, Murcia, Spain. <sup>229</sup>VU University Medical Center, Amsterdam, The Netherlands. <sup>230</sup>International Agency for Research on Cancer, Lyon, France. <sup>231</sup>American University of Beirut, Beirut, Lebanon. <sup>232</sup>Cairo University, Cairo, Egypt. <sup>233</sup>University of Zaragoza, Zaragoza, Spain. <sup>234</sup>National Institutes of Biomedical Innovation, Health and Nutrition, Tokyo, Japan. <sup>235</sup>Medical University Varna, Varna, Bulgaria. <sup>236</sup>The University of Tokyo, Tokyo, Japan. <sup>237</sup>The Hospital for Sick Children, Toronto, Ontario, Canada. <sup>238</sup>Niigata University, Niigata, Japan. <sup>239</sup>Hadassah University Medical Center, Jerusalem, Israel. <sup>240</sup>Norwegian University of Science and Technology, Trondheim, Norway. <sup>241</sup>University of Melbourne, Melbourne, Victoria, Australia. <sup>242</sup>University Hospital Centre Zagreb, Zagreb, Croatia. <sup>243</sup>University of Zagreb School of Medicine, Zagreb, Croatia. <sup>244</sup>Heart Foundation, Melbourne, Victoria, Australia. <sup>245</sup>Guangzhou 12th Hospital, Guangzhou, China. <sup>246</sup>Universidad Eugenio María de Hostos, Santo Domingo, Dominican Republic. <sup>247</sup>Simon Fraser University, Burnaby, British Columbia, Canada. <sup>248</sup>Ruprecht-Karls-University of Heidelberg, Heidelberg, Germany. <sup>249</sup>World Health Organization Country Office, Delhi, India. <sup>250</sup>Guilan University of Medical Sciences, Rasht, Iran. <sup>251</sup>University of Opole, Opole, Poland. <sup>252</sup>University of Crete, Heraklion, Greece. <sup>253</sup>Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia. <sup>254</sup>Maharajgunj Medical Campus, Kathmandu, Nepal. <sup>255</sup>Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. <sup>256</sup>National Institute of Epidemiology, Chennai, India. <sup>257</sup>University of Münster, Münster, Germany. <sup>258</sup>Research Institute for Primordial Prevention of Non-communicable Disease, Isfahan, Iran. <sup>259</sup>Kyrgyz State Medical Academy, Bishkek, Kyrgyzstan. <sup>260</sup>Research Institute of Child Nutrition, Dortmund, Germany. <sup>261</sup>Mazandaran University of Medical Sciences, Sari, Iran. <sup>262</sup>Hypertension Research Center, Isfahan, Iran. <sup>263</sup>Medical University of Innsbruck, Innsbruck, Austria. <sup>264</sup>Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania. <sup>265</sup>Yonsei University College of Medicine, Seoul, Republic of Korea. <sup>266</sup>National Cancer Center, Goyang-si, Republic of Korea. <sup>267</sup>University of Pennsylvania, Philadelphia, PA, USA. <sup>268</sup>University of Vienna, Vienna, Austria. <sup>269</sup>Oulu Deaconess Institute Foundation, Oulu, Finland. <sup>270</sup>Tartu University Clinics, Tartu, Estonia. <sup>271</sup>Kansai Medical University, Hirakata, Japan. <sup>272</sup>Ministry of Health and Quality of Life, Port Louis, Mauritius. <sup>273</sup>University Hospital Ulm, Ulm, Germany. <sup>274</sup>University of Groningen, Groningen, The Netherlands. <sup>275</sup>University of Jyväskylä, Jyväskylä, Finland. <sup>276</sup>African Population and Health Research Center, Nairobi, Kenya. <sup>277</sup>Higher Institute of Health Sciences of Settat, Settat, Morocco. <sup>278</sup>Ministry of Health, Algiers, Algeria. <sup>279</sup>Harokopio University, Athens, Greece. <sup>280</sup>Sahlgrenska Academy, Gothenburg, Sweden. <sup>281</sup>Endocrinology and Metabolism Research Center, Tehran, Iran. <sup>282</sup>University of Public Health, Yangon, Myanmar. <sup>283</sup>National University of Singapore, Singapore, Singapore. <sup>284</sup>Tampere University Hospital, Tampere, Finland. <sup>285</sup>Tampere University, Tampere, Finland. <sup>286</sup>University of Cape Town, Cape Town, South Africa. <sup>287</sup>West Virginia University, Morgantown, WV, USA. <sup>288</sup>Oswaldo Cruz Foundation René Rachou Research Institute, Belo Horizonte, Brazil. <sup>289</sup>National Taiwan University, Taipei, Taiwan. <sup>290</sup>University of Chinese Academy of Sciences, Shanghai, China. <sup>291</sup>Uppsala University, Uppsala, Sweden. <sup>292</sup>National Institute for Health Development, Tallinn, Estonia. <sup>293</sup>Universidad San Martín de Porres, Lima, Peru. <sup>294</sup>Consejería de Sanidad Junta de Castilla y León, Valladolid, Spain. <sup>295</sup>Norrbotten County Council, Luleå, Sweden. <sup>296</sup>Peking University, Beijing, China. <sup>297</sup>Universidade Federal de Ouro Preto, Ouro Preto, Brazil. <sup>298</sup>University of Coimbra, Coimbra, Portugal. <sup>299</sup>The Jikei University School of Medicine, Tokyo, Japan. <sup>300</sup>Institute of Neuroscience of the National Research Council, Padua, Italy. <sup>301</sup>Baker Heart and Diabetes Institute, Melbourne, Victoria, Australia. <sup>302</sup>Agricultural University of Athens, Athens, Greece. <sup>303</sup>French National Research Institute for Sustainable Development, Montpellier, France. <sup>304</sup>Hospital Israelita Albert Einstein, São Paulo, Brazil. <sup>305</sup>University of Otago, Dunedin, New Zealand. <sup>306</sup>University of Padua, Padua, Italy. <sup>307</sup>Lausanne University Hospital, Lausanne, Switzerland. <sup>308</sup>Emory University, Atlanta, GA, USA. <sup>309</sup>Universidade Estadual do Centro-Oeste, Guarapuava, Brazil. <sup>310</sup>UiT The Arctic University of Norway, Tromsø, Norway. <sup>311</sup>Cape Peninsula University of Technology, Cape Town, South Africa. <sup>312</sup>Brown University, Providence, RI, USA. <sup>313</sup>University of Edinburgh, Edinburgh, UK. <sup>314</sup>University College Dublin, Dublin, Ireland. <sup>315</sup>Shahid Sadoughi University of Medical Sciences, Yazd, Iran. <sup>316</sup>Institut National de la Santé et de la Recherche Médicale, Lille, France. <sup>317</sup>ICMR-National Institute of Medical Statistics, New Delhi, India. <sup>318</sup>Lusófona University, Lisbon, Portugal. <sup>319</sup>Women's Social and Health Studies Foundation, Trivandrum, India. <sup>320</sup>Università degli Studi di Firenze, Florence, Italy. <sup>321</sup>Ain Shams University, Cairo, Egypt. <sup>322</sup>Isfahan Cardiovascular Research Center, Isfahan, Iran. <sup>323</sup>University of Strasbourg, Strasbourg, France. <sup>324</sup>Strasbourg University Hospital, Strasbourg, France. <sup>325</sup>University of Pécs, Pécs, Hungary. <sup>326</sup>Mulago Hospital, Kampala, Uganda. <sup>327</sup>University of Medical Sciences of Cienfuegos, Cienfuegos, Cuba. <sup>328</sup>Royal College of Surgeons in Ireland Dublin, Dublin, Ireland. <sup>329</sup>La Trobe University, Melbourne, Victoria, Australia. <sup>330</sup>International Institute of Molecular and Cell Biology, Warsaw, Poland. <sup>331</sup>Instituto Commemorativo Gorgas de Estudios de la Salud, Panama City, Panama. <sup>332</sup>World Health Organization Country Office, Lilongwe, Malawi. <sup>333</sup>University of Brescia, Brescia, Italy. <sup>334</sup>Bushehr University of Medical Sciences, Bushehr, Iran. <sup>335</sup>Ulm University, Ulm, Germany. <sup>336</sup>Kobe University, Kobe, Japan. <sup>337</sup>Suraj Eye Institute, Nagpur, India. <sup>338</sup>University Medicine of Greifswald, Greifswald, Germany. <sup>339</sup>University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, Vietnam. <sup>340</sup>Hanoi Medical University, Hanoi, Vietnam. <sup>341</sup>Miami Veterans Affairs Healthcare System, Miami, FL, USA. <sup>342</sup>University of Turku, Turku, Finland. <sup>343</sup>Eastern Mediterranean Public Health Network, Amman, Jordan. <sup>344</sup>University of Manchester, Manchester, UK. <sup>345</sup>Japan Wildlife Research Center, Tokyo, Japan. <sup>346</sup>Istanbul University, Istanbul, Turkey. <sup>347</sup>Ministry of Health, Bandar Seri Begawan, Brunei. <sup>348</sup>University of Madeira, Funchal, Portugal. <sup>349</sup>MRC Lifecourse Epidemiology Unit, Southampton, UK. <sup>350</sup>University of Novi Sad, Novi Sad, Serbia. <sup>351</sup>Kwame Nkrumah University of Science and Technology, Kumasi, Ghana. <sup>352</sup>Institute for Social and Preventive Medicine, Ottawa, Ontario, Canada. <sup>353</sup>IRCCS Ente Ospedaliero Specializzato in Gastroenterologia S. de Bellis, Bari, Italy. <sup>354</sup>Jivandeep Hospital, Anand, India. <sup>355</sup>South African Medical Research Council, Durban, South Africa. <sup>356</sup>Vietnam National Heart Institute, Hanoi, Vietnam. <sup>357</sup>Clinica de Medicina Avanzada Dr. Abel González, Santo Domingo, Dominican Republic. <sup>358</sup>Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany. <sup>359</sup>University of Sarajevo, Sarajevo, Bosnia and Herzegovina. <sup>360</sup>Cardiovascular Prevention Centre, Udine, Italy. <sup>361</sup>Public Health Agency of Catalonia, Barcelona, Spain. <sup>362</sup>Observatorio de Salud Pública de Santander, Bucaramanga, Colombia. <sup>363</sup>Ardabil University of Medical Sciences, Ardabil, Iran. <sup>364</sup>Alborz University of Medical Sciences, Karaj, Iran. <sup>365</sup>Ministry of Health, Hanoi, Vietnam. <sup>366</sup>India Diabetes Research Foundation, Chennai, India. <sup>367</sup>Institut Universitari d'Investigació en Atenció Primària Jordi Gol, Girona, Spain. <sup>368</sup>Universiti Putra Malaysia, Serdang, Malaysia. <sup>369</sup>University of Malaya, Kuala Lumpur, Malaysia. <sup>370</sup>University of Valencia, Valencia, Spain. <sup>371</sup>University of the Philippines, Manila, The Philippines. <sup>372</sup>Minas Gerais State Secretariat for Health, Belo Horizonte, Brazil. <sup>373</sup>CS S. Agustín Ibsalut, Palma, Spain. <sup>374</sup>Amsterdam Institute for Global Health and Development, Amsterdam, The Netherlands. <sup>375</sup>Canarian Health Service, Tenerife, Spain. <sup>376</sup>Universidad Industrial de Santander, Bucaramanga, Colombia. <sup>377</sup>Associazione Calabrese di Epatologia, Reggio Calabria, Italy. <sup>378</sup>Sahlgrenska University Hospital, Gothenburg, Sweden. <sup>379</sup>Toulouse University Hospital, Toulouse, France. <sup>380</sup>Institute of Food Sciences of the National Research Council, Avellino, Italy. <sup>381</sup>Sitaram Bhartia Institute of Science and Research, New Delhi, India. <sup>382</sup>Faculty of Medicine of Tunis, Tunis, Tunisia. <sup>383</sup>National Institute of Health, Lima, Peru. <sup>384</sup>Catalan Department of Health, Barcelona, Spain. <sup>385</sup>Universidade de Lisboa, Lisbon, Portugal. <sup>386</sup>South Karelia Social and Health Care District, Lappeenranta, Finland. <sup>387</sup>Cardiovascular Research Institute, Isfahan, Iran. <sup>388</sup>National Cancer Center, Tokyo, Japan. <sup>389</sup>University of São Paulo Clinics Hospital, São Paulo, Brazil. <sup>390</sup>Hospital Italiano de Buenos Aires, Buenos Aires, Argentina. <sup>391</sup>Center for Oral Health Services and Research Mid-Norway, Trondheim, Norway. <sup>392</sup>King's College London, London, UK. <sup>393</sup>National Center for Global Health and Medicine, Tokyo, Japan. <sup>394</sup>Sungkyunkwan University, Seoul, Republic of Korea. <sup>395</sup>Finnish Institute of Occupational Health, Helsinki, Finland. <sup>396</sup>St Vincent's Hospital, Sydney, New South Wales, Australia. <sup>397</sup>University of New South Wales, Sydney, New South Wales, Australia. <sup>398</sup>Karolinska Institutet, Stockholm, Sweden. <sup>399</sup>Research Centre for Prevention and Health, Glostrup, Denmark. <sup>400</sup>London School of Hygiene & Tropical Medicine, London, UK. <sup>401</sup>Diponegoro University, Semarang, Indonesia. <sup>402</sup>University of Bari, Bari, Italy. <sup>403</sup>University of Copenhagen, Copenhagen, Denmark. <sup>404</sup>Institut Régional de Santé Publique, Ouidah, Benin. <sup>405</sup>University of Bordeaux, Bordeaux, France. <sup>406</sup>Izmir Katip Çelebi University, Izmir, Turkey. <sup>407</sup>University of Leuven, Leuven, Belgium. <sup>408</sup>Institut National de la Santé et de la Recherche Médicale, Nancy, France. <sup>409</sup>Bonn University, Bonn, Germany. <sup>410</sup>Croatian Institute of Public Health, Zagreb, Croatia. <sup>411</sup>National Institute of Public Health–National Institute of Hygiene, Warsaw, Poland. <sup>412</sup>National Institute of Hygiene, Epidemiology and Microbiology, Havana, Cuba. <sup>413</sup>Fu Jen Catholic University, Taipei, Taiwan. <sup>414</sup>National Statistic Office of Cabo Verde, Praia, Cabo Verde. <sup>415</sup>Ministry of Health, Amman, Jordan. <sup>416</sup>Central University of Kerala,

Kasaragod, India.<sup>417</sup>Health Service of Murcia, Murcia, Spain.<sup>418</sup>Institut d'Investigacio Sanitaria Illes Balears, Menorca, Spain.<sup>419</sup>Universidad Centro-Occidental Lisandro Alvarado, Barquisimeto, Venezuela.<sup>420</sup>Dokuz Eylul University, Izmir, Turkey.<sup>421</sup>University of Tampere Tays Eye Center, Tampere, Finland.<sup>422</sup>Icahn School of Medicine at Mount Sinai, New York City, NY, USA.<sup>423</sup>Utrecht University, Utrecht, The Netherlands.<sup>424</sup>Hanoi University of Public Health, Hanoi, Vietnam.<sup>425</sup>University Medical Center Utrecht, Utrecht, The Netherlands.<sup>426</sup>Universitas Indonesia, Jakarta, Indonesia.<sup>427</sup>Instituto de Investigación Sanitaria y Biomédica de Alicante, Alicante, Spain.<sup>428</sup>North Karelian Center for Public Health, Joensuu, Finland.<sup>429</sup>University of the Witwatersrand, Johannesburg, South Africa.<sup>430</sup>Cork Institute of Technology, Cork, Ireland.<sup>431</sup>Institute for Medical Research, Kuala Lumpur, Malaysia.<sup>432</sup>Health Canada, Ottawa, Ontario, Canada.<sup>433</sup>Beijing Institute of Ophthalmology, Beijing, China.<sup>434</sup>Xinjiang Medical University, Urumqi, China.<sup>435</sup>Capital Medical University, Beijing, China.<sup>436</sup>St George's, University of London, London, UK.<sup>437</sup>Medical University of Vienna, Vienna, Austria.<sup>438</sup>University of Oxford, Oxford, UK.<sup>439</sup>Institute of Food and Nutrition Development of Ministry of Agriculture and Rural Affairs, Beijing, China.<sup>440</sup>Children's Hospital of Fudan University, Shanghai, China.<sup>441</sup>Penang Medical College, Penang, Malaysia.<sup>442</sup>University of Cyprus, Nicosia, Cyprus.<sup>443</sup>Iran University of Medical Sciences, Tehran, Iran.<sup>444</sup>Jiangsu Provincial Center for Disease Control and Prevention, Nanjing, China.<sup>445</sup>Sun Yat-sen University, Guangzhou, China.<sup>446</sup>West Kazakhstan State Medical University, Aktobe, Kazakhstan.<sup>447</sup>University of Ghana, Accra, Ghana.<sup>448</sup>Deceased: Konrad Jamrozik, Altan Onat, Robespierre Ribeiro, Jutta Stieber.✉e-mail: majid.ezzati@imperial.ac.uk

# Article

## Methods

Our aim was to estimate trends in mean total, HDL and non-HDL cholesterol for 200 countries and territories (Supplementary Table 2). We used non-HDL cholesterol rather than low-density lipoprotein (LDL) cholesterol because most studies in our analysis had measured total cholesterol and HDL cholesterol, from which non-HDL cholesterol can be calculated through subtraction. By contrast, LDL cholesterol was directly measured in only around 14% of studies. When LDL cholesterol is not directly measured, its calculation requires data on triglycerides, which were available in approximately 64% of the studies. Furthermore, the most-commonly used estimation method—that is, the Friedewald equation—can be inaccurate, particularly at high levels of triglycerides<sup>34</sup>. Non-HDL and LDL cholesterol were highly correlated (Pearson correlation coefficient = 0.94) in studies with data on both variables (Extended Data Fig. 9), because LDL cholesterol constitutes most of non-HDL cholesterol. Furthermore, non-HDL cholesterol predicts IHD risk at least as well as LDL cholesterol<sup>5,35</sup>, and can be measured at a lower cost than LDL cholesterol, which is relevant for how widely it can be used in low- and middle-income countries. Although non-HDL cholesterol is now commonly used in clinical guidelines<sup>36–38</sup>, LDL cholesterol continues to be a key target for treatment<sup>36,37</sup>, possibly because the interpretation of non-HDL cholesterol is more complex than LDL cholesterol alone. Specifically, an increase in non-HDL cholesterol could be due to the increase in LDL cholesterol or very-low-density lipoprotein cholesterol<sup>39</sup>. Furthermore, there is some evidence that triglyceride levels are high in Asian populations, compared to levels seen in high-income western countries<sup>40</sup>. Therefore, data on non-HDL cholesterol can motivate dietary interventions to both reduce LDL cholesterol (for example, reducing saturated and trans fat intake) and triglyceride levels (for example, reducing refined carbohydrates and increasing omega-3 fatty acids) as well as treatments that lower LDL cholesterol (statins), alongside those that lower triglycerides (for example, fibrates).

### Data sources

We used a database of population-based data on cardiometabolic risk factors collated by the NCD Risk Factor Collaboration (NCD-RisC), a worldwide network of health researchers and practitioners that systematically monitors the worldwide trends and variations in non-communicable disease (NCD) risk factors. The database was collated through multiple routes for identifying and accessing data. We accessed publicly available population-based multi-country and national measurement surveys (for example, Demographic and Health Surveys and surveys identified through the Inter-University Consortium for Political and Social Research and European Health Interview & Health Examination Surveys Database). We requested, via the World Health Organization (WHO) and its regional and country offices, from ministries of health and other national health and statistical agencies to identify and access population-based surveys. Requests were also sent via the World Heart Federation to its national partners. We made a similar request to the co-authors of an earlier pooled analysis of cardiometabolic risk factors<sup>7,41–43</sup>, and invited the co-authors of the analysis to reanalyse data from their studies and join NCD-RisC. Finally, to identify major sources that were not accessed through the above routes, we searched and reviewed published studies as described in the Supplementary Information and invited all eligible studies to join NCD-RisC.

For each data source, we recorded the available information about the study population, start year and duration of measurement, sampling approach and measurement methods. The information about study population was used to establish that each data source was population-based, and to assess whether it covered the whole country, multiple subnational regions or one or a small number of communities, and whether it was rural, urban or combined.

We carefully checked all data sources in terms of how they met our inclusion and exclusion criteria listed below. We identified duplicate data sources by comparing studies from the same country and year. Additionally, all NCD-RisC members are asked periodically to review the list of sources from their country, to suggest additional sources not in the database, and to verify that the included data meet the inclusion criteria listed below and are not duplicates. The NCD-RisC database is continuously updated through the above routes and through regular contact with NCD-RisC members.

Anonymized individual record data from sources included in NCD-RisC were reanalysed according to a common protocol. Within each survey, we included participants aged 18 years and older who were not pregnant. We removed participants with implausible total cholesterol levels (defined as total cholesterol levels of  $<1.75 \text{ mmol l}^{-1}$  or  $>20 \text{ mmol l}^{-1}$ , or total cholesterol values that were lower than HDL cholesterol values) ( $<0.05\%$  of all participants with total cholesterol measurements) or HDL cholesterol levels (defined as HDL cholesterol levels of  $<0.4 \text{ mmol l}^{-1}$  or  $>5 \text{ mmol l}^{-1}$ , or total cholesterol values that were lower than HDL cholesterol values) ( $<0.15\%$  of all participants with HDL cholesterol measurements). When data on LDL cholesterol were also available, we removed individuals for whom the sum of LDL and HDL cholesterol level surpassed total cholesterol level by more than is plausible based on the limits to errors in their measurement (following the CDC Cholesterol Reference Method Laboratory Network (CRMLN) standards, these errors were set at 8.9% for total cholesterol, 13% for HDL cholesterol and 12% for LDL cholesterol) ( $<0.06\%$  of all participants with total cholesterol and HDL cholesterol measurements)<sup>44–46</sup>.

We calculated mean total cholesterol, mean HDL cholesterol and mean non-HDL cholesterol, and associated standard errors and sample sizes, by sex and age group (18–19 years, 20–29 years, followed by 10-year age groups and 80+ years). All analyses incorporated appropriate sample weights and complex survey design in calculating age–sex-specific means when applicable. To ensure summaries were prepared according to the study protocol, computer code was provided to NCD-RisC members who requested assistance. All submitted data were checked independently by at least two researchers. Questions and clarifications were discussed with NCD-RisC members and resolved before the data were incorporated in the database.

Finally, we obtained data not accessed through the above routes by extracting data from published reports of all additional national health surveys identified through the above-described strategies, as well as eight sites of the WHO Multinational MONItoring of trends and determinants in CArdiovascular disease (MONICA) project that were not deposited in the MONICA Data Centre. Data were extracted from published reports only when reported by sex and in age groups no wider than 20 years. We also used data from a previous pooling study<sup>7</sup> when such data did not overlap with those accessed through the above routes.

### Data inclusion and exclusion

Data sources were included in NCD-RisC database if: (1) measured data on total, LDL, HDL cholesterol and/or triglycerides were available; (2) study participants were 10 years of age or older; (3) data were collected using a probabilistic sampling method with a defined sampling frame; (4) data were from population samples at the national, subnational (covering one or more subnational regions, more than three urban communities or more than five rural communities) or community (one or a small number of communities) level; (5) data were collected in or after 1950; and (6) data were from the countries and territories listed in Supplementary Table 2.

We excluded all data sources that included only hypercholesterolaemia or dyslipidaemia diagnosis history or medication status without measurement of cholesterol levels. We also excluded data sources on population subgroups for which the lipid profile may differ systematically from the general population, including: (1) studies that had included or excluded people on the basis of their health status or

cardiovascular risk; (2) studies for which the participants were only from ethnic minorities; (3) studies that had recruited only specific educational, occupational or socioeconomic subgroups, with the exception noted below; and (4) studies that had recruited participants through health facilities, with the exception noted below.

We used school-based data in countries and for age–sex groups, for which secondary school enrolment was 70% or higher. We used data for which the sampling frame was health insurance schemes in countries in which at least 80% of the population was insured. Finally, we used data collected through general practice and primary-care systems in high-income and central European countries with universal insurance, because contact with the primary-care systems tends to be as good as or better than response rates for population-based surveys. We used data sources regardless of fasting status, because the differences between fasting and non-fasting measurements are negligible for total, non-HDL and HDL cholesterol<sup>39</sup>, and therefore non-fasting lipid profiles are now widely endorsed for the estimation of cardiovascular risk<sup>36,37</sup>.

### Data used in the analysis

For this paper, we used data from the NCD-RisC database for years 1980 to 2018 and individuals aged 18 years and older. A list of the data sources that we used in this analysis and their characteristics is provided in Supplementary Table 1. The data comprised 1,127 population-based measurement surveys and studies that included measurements of blood lipids on 102.6 million participants aged 18 years and older. We had at least one data source for 161 of the 200 countries that we made estimates for, covering 92.4% of the world's population in 2018 (Extended Data Fig. 1); and at least two data sources for 104 countries (87.5% of the world population). Of these 1,127 sources, 409 (36.3%) sampled from national populations, 250 (22.2%) covered one or more subnational regions, and the remaining 468 (41.5%) were from one or a small number of communities. Regionally, data availability ranged from around 2 data sources per country in sub-Saharan Africa to approximately 35 sources per country in the high-income Asia–Pacific region. In total, 454 data sources (40.3%) were from years before 2000 and the remaining 673 (59.7%) were collected from 2001 onwards.

### Adjusting for the differences in mean cholesterol between portable device and laboratory measurements

In 112 (10%) of the 1,127 data sources used in our analysis (11.5% and 5.8% of age–sex-specific data points for total and HDL cholesterol, respectively) lipids were measured using a portable device. Some portable devices have narrower analytical ranges than laboratory methods, which results in truncations of blood cholesterol data that are outside their range (Supplementary Table 3). This may in turn affect the population mean. Although cholesterol concentrations that fall outside the analytical range are displayed as ‘high’ (above the measurement range) or ‘low’ (below the measurement range) by these devices, different surveys record and code cholesterol concentrations outside the analytical range in different ways, for example using ‘too low’, ‘too high’ and ‘error’ codes; assigning the minimum or maximum value to individuals whose cholesterol was below or above the analytical range, respectively; setting values outside the analytical range to missing; and so on. We used an approach that treated surveys with such data consistently.

Specifically, we first dropped all participants with cholesterol levels below and at the minimum, and at and above the maximum, values of the analytical range of each portable device before calculating the mean cholesterol (Supplementary Table 3). We then developed conversion regressions to adjust the mean cholesterol levels measured using a portable device (calculated over the restricted range, Supplementary Table 3) to the levels expected using laboratory measurements. The dependent variable in each regression was mean total, non-HDL or HDL cholesterol for the full range, and the main independent variable was mean total, non-HDL or HDL cholesterol over the above-mentioned restricted cholesterol range of the portable devices. The regression

coefficients were estimated from data sources for which lipids were measured in a laboratory, and thus had the full range of measurement and could be used to calculate both dependent and independent variables. When estimating the regression coefficients, we constructed the dependent variable using the full data, and the independent variable by dropping the values outside the above-mentioned restricted cholesterol range of each device, mimicking those that would be expected if a portable device had been used. Separate models were developed according to the specific range of the different portable devices. All regressions included terms for age and sex, as well as interactions between predictors and age and sex, based on the Bayesian information criterion<sup>47</sup>. The regressions for mean non-HDL cholesterol also included mean total cholesterol and mean HDL cholesterol because non-HDL cholesterol is calculated from total cholesterol and HDL cholesterol. We excluded data points for which there were fewer than 25 individuals for the purpose of estimating the coefficients of these regressions. All sources of uncertainty in the conversion—including the sampling uncertainty of the original data, the uncertainty of the regression coefficients and residuals—were carried forward by using repeated draws from their respective distributions. The regression coefficients and number of data points used to estimate the coefficients are shown in Supplementary Table 4.

### Statistical analysis

We used a statistical model to estimate mean total, non-HDL and HDL cholesterol by country, year, sex and age using all of the available data. The model is described in detail in a statistical paper and related substantive papers<sup>8,32,33,48</sup>; the computer code is available at <http://www.ncdrisc.org/>. In summary, we organized countries into 21 regions, mainly based on geography and national income; these regions were further aggregated into 9 ‘super-regions’ (Supplementary Table 2). The model had a hierarchical structure in which estimates for each country and year were informed by its own data, if available, and by data from other years in the same country and from other countries, especially countries in the same region or super-region with data for similar time periods. The extent to which estimates for each country-year are influenced by data from other years and other countries depends on whether the country has data, the sample size of data, whether or not they are national, and the within-country and within-region data variability. The model incorporated nonlinear time trends comprising linear terms and a second-order random walk. The age association of blood lipids was modelled using a cubic spline to allow nonlinear age patterns, which might vary across countries. The model accounted for the possibility that blood lipids in subnational and community samples might systematically differ from nationally representative ones; and/or have larger variation. These features were implemented by including data-driven fixed-effect and random-effect terms for subnational and community data. The fixed effects adjust for systematic differences between subnational or community studies and national studies. The random effects allow national data to have larger influence on the estimates than subnational or community data with similar sample sizes. The model also accounted for rural–urban differences in blood lipids, through the use of data-driven fixed effects for rural-only and urban-only studies. These rural and urban effects were weighted by the difference between study-level and country-level urbanization in the year in which the study was done. The proportion of the national population living in urban areas was also included as a predictor (covariate) in the model. The model for mean non-HDL and HDL cholesterol also used age-standardized mean total cholesterol as a covariate.

We fitted the statistical model with the Markov chain Monte Carlo (MCMC) algorithm, and obtained 5,000 post-burn-in samples from the posterior distribution of model parameters, which were in turn used to obtain the posterior distributions of mean total, non-HDL and HDL cholesterol. We calculated average change in mean total, HDL and non-HDL cholesterol across the 39 years of analysis (reported as change

## Article

per decade). Age-standardized estimates were generated by taking weighted averages of age–sex-specific estimates, using the WHO standard population. Estimates for regions and the world were calculated as population-weighted averages of the constituent country estimates by age group and sex. The reported credible intervals represent the 2.5–97.5th percentiles of the posterior distributions. We also report the posterior probability that an estimated increase or decrease represents a truly increasing or decreasing trend as opposed to a chance observation. We performed all analyses by sex, because blood lipids levels and trends are different in men and women.

### Validation of statistical model

We tested how well our statistical model predicts missing data, known as external predictive validity, in two different tests. In the first test, we held out all data from 10% of countries with data (that is, created the appearance of countries with no data where we actually had data). The countries for which the data were withheld were randomly selected from the following three groups: data rich (5 or more data sources, with at least one data source after the year 2000), data poor (1 data source) and average data availability (2–4 data sources). In the second test, we assessed other patterns of missing data by holding out 10% of our data sources, again from a mix of data-rich, data-poor and average-data countries, as defined above. For a given country, we either held out a random half of the data of a country or all of the 2000–2018 data of the country to determine, respectively, how well we filled in the gaps for countries with intermittent data and how well we estimated in countries without recent data. In both tests, we then fitted the model to the remaining 90% of the countries (test 1) or data sources (test 2) and made estimates of the held-out observations. We repeated each test five times, holding out a different subset of data in each repetition. In both tests, we calculated the differences between the held-out data and the estimates. We also calculated the 95% credible intervals of the estimates; in a model with good external predictive validity, 95% of held-out values would be included in the 95% credible intervals.

Our statistical model performed well in the external validation tests, that is, in estimating mean cholesterol when data were missing. The estimates of mean total, non-HDL and HDL cholesterol were unbiased, as evidenced with median errors that were very close to zero globally for every outcome and test, and less than  $\pm 0.30 \text{ mmol l}^{-1}$  in every subset of withheld data except for women in the high-income Asia–Pacific region in test 1 for non-HDL cholesterol (median error  $0.47 \text{ mmol l}^{-1}$ ) and men in south Asia in test 2 for non-HDL cholesterol (median error  $-0.33 \text{ mmol l}^{-1}$ ) (Supplementary Table 5). The 95% credible intervals of estimated means covered 83–92% and 75–83% of true data globally in the first and second tests, respectively. In subsets, coverage ranged from 47% to 100%, but was mostly greater than 75%, with coverage generally lower in test 2 than test 1. Median absolute errors ranged from 0.07 to  $0.23 \text{ mmol l}^{-1}$  globally for different outcomes and sexes, and were no more than  $0.45 \text{ mmol l}^{-1}$  in all subsets of withheld data, except for women in the high-income Asia–Pacific region for non-HDL cholesterol in test 1 (median absolute error  $0.47 \text{ mmol l}^{-1}$ ).

### Calculation of the number of deaths attributable to high cholesterol

We estimated the number of deaths from IHD and ischaemic stroke attributable to high non-HDL cholesterol. For each country, year, sex and age group, we first calculated the population attributable fractions—that is, the proportion of deaths from IHD and ischaemic stroke that would have been prevented if non-HDL cholesterol levels were at an optimal level (defined as a mean of  $1.8\text{--}2.2 \text{ mmol l}^{-1}$ ) in the population<sup>6,49</sup>. For these calculations, we used age-specific relative risks from meta-analyses of prospective cohort studies<sup>4,5,50</sup>. The number of IHD and ischaemic stroke deaths attributable to high non-HDL cholesterol was calculated for each country–year–age–sex group by multiplying the cause-specific population attributable fractions by the cause-specific

deaths from the Global Burden of Disease study in 1990 and 2017 (the earliest and latest years with cause-specific mortality data).

### Strengths and limitations

The strengths of our study include its scope in making consistent and comparable estimates of trends in blood cholesterol and its cardiovascular disease mortality burden, over almost four decades for all of the countries in the world, including global estimates of non-HDL and HDL cholesterol. We used a large amount of population-based data, which came from countries in which 92% of the global adult population lives. We used only data from studies that had measured blood lipids to avoid bias in self-reported data. Data were analysed according to a consistent protocol, and the characteristics and quality of data from each country were rigorously verified through repeated checks by NCD-RisC members. We pooled data using a statistical model that took into account the epidemiological features of cholesterol, including nonlinear time trends and age associations. Our statistical model used all available data while giving more weight to national data than to subnational and community sources.

Similar to all global analyses, our study is affected by some limitations. Despite our extensive efforts to identify and access worldwide population-based data, some countries had no or few data sources, especially those in sub-Saharan Africa, the Caribbean, central Asia and Melanesia. Estimates for these countries relied mostly or entirely on the statistical model, which shares information across countries and regions through its hierarchy. Data scarcity is reflected in wider uncertainty intervals of our estimates for these countries and regions, highlighting the need for national NCD-oriented surveillance. The distribution of lipids measured in a population using a portable device, which was used in 10% of our studies, may be truncated and may therefore affect the population mean. To overcome this issue, we developed conversion regressions to adjust mean cholesterol levels measured using a portable device to the levels expected in laboratory measurements; the conversion regressions used for this purpose had good predictive accuracy. Although most studies had measured cholesterol in serum samples, around 7% had used plasma samples. As cholesterol measured in plasma and serum samples differ<sup>51</sup> by only about 3%, adjusting for plasma-serum differences would have little effect on our results, as seen in a previous analysis<sup>14</sup>. Although methods to measure total and HDL cholesterol have evolved over time, since the 1950s there have been systematic efforts to standardize lipid measurements that have resulted in increased comparability between different methods. In our analysis, 90% of studies measured lipids in a laboratory; of these studies more than 60% for total cholesterol and more than 70% for HDL cholesterol participated in a lipid standardization programme or quality control scheme. We did not analyse emerging lipid markers such as apolipoprotein B and apolipoprotein A-I, because they are neither commonly measured in population-based health surveys, nor routinely used in clinical practice<sup>36</sup>.

### Comparison with other studies

There are no global analyses on trends in lipid fractions for comparison with our results. Our findings for total cholesterol were largely consistent with the only other previous analysis<sup>7</sup>, but we estimated a larger decrease in mean total cholesterol in high-income western countries and central Europe, and a larger increase in southeast Asia, because we had an additional decade of data compared with the earlier global analysis. Therefore, although the highest mean total cholesterol levels reported previously<sup>7</sup>, for 2008, were still in high-income western countries, we estimated that in 2018 total cholesterol was equally high or higher in southeast Asia. Our findings on mean total cholesterol trends are also largely consistent with previous multi- and single-country reports<sup>14,15,17–21,52–73</sup>. Differences from previous studies—for example, in Italy<sup>61</sup>, Lithuania<sup>63</sup>, the Netherlands<sup>65</sup>, Russian Federation<sup>69</sup> and in some countries that participated in the MONICA Project<sup>52</sup>—mostly arise

because our study covered a longer period and used a larger number of data sources. Studies<sup>15,18,54,63,66,70,74–77</sup> that have reported trends in lipid fractions for a period longer than 15 years have found changes in non-HDL cholesterol (or in LDL cholesterol for some studies) that were consistent with our results.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

Estimates of mean total, non-HDL and HDL cholesterol by country, year and sex are available at <http://www.ncdrisc.org/>. Input data from publicly available sources can also be downloaded from <http://www.ncdrisc.org/>. For other data sources, contact information for data providers can be obtained from <http://www.ncdrisc.org/>.

## Code availability

The computer code for the Bayesian hierarchical model used in this work is available at <http://www.ncdrisc.org/>.

34. Martin, S. S. et al. Friedewald-estimated versus directly measured low-density lipoprotein cholesterol and treatment implications. *J. Am. Coll. Cardiol.* **62**, 732–739 (2013).
35. Cui, Y. et al. Non-high-density lipoprotein cholesterol level as a predictor of cardiovascular disease mortality. *Arch. Intern. Med.* **161**, 1413–1419 (2001).
36. Grundy, S. M. et al. 2018 AHA/ACC/AACVPR/APA/ABC/ACPM/ADA/AGS/APhA/NLA/PCNA guideline on the management of blood cholesterol: a report of the American College of Cardiology/American Heart Association Task Force on clinical practice guidelines. *Circulation* **139**, e1082–e1143 (2019).
37. Mach, F. et al. 2019 ESC/EAS guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. *Eur. Heart J.* **41**, 111–188 (2020).
38. Expert Dyslipidemia Panel of the International Atherosclerosis Society. An International Atherosclerosis Society Position Paper: global recommendations for the management of dyslipidemia—full report. *J. Clin. Lipidol.* **8**, 29–60 (2014).
39. Nordestgaard, B. G. et al. Fasting is not routinely required for determination of a lipid profile: clinical and laboratory implications including flagging at desirable concentration cut-points—a joint consensus statement from the European Atherosclerosis Society and European Federation of Clinical Chemistry and Laboratory Medicine. *Eur. Heart J.* **37**, 1944–1958 (2016).
40. Bilen, O., Kamal, A. & Virani, S. S. Lipoprotein abnormalities in South Asians and its association with cardiovascular disease: current state and future directions. *World J. Cardiol.* **8**, 247–257 (2016).
41. Danaei, G. et al. National, regional, and global trends in systolic blood pressure since 1980: systematic analysis of health examination surveys and epidemiological studies with 786 country-years and 5·4 million participants. *Lancet* **377**, 568–577 (2011).
42. Danaei, G. et al. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants. *Lancet* **378**, 31–40 (2011).
43. Finucane, M. M. et al. National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9·1 million participants. *Lancet* **377**, 557–567 (2011).
44. Cholesterol Reference Method Laboratory Network. *Total Cholesterol Certification Protocol for Manufacturers*. <https://www.cdc.gov/labstandards/pdf/crmln/RevisedTCProtocolOct04.pdf> (Cholesterol Reference Method Laboratory Network, 2004).
45. Cholesterol Reference Method Laboratory Network. *HDL Cholesterol Certification Protocol for Manufacturers*. [https://www.cdc.gov/labstandards/pdf/crmln/HDL\\_Certification\\_Protocol-508.pdf](https://www.cdc.gov/labstandards/pdf/crmln/HDL_Certification_Protocol-508.pdf) (Cholesterol Reference Method Laboratory Network, 2018).
46. Cholesterol Reference Method Laboratory Network. *LDL Cholesterol Certification Protocol for Manufacturers*. [https://www.cdc.gov/labstandards/pdf/crmln/LDL\\_Certification\\_Protocol-508.pdf](https://www.cdc.gov/labstandards/pdf/crmln/LDL_Certification_Protocol-508.pdf) (Cholesterol Reference Method Laboratory Network, 2018).
47. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **6**, 461–464 (1978).
48. Finucane, M. M., Paciorek, C. J., Danaei, G. & Ezzati, M. Bayesian estimation of population-level trends in measures of health status. *Stat. Sci.* **29**, 18–25 (2014).
49. Ezzati, M., Lopez, A. D., Rodgers, A., Vander Hoorn, S. & Murray, C. J. Selected major risk factors and global and regional burden of disease. *Lancet* **360**, 1347–1360 (2002).
50. Singh, G. M. et al. The age-specific quantitative effects of metabolic risk factors on cardiovascular diseases and diabetes: a pooled analysis. *PLoS ONE* **8**, e65174 (2013).
51. National Cholesterol Education Program. *Recommendations on Lipoprotein Measurement*. From the Working Group on Lipoprotein Measurement. NIH Publication No. 95-3044 (National Institutes of Health, National Heart, Lung, and Blood Institute, 1995).
52. Evans, A. et al. Trends in coronary risk factors in the WHO MONICA project. *Int. J. Epidemiol.* **30**, S35–S40 (2001).
53. Bennett, S. A. & Magnus, P. Trends in cardiovascular risk factors in Australia. Results from the National Heart Foundation's Risk Factor Prevalence Study, 1980–1989. *Med. J. Aust.* **161**, 519–527 (1994).
54. Cifková, R. et al. Longitudinal trends in major cardiovascular risk factors in the Czech population between 1985 and 2007/8. Czech MONICA and Czech post-MONICA. *Atherosclerosis* **211**, 676–681 (2010).
55. Sun, J. Y. et al. The changing trend of serum total cholesterol in Beijing population aged 25–64 years during 1984–1999 (article in Chinese). *Zhonghua Nei Ke Za Zhi* **45**, 980–984 (2006).
56. Afzal, S., Tybjærg-Hansen, A., Jensen, G. B. & Nordestgaard, B. G. Change in body mass index associated with lowest mortality in Denmark, 1976–2013. *J. Am. Med. Assoc.* **315**, 1989–1996 (2016).
57. Ferrières, J. et al. Trends in plasma lipids, lipoproteins and dyslipidaemias in French adults, 1996–2007. *Arch. Cardiovasc. Dis.* **102**, 293–301 (2009).
58. Truthmann, J. et al. Changes in mean serum lipids among adults in Germany: results from National Health Surveys 1997–99 and 2008–11. *BMC Public Health* **16**, 240 (2016).
59. Sigfusson, N. et al. Decline in ischaemic heart disease in Iceland and change in risk factor levels. *Br. Med. J.* **302**, 1371–1375 (1991).
60. Gupta, R. et al. Twenty-year trends in cardiovascular risk factors in India and influence of educational status. *Eur. J. Prev. Cardiol.* **19**, 1258–1271 (2012).
61. Giampaoli, S. et al. Cardiovascular health in Italy. Ten-year surveillance of cardiovascular diseases and risk factors: Osservatorio Epidemiologico Cardiovascolare/Health Examination Survey 1998–2012. *Eur. J. Prev. Cardiol.* **22**, 9–37 (2015).
62. Iso, H. Changes in coronary heart disease risk among Japanese. *Circulation* **118**, 2725–2729 (2008).
63. Luksiene, D. et al. Trends in prevalence of dyslipidaemias and the risk of mortality in Lithuanian urban population aged 45–64 in relation to the presence of the dyslipidaemias and the other cardiovascular risk factors. *PLoS ONE* **9**, e100158 (2014).
64. Uusitalo, U. et al. Fall in total cholesterol concentration over five years in association with changes in fatty acid composition of cooking oil in Mauritius: cross sectional survey. *Br. Med. J.* **313**, 1044–1046 (1996).
65. Koopman, C. et al. Trends in risk factors for coronary heart disease in the Netherlands. *BMC Public Health* **16**, 835 (2016).
66. Metcalf, P. et al. Trends in major cardiovascular risk factors in Auckland, New Zealand: 1982 to 2002–2003. *N. Z. Med. J.* **119**, U2308 (2006).
67. Sharashova, E., Wilsgaard, T. & Brenn, T. Resting heart rate on the decline: the Tromsø Study 1986–2007. *Int. J. Epidemiol.* **44**, 1007–1017 (2015).
68. Pajak, A. et al. Changes over time in blood lipids and their correlates in Polish rural and urban populations: the Poland–United States Collaborative Study in cardiopulmonary disease epidemiology. *Ann. Epidemiol.* **7**, 115–124 (1997).
69. Vlasoff, T. et al. Ten year trends in chronic disease risk factors in the Republic of Karelia, Russia. *Eur. J. Public Health* **18**, 666–673 (2008).
70. Bovet, P. et al. Divergent fifteen-year trends in traditional and cardiometabolic risk factors of cardiovascular diseases in the Seychelles. *Cardiovasc. Diabetol.* **8**, 34 (2009).
71. Serra-Majem, L. et al. Trends in blood lipids and fat soluble vitamins in Catalonia, Spain (1992–2003). *Public Health Nutr.* **10**, 1379–1388 (2007).
72. Berg, C. M. et al. Trends in blood lipid levels, blood pressure, alcohol and smoking habits from 1985 to 2002: results from INTERGENE and GOT-MONICA. *Eur. J. Cardiovasc. Prev. Rehabil.* **12**, 115–125 (2005).
73. Wietlisbach, V., Paccaud, F., Rickenbach, M. & Gutzwiller, F. Trends in cardiovascular risk factors (1984–1993) in a Swiss region: results of three population surveys. *Prev. Med.* **26**, 523–533 (1997).
74. Leiviskä, J. et al. What have we learnt about high-density lipoprotein cholesterol measurements during 32 years? Experiences in Finland 1980–2012. *Clin. Chim. Acta* **415**, 118–123 (2013).
75. Yokoyama, S. et al. High-density lipoprotein levels have markedly increased over the past twenty years in Japan. *J. Atheroscler. Thromb.* **21**, 151–160 (2014).
76. Hardoon, S. L. et al. Assessing the impact of medication use on trends in major coronary risk factors in older British men: a cohort study. *Eur. J. Cardiovasc. Prev. Rehabil.* **17**, 502–508 (2010).
77. Carroll, M. D. et al. Trends in serum lipids and lipoproteins of adults, 1960–2002. *J. Am. Med. Assoc.* **294**, 1773–1781 (2005).

**Acknowledgements** This study was funded by a Wellcome Trust (Biomedical Resource & Multi-User Equipment grant 01506/Z/13/Z) and the British Heart Foundation (Centre of Research Excellence grant RE/18/4/34215). C.T. was supported by a Wellcome Trust Research Training Fellowship (203616/Z/16/Z). The authors alone are responsible for the views expressed in this Article and they do not necessarily represent the views, decisions, or policies of the institutions with which they are affiliated.

**Author contributions** M.E. and G.D. designed the study and oversaw research. C.T., B.Z., H.B. and R.C.L. led the data collection. The other authors contributed to study design; and collected, reanalysed, checked and pooled data. C.T. analysed pooled data and prepared results. C.T., E.G. and M.E. wrote the first draft of the manuscript with input from the other authors.

**Competing interests** M.E. reports a charitable grant from the AstraZeneca Young Health Programme, and personal fees from Prudential, Scor and Third Bridge, outside the submitted work. The other authors declare no competing interests.

## Additional information

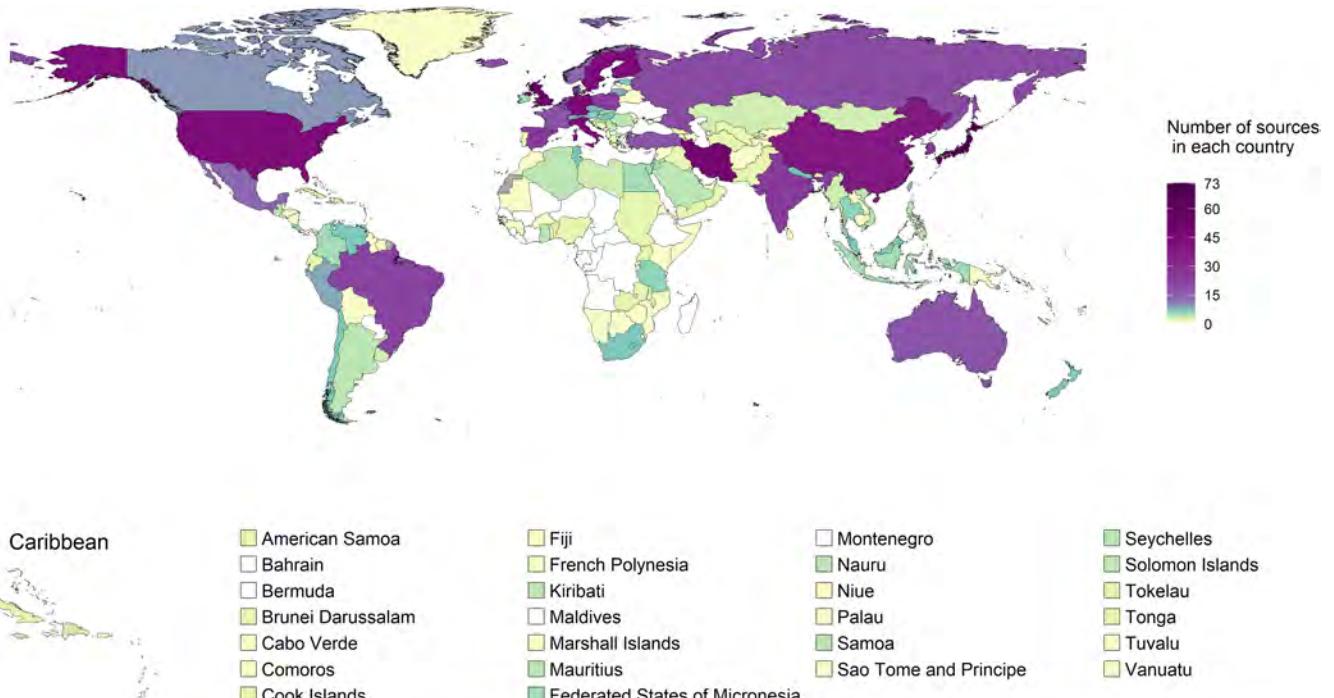
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2338-1>.

**Correspondence and requests for materials** should be addressed to M.E.

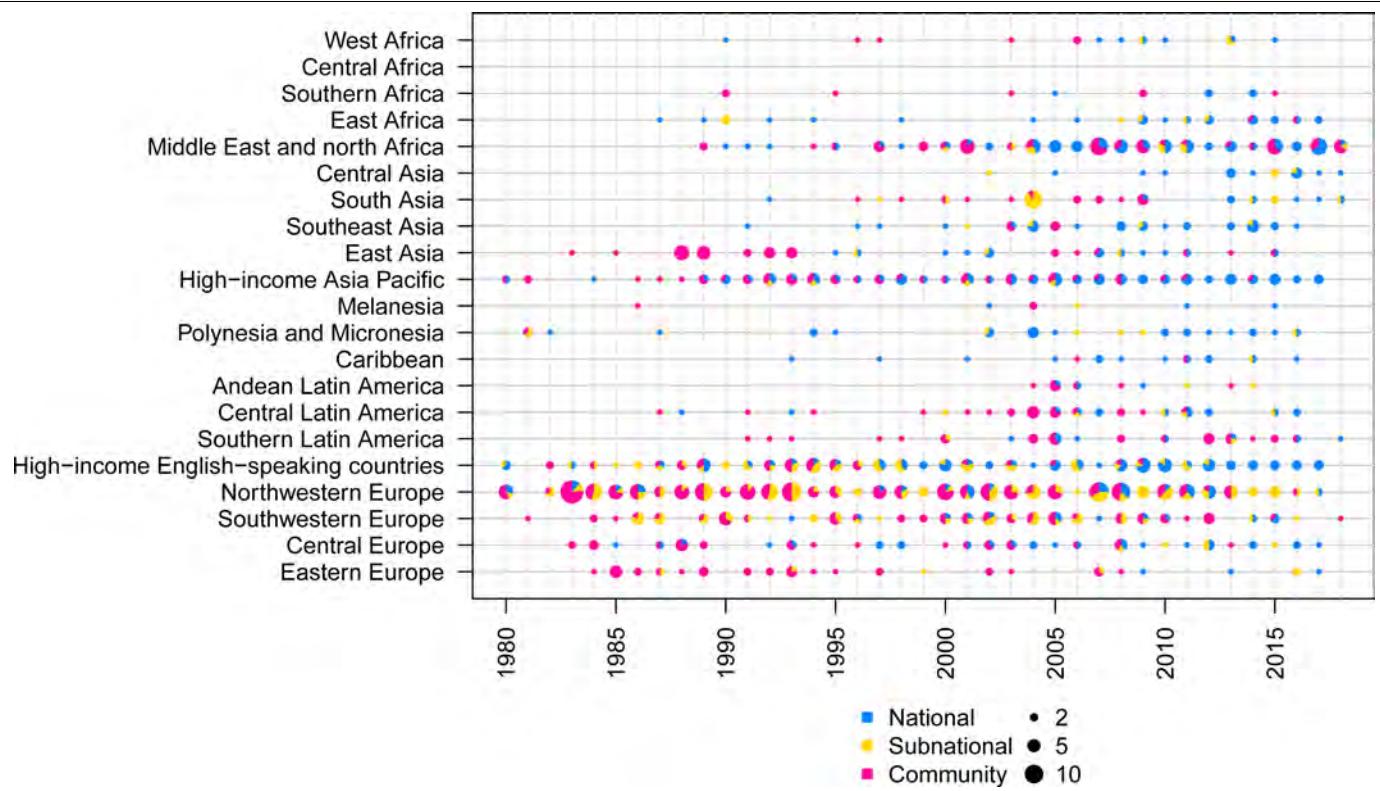
**Peer review information** *Nature* thanks Frank Hu and Pekka Jousilahti for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

# Article

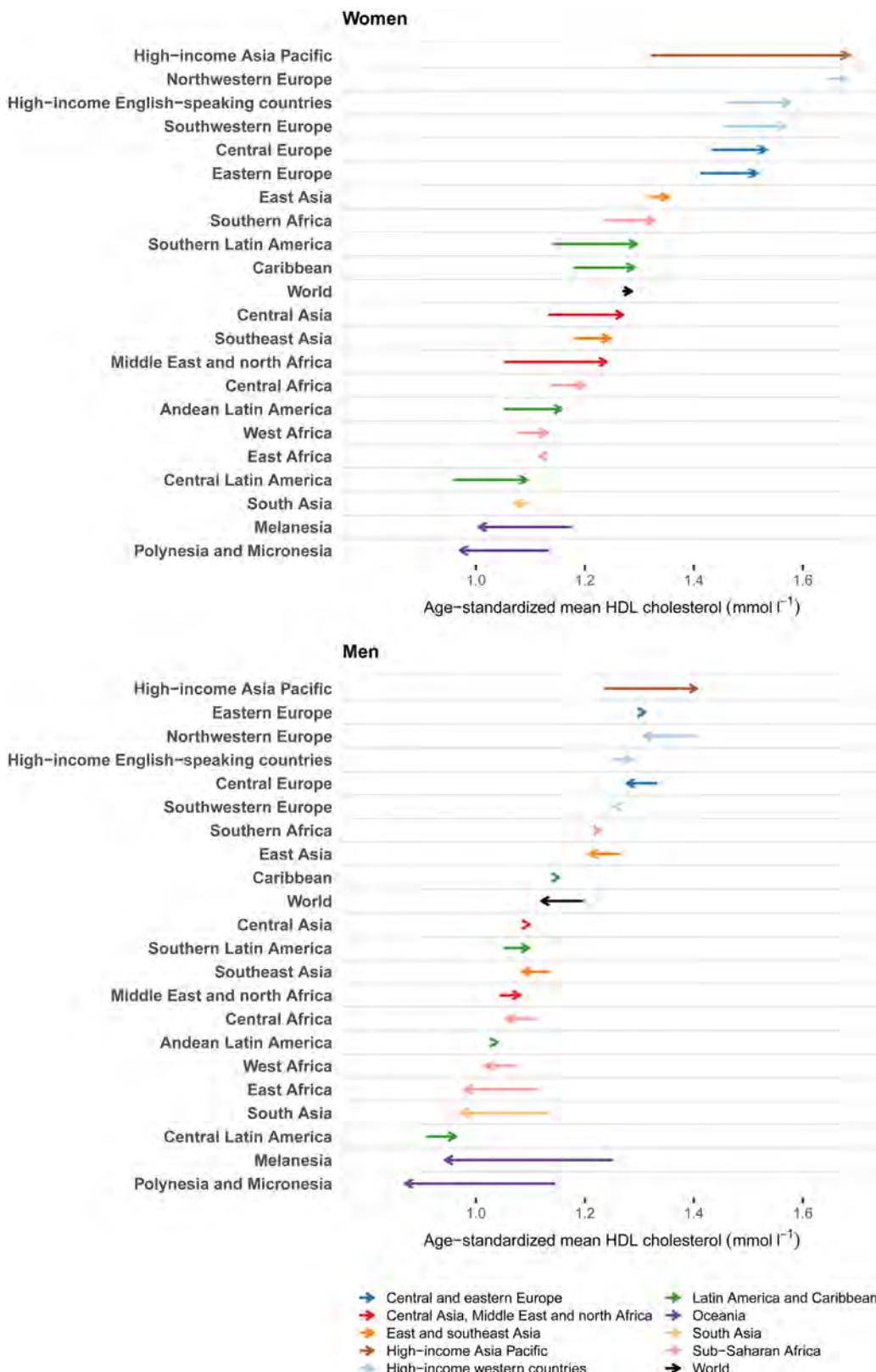


**Extended Data Fig. 1 | Number of data sources by country.** The colour indicates the number of data sources for each country used in the analysis. Countries and territories that were not included in the analysis are coloured in grey.

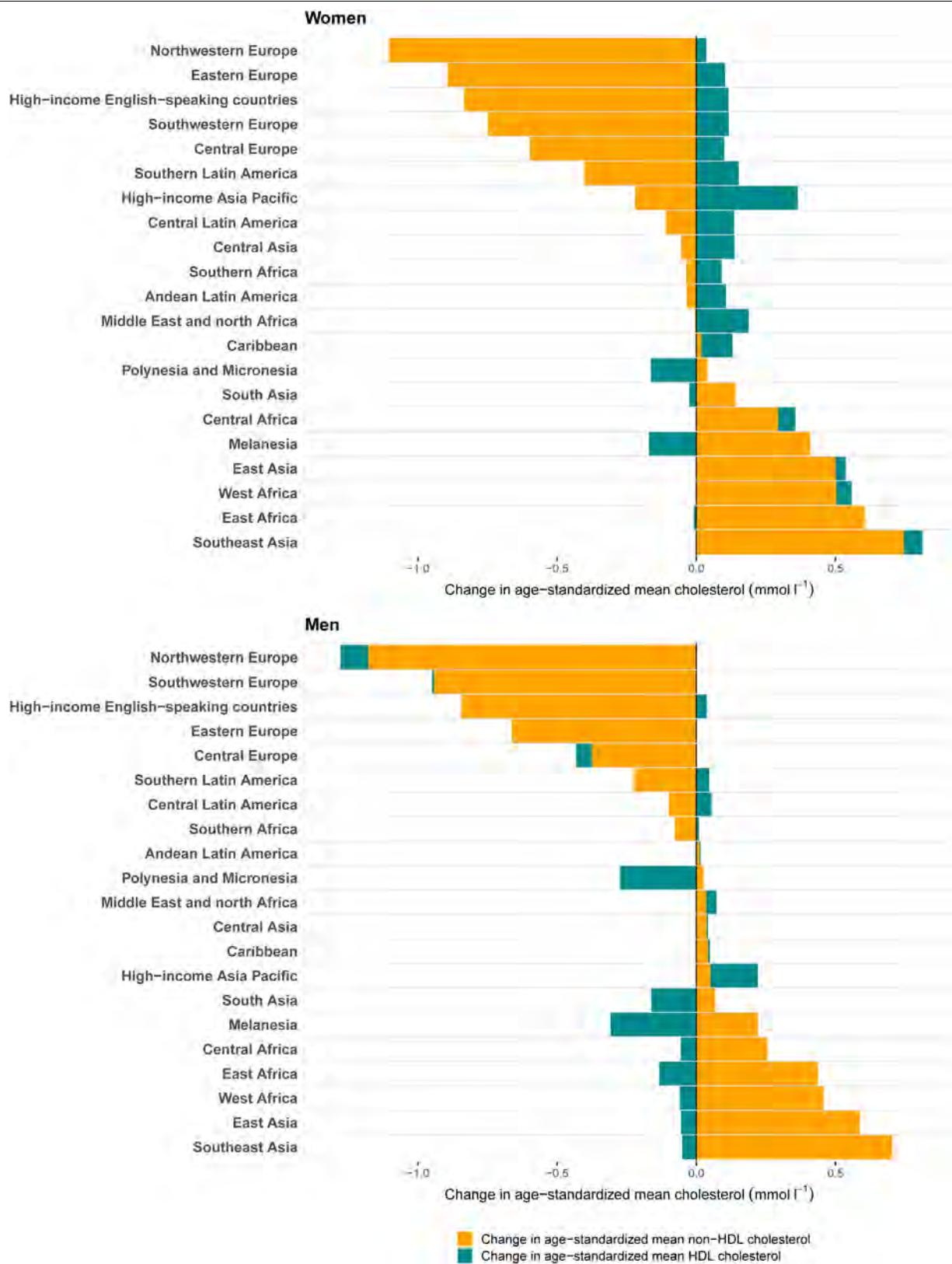


**Extended Data Fig. 2 | Number of data sources by region and year.** The size of each circle shows the number of data sources for each region and year, and the colours indicate the relative size of national, subnational and community data sources.

# Article

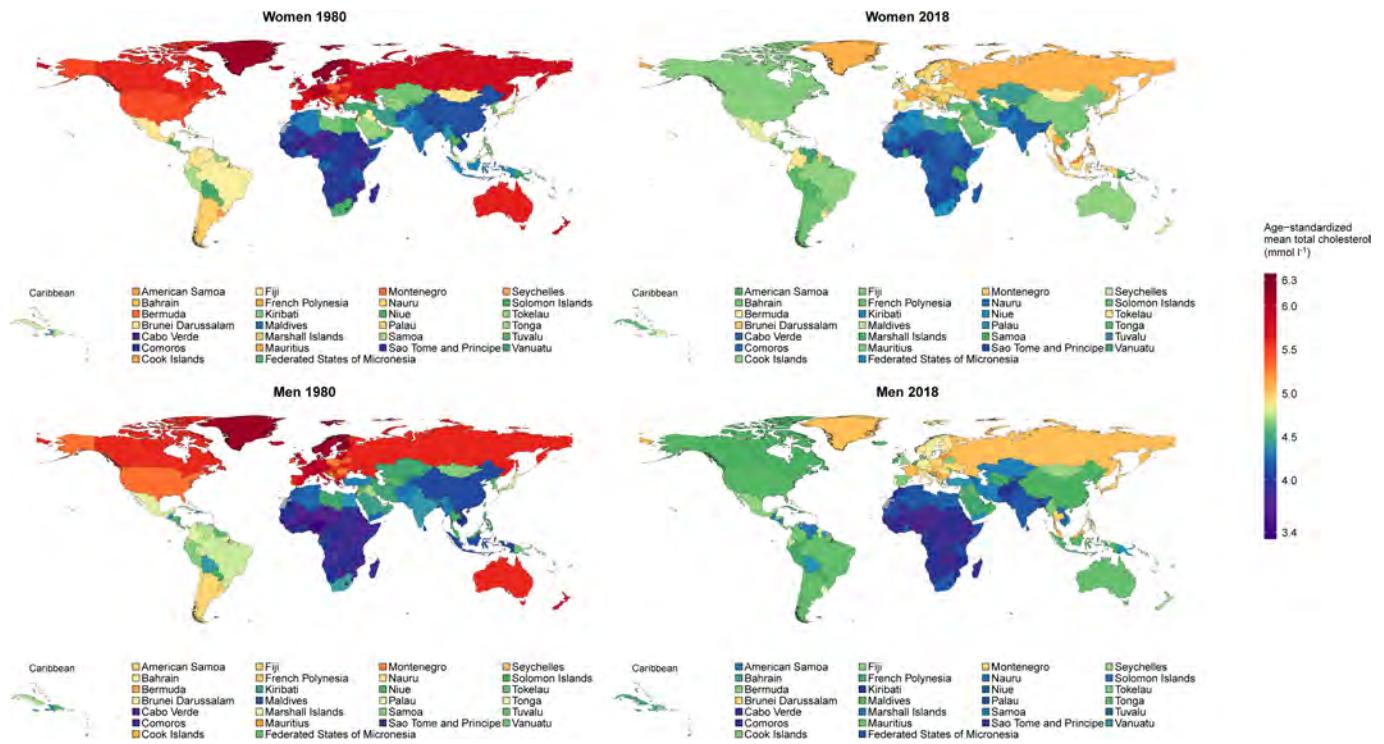


**Extended Data Fig. 3 | Change in age-standardized mean HDL cholesterol between 1980 and 2018 by region for women and men.** The start of the arrow shows the level in 1980 and the head shows the level in 2018. One  $\text{mmol L}^{-1}$  is equivalent to  $38.61 \text{ mg dL}^{-1}$ .

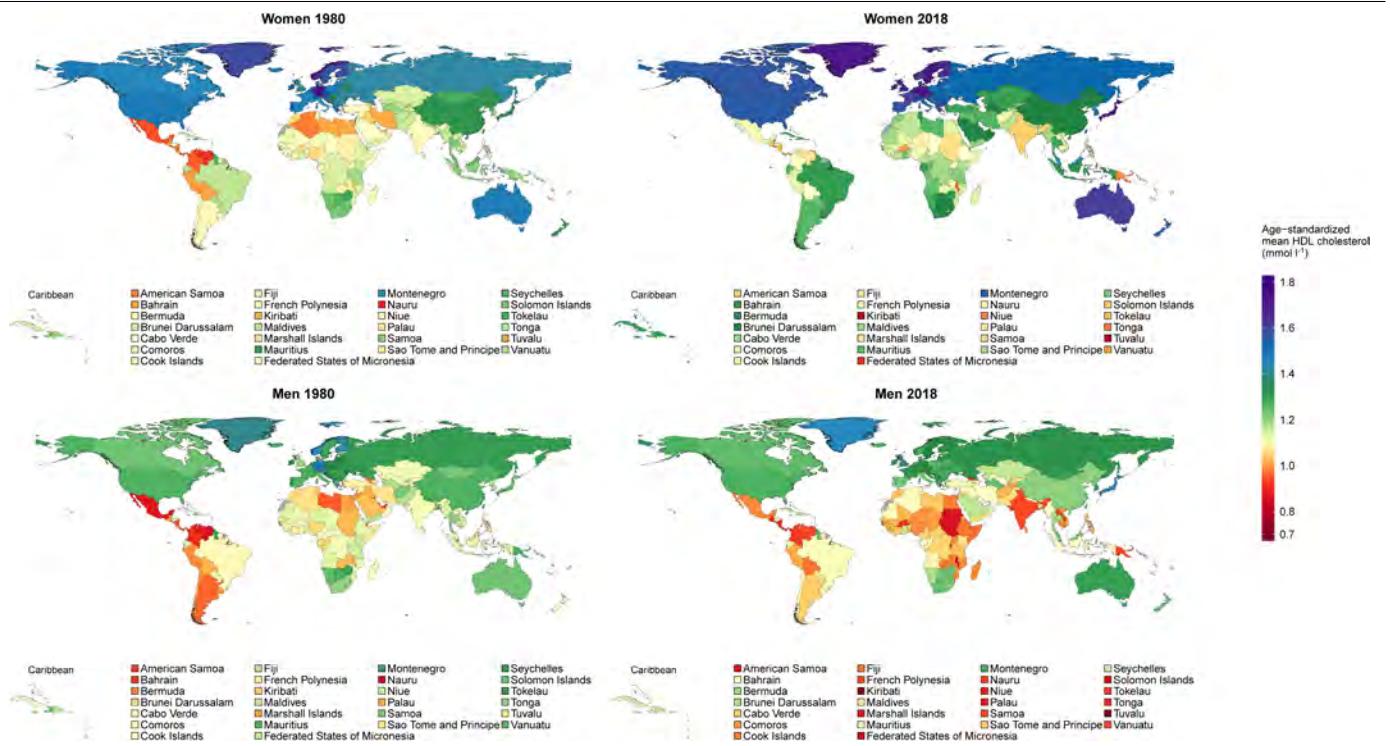


**Extended Data Fig. 4 | Change in age-standardized mean HDL and non-HDL cholesterol between 1980 and 2018 by region for women and men.** One  $\text{mmol L}^{-1}$  is equivalent to  $38.61 \text{ mg dL}^{-1}$ .

# Article

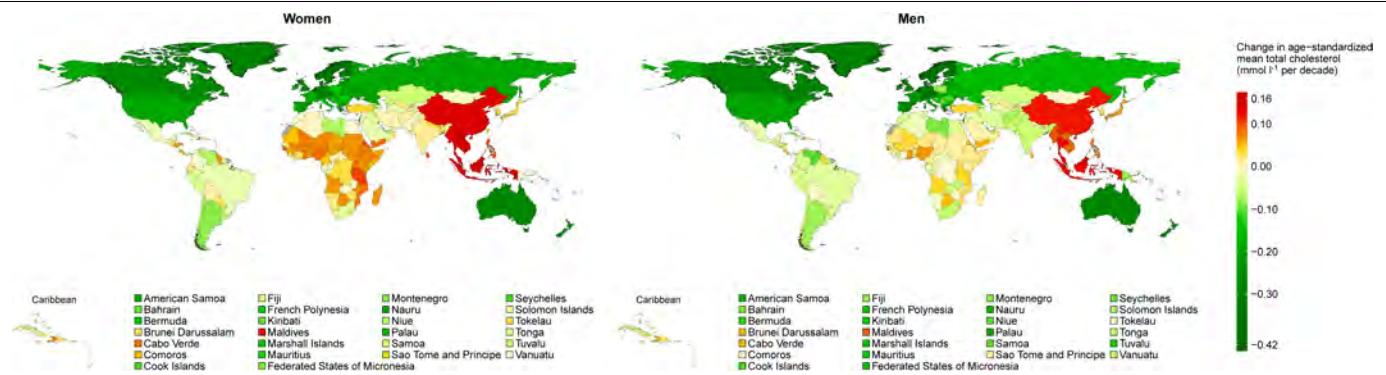


**Extended Data Fig. 5 | Age-standardized mean total cholesterol by country in 1980 and 2018 for women and men.** One  $\text{mmol l}^{-1}$  is equivalent to 38.61  $\text{mg dl}^{-1}$ .

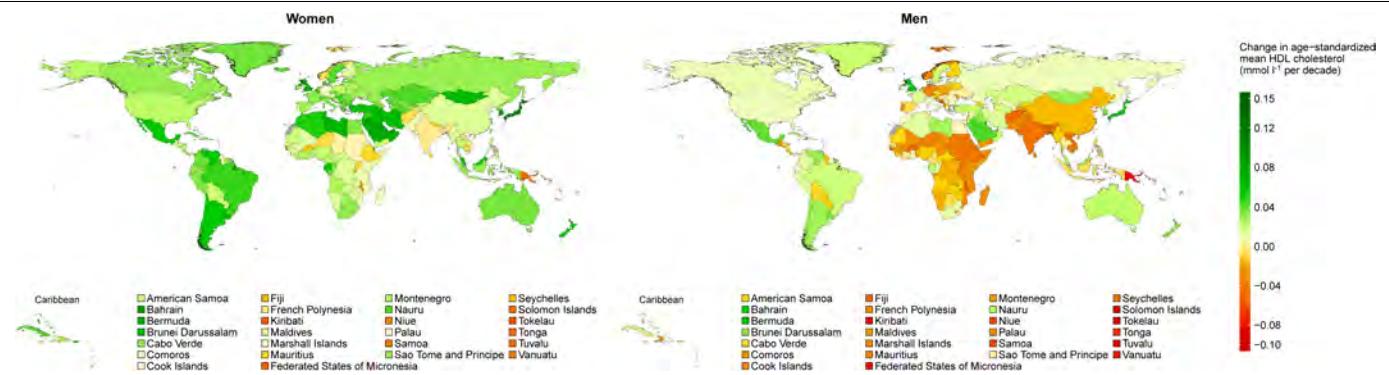


**Extended Data Fig. 6 | Age-standardized mean HDL cholesterol by country in 1980 and 2018 for women and men.** One  $\text{mmol l}^{-1}$  is equivalent to  $38.61 \text{ mg dl}^{-1}$ .

# Article

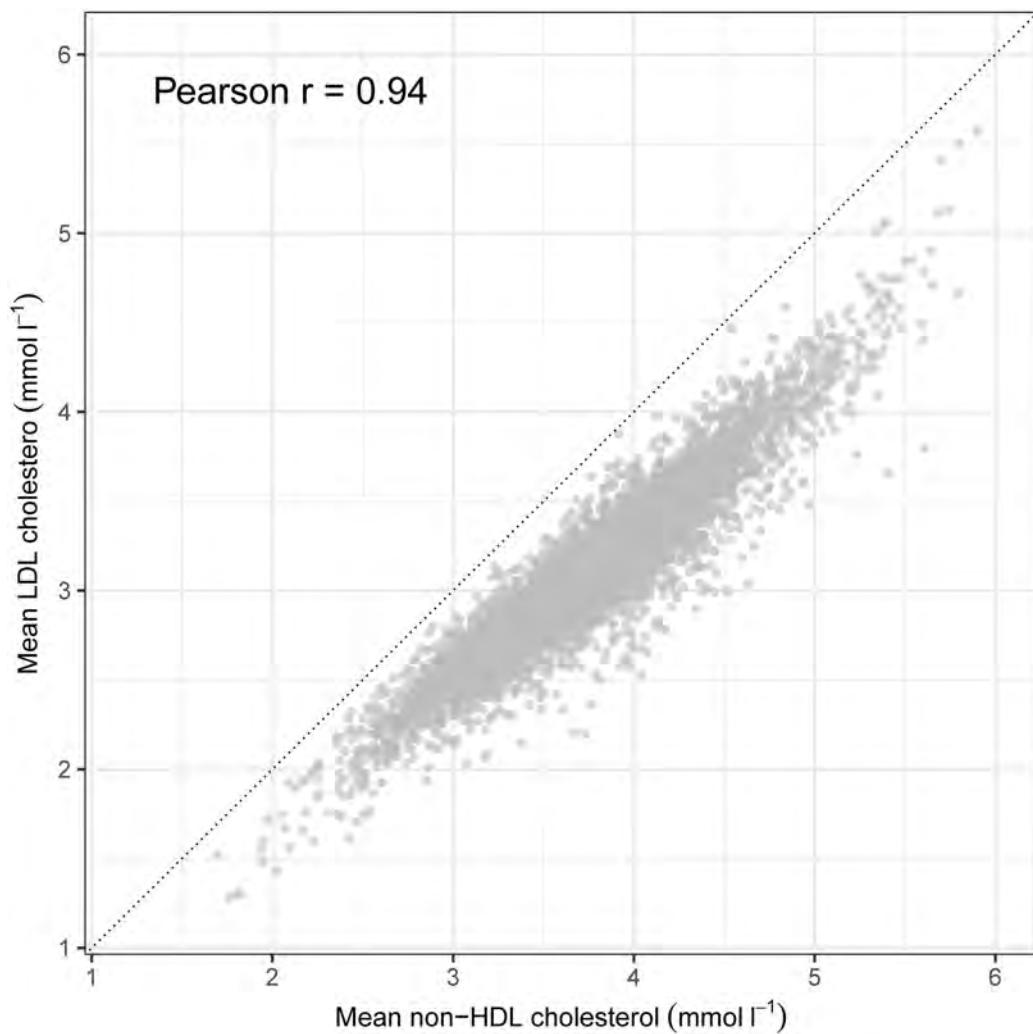


**Extended Data Fig. 7 | Change per decade in age-standardized mean total cholesterol by country for women and men.** One mmol l<sup>-1</sup> is equivalent to 38.61 mg dl<sup>-1</sup>.



**Extended Data Fig. 8 | Change per decade in age-standardized mean HDL cholesterol by country for women and men.** One mmol L<sup>-1</sup> is equivalent to 38.61 mg dL<sup>-1</sup>.

## Article



**Extended Data Fig. 9 | The association between mean LDL and non-HDL cholesterol in studies that measured lipids in a laboratory that had data for both variables.** Each data point is one study–age–sex group ( $n = 6,864$ ). One  $\text{mmol l}^{-1}$  is equivalent to  $38.61 \text{ mg dl}^{-1}$ .

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

## Software and code

Policy information about [availability of computer code](#)

Data collection	Processing of secondary data was conducted using the statistical software R (version 3.6.0).
Data analysis	All analyses were conducted using the statistical software R (version 3.6.0). The code for estimation of mean risk factor trends is available at <a href="http://www.ncdrisc.org">www.ncdrisc.org</a> .

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

This is a data-pooling study that brings together more than 1000 disparate data sources and uses a Bayesian hierarchical model to estimate population risk factor trends. Estimates of mean total, non-HDL and HDL cholesterol by country, year, and sex will be available from [www.ncdrisc.org](http://www.ncdrisc.org) upon the publication of the paper. Some of the input data sources are publicly available, for which we will add links in the final version of the paper. Others are the property of specific research groups and agencies, for which we will provide contact information.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](http://nature.com/documents/nr-reporting-summary-flat.pdf)

## Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We pooled and re-analysed population-based data that had measured blood lipids in adults to estimate trends in mean total, non-HDL and HDL cholesterol from 1980 to 2018 for 200 countries and territories, using a Bayesian hierarchical model.
Research sample	We pooled data from 1,127 population-based studies of blood lipids conducted in 161 countries, with measurement of blood lipids in over 102 million adults aged 18 years and older. Studies were representative of a national, subnational or community population.
Sampling strategy	We included data collected using a probabilistic sampling method with a defined sampling frame. We therefore included studies with simple random and complex survey designs but excluded convenience samples.
Data collection	We used data on measured blood lipids to calculate mean total, non-HDL and HDL cholesterol. We excluded self-reported data.
Timing	We pooled data collected from 1980 to 2018. We also included national studies for the 3 years prior to 1980 (n=1), assigning them to 1980, so that they can inform the estimates in countries with slightly earlier national data.
Data exclusions	<p>We excluded all data sources that included only hypercholesterolemia or dyslipidaemia diagnosis history or medication status without measurement of cholesterol levels. We also excluded data sources on population subgroups whose lipid profile may differ systematically from the general population, including:</p> <ul style="list-style-type: none"> <li>• studies that had included or excluded people based on their health status or cardiovascular risk;</li> <li>• studies whose participants were only ethnic minorities;</li> <li>• specific educational, occupational, or socioeconomic subgroups, with the exception noted below;</li> <li>• those recruited through health facilities, with the exception noted below.</li> </ul> <p>We used school-based data in countries, and in age-sex groups, where secondary school enrollment was 70% or higher. We used data whose sampling frame was health insurance schemes in countries where at least 80% of the population were insured. Finally, we used data collected through general practice and primary care systems in high-income and central European countries with universal insurance, because contact with the primary care systems tends to be as good as or better than response rates for population-based surveys. Our exclusion criteria were established at the initiation of the study to ensure all data were representative.</p>
Non-participation	Our inclusion/exclusion criteria were designed to ensure participants of the surveys included were representative of the general population from which each sample was drawn.
Randomization	Our study is descriptive, and we did not carry out experiments.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Article

# The nature of Neanderthal introgression revealed by 27,566 Icelandic genomes

<https://doi.org/10.1038/s41586-020-2225-9>

Received: 4 July 2019

Accepted: 21 February 2020

Published online: 22 April 2020

 Check for updates

Laurits Skov<sup>1,2,7</sup>✉, Moisès Coll Macià<sup>1,7</sup>, Garðar Sveinbjörnsson<sup>3</sup>, Fabrizio Mafessoni<sup>2</sup>, Elise A. Lucotte<sup>1</sup>, Margret S. Einarsdóttir<sup>3</sup>, Hakon Jonsson<sup>3</sup>, Bjarni Halldorsson<sup>3,4</sup>, Daniel F. Gudbjartsson<sup>3</sup>, Agnar Helgason<sup>3,5</sup>, Mikkel Heide Schierup<sup>1</sup>✉ & Kari Stefansson<sup>3,6</sup>✉

Human evolutionary history is rich with the interbreeding of divergent populations. Most humans outside of Africa trace about 2% of their genomes to admixture from Neanderthals, which occurred 50–60 thousand years ago<sup>1</sup>. Here we examine the effect of this event using 14.4 million putative archaic chromosome fragments that were detected in fully phased whole-genome sequences from 27,566 Icelanders, corresponding to a range of 56,388–112,709 unique archaic fragments that cover 38.0–48.2% of the callable genome. On the basis of the similarity with known archaic genomes, we assign 84.5% of fragments to an Altai or Vindija Neanderthal origin and 3.3% to Denisovan origin; 12.2% of fragments are of unknown origin. We find that Icelanders have more Denisovan-like fragments than expected through incomplete lineage sorting. This is best explained by Denisovan gene flow, either into ancestors of the introgressing Neanderthals or directly into humans. A within-individual, paired comparison of archaic fragments with syntenic non-archaic fragments revealed that, although the overall rate of mutation was similar in humans and Neanderthals during the 500 thousand years that their lineages were separate, there were differences in the relative frequencies of mutation types—perhaps due to different generation intervals for males and females. Finally, we assessed 271 phenotypes, report 5 associations driven by variants in archaic fragments and show that the majority of previously reported associations are better explained by non-archaic variants.

Eurasians are estimated to trace around 2% of their genomes to introgression from archaic humans<sup>2</sup> related to the Vindija Neanderthal<sup>1</sup>. In Europeans, archaic ancestry is consistent with contributions by a single population of Neanderthals<sup>3</sup> through one or more pulses<sup>4–7</sup>. As many archaic fragments still segregate in contemporary populations, it follows that archaic genomes can be partially mined from whole-genome sequence data from Europeans. Once such archaic fragments are identified, it is possible to determine the effect that variants in these fragments have on phenotypic diversity in modern humans<sup>8–13</sup>.

Previous attempts to identify archaic fragments in contemporary humans have generally relied on existing archaic genomes<sup>14–16</sup>, which partially represent the haplotypic diversity of the archaic populations. Moreover, the introgressing Neanderthals were probably from populations that were separated both in time and space from the three archaic individuals who have so far been sequenced to high coverage—hereafter referred to as DAV (Denisovan, Altai Neanderthal and Vindija Neanderthal)<sup>1,17,18</sup>. Some archaic fragments in contemporary humans will be missed by methods that depend on DAV genomes. Even more would be missed if the genetic information of introgressing Neanderthals is not well represented by one of the DAV genomes, for example, owing to admixture between archaic groups<sup>18,19</sup>. Here we investigated the

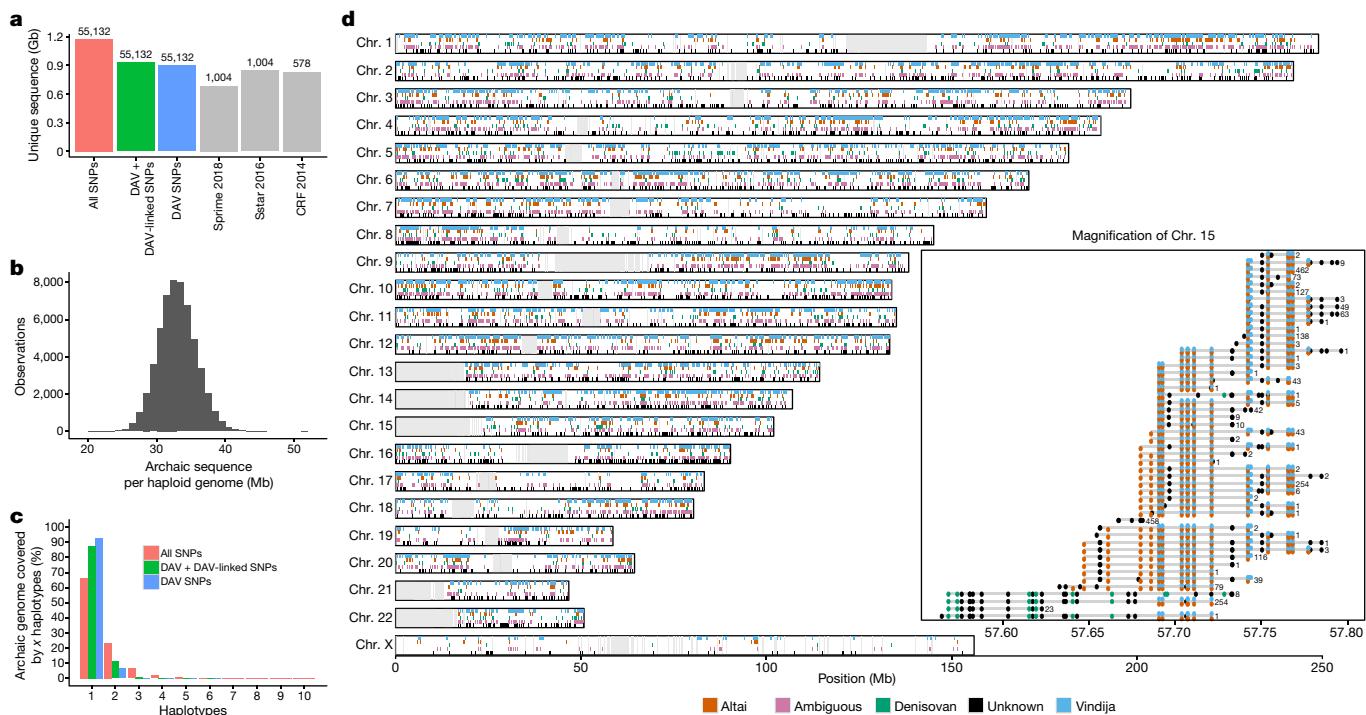
archaic diversity in contemporary Europeans and its relationship to the DAV genomes.

## A diverse set of introgressed fragments

We used a two-state hidden Markov model that is not conditioned on DAV genomes<sup>20</sup> to search for archaic fragments independently in 55,132 haploid genomes from 27,566 sequenced Icelanders that were sequenced to 30X and accurately phased using long-range phasing<sup>21</sup> (Supplementary Information 1, 2.1). The model detects fragments with a high density of derived alleles that were not found in an out-group population of 292 sub-Saharan Africans (hereafter referred to as Africans) from the 1000 Genomes Project<sup>22</sup>. For all subsequent analyses, we retained 14,422,595 fragments with a posterior probability of more than 90% of being archaic (Supplementary Information 2). This threshold was used because simulations indicate that our model has a false-positive rate of less than 4% while still recovering around 71.5% of all archaic sequences per haploid genome (Extended Data Fig. 1 and Supplementary Information 2.3, 3.2).

The archaic fragments carried derived alleles at 395,304 single-nucleotide polymorphisms (SNPs) that are absent in Africans (Supplementary Data 1). Of these variants, 147,925 were found in the

<sup>1</sup>Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark. <sup>2</sup>Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. <sup>3</sup>deCODE Genetics, Amgen, Reykjavik, Iceland. <sup>4</sup>School of Science and Engineering, Reykjavik University, Reykjavik, Iceland. <sup>5</sup>Department of Anthropology, University of Iceland, Reykjavik, Iceland. <sup>6</sup>Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland. <sup>7</sup>These authors contributed equally: Laurits Skov, Moisès Coll Macià. ✉e-mail: laurits\_skov@eva.mpg.de; mheide@birc.au.dk; kari.stefansson@decode.is



**Fig. 1 | Archaic diversity in 55,132 Icelandic haploid genomes.** **a**, The total length of the genome covered by archaic fragments using: all SNPs, DAV-linked and DAV SNPs, and only DAV SNPs. The amount of archaic sequences from previous studies based on the 1000 Genomes Project is also shown (Sprime<sup>3</sup>, Sstar<sup>14</sup> and CRF<sup>15</sup>). Sample sizes are shown above the bar. **b**, The distribution of the average amount of archaic sequences per haploid genome. **c**, Fraction of the introgressing genome covered by different numbers of haplotypes. We classify fragments into groups of haplotypes for which the mean difference between fragments in the cluster is less than 1 difference per 10 kb. We do this

for all SNPs, DAV-linked and DAV SNPs, and DAV SNPs. **d**, The genome-wide distribution of archaic haplotypes, coloured according to the DAV genome that they are closest to. Unknown, no derived alleles shared with a DAV genome; ambiguous, alleles that are equidistant to more than 1 DAV genome. The inset shows a 225-kb region on chromosome 15 that has a high fragment diversity. SNPs are coloured according to the DAV genome that they are found in; non-DAV and DAV-linked/unlinked SNPs are shown in black. The absolute frequency of each fragment is shown on the right. Some fragments appear to be Denisovan, while others are a mosaic of Denisovan and Neanderthal.

DAV genomes (hereafter DAV variants). These variants and the fragments that carried them are very likely to be archaic in origin<sup>23</sup>. The remaining 247,379 SNPs have derived alleles that were not observed in the DAV genomes. Some may have been present in introgressing Neanderthals, but are not captured by the three DAV genomes. Others could have arisen on archaic fragments in humans after the introgression. The remainder are non-archaic variants that were not observed in our African reference set. Archaic variants that are not observed in the DAV genomes are expected to be in strong linkage disequilibrium with DAV variants from the same archaic fragments. Of the 247,379 putative archaic variants, 47,193 were found to be linked ( $r^2 > 0.9$ ) to at least one nearby DAV variant (DAV-linked variants), whereas 145,153 were found on fragments with DAV variants but were not linked to them. The latter DAV-unlinked variants tend to have a lower frequency than DAV and DAV-linked variants, suggesting that many of the underlying mutations occurred on archaic fragments after Neanderthal introgression (Extended Data Fig. 2a). The remaining 55,033 variants were found on putative archaic fragments with no DAV variants and are hereafter called nonDAV variants.

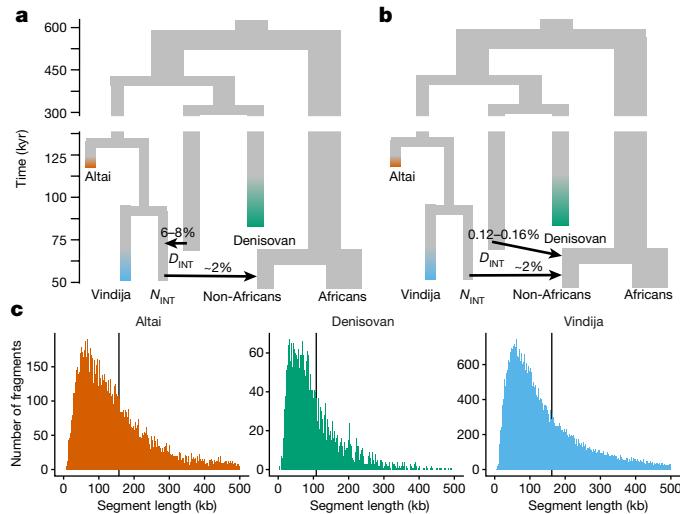
As introgressing archaic variants must be at least 50 thousand years old, they have a greater chance of being found in multiple populations than non-archaic variants of the same frequency. Both DAV-unlinked and nonDAV variants are found less often in European populations from the 1000 Genomes Project (55% and 64%, respectively) (Extended Data Fig. 2b) than DAV and DAV-linked variants (84% and 77%, respectively) after controlling for allele frequency. This is consistent with fewer DAV-unlinked and nonDAV variants having an archaic origin. DAV-unlinked variants are closer to the ends of putative archaic fragments than DAV variants: 12.3 kilobases (kb) compared with

26.8 kb median distance from the edge (Wilcoxon signed-ranked test,  $P < 2.2 \times 10^{-16}$ ) (Supplementary Fig. 2.5.3), indicating a slight increase in false positives at the ends of fragments.

The 14,422,595 candidate archaic fragments correspond to 112,709 unique fragments (based on start–end positions), have a combined length of 1,818 Gb and cover 1.179 Gb (48.2%) of the callable genome (2.445 Gb) (Fig. 1a), exceeding previous studies using Sprime (0.688 Gb)<sup>3</sup>, Sstar (0.846 Gb)<sup>24</sup> and CRF (0.834 Gb)<sup>15</sup> (Fig. 1a). After pruning fragments to start and end with DAV-linked and DAV variants, archaic coverage is estimated to be 0.962 Gb (39.3% of the callable genome, spanned by 59,124 unique fragments). The most conservative estimate of archaic coverage, based on ends defined by only DAV variants, is 0.929 Gb (38% of the callable genome, spanned by 56,388 unique fragments) (Fig. 1a and Supplementary Table 2.4.1). Even after analysing 55,132 haploid genomes, it may be deduced that not all archaic fragments have been detected in Icelanders (Supplementary Figs. 2.4.3, 2.5.4), and even more remain to be found in other populations.

An average of 261.6 archaic fragments, corresponding to around 33 Mb, were identified per haploid genome (Fig. 1b and Supplementary Information 2.3). Callable positions in the genome were, on average, covered by 743 archaic fragments, corresponding to an average frequency of 1.34% in the 55,132 haploid genomes analysed. These are probably underestimates, because our method misses archaic fragments that are short or not sufficiently divergent from non-archaic fragments. Simulations suggest a false-negative rate of 28.5%, suggesting that the frequency of archaic fragments may be closer to 1.9% (Supplementary Information 3.2).

The average nucleotide diversity between overlapping archaic fragments, when considering DAV and DAV-linked variants is  $2.00 \times 10^{-5}$



**Fig. 2 | Phylogenetic relationships of archaic fragments to sequenced archaic genomes.** Throughout this figure Vindija is shown in blue, Altai in red and Denisova in green.  $D_{INT}$ , introgressing Denisovan population;  $N_{INT}$ , introgressing Neanderthal population. Branch thickness in the phylogenetic trees reflects the effective population size, in which the human ancestral branch is 10,000. Split times were obtained from a previous study<sup>1</sup> and the  $N_e$  is calculated according to Supplementary Information 2.7.2. **a, b**, Two different scenarios to account for the presence of Denisovan fragments in contemporary Icelanders are shown. **a**, Denisovan introgression into Neandertals is followed by Neanderthal introgression into humans. **b**, Direct Denisovan introgression into the ancestor of all non-Africans. kyr, thousand years ago. **c**, Length distribution of Altai, Denisovan and Vindija fragments. The dashed lines indicate the mean segment lengths. Altai Neanderthal, length =  $157 \pm 1.045$  kb; Denisovan, length =  $108 \pm 1.498$  kb; Vindija Neanderthal, length =  $162 \pm 0.626$  kb; lengths are mean  $\pm$  s.e.m.

(95% confidence interval,  $1.86 \times 10^{-5}$ – $2.12 \times 10^{-5}$ ), which is similar to the heterozygosity in Neanderthals (Altai Neanderthal,  $1.63 \times 10^{-5}$ ; Vindija Neanderthal,  $1.71 \times 10^{-5}$ ) and the Altai Denisovan ( $1.89 \times 10^{-5}$ ). Some genomic regions contained highly diverse archaic fragments indicative of multiple introgressing haplotypes (Extended Data Fig. 3). To shed light on the number of archaic individuals who contributed to contemporary humans, we clustered overlapping fragments and divided them into subgroups that did not exceed a mean pairwise distance of 1 variant per 10 kb (Supplementary Information 5). We postulate that each subgroup approximates a different introgressing archaic haplotype. When only DAV and DAV-linked variants were used to define subgroups, the percentage of fragments assigned to just one subgroup is 88.1% (Fig. 1c and Supplementary Information 5.2), which increases to 93.2% using only DAV variants. However, even with such stringent criteria, we find genomic regions with up to six different archaic haplotypes—indicating that multiple archaic individuals were involved in the introgression (Supplementary Fig. 5.1.1).

Figure 1d shows the genomic distribution of archaic fragments and the DAV genome that they share the most variants with: Vindija Neanderthal (50.8%), Altai Neanderthal (13.1%), Denisovan (3.3%), two or more DAV genomes (20.4%) or not shared with a DAV genome (unknown, 12.2%). Fragments of unknown origin are longer (mean = 77.9 kb, s.e.m. = 0.5 kb) than false-positive archaic fragments in simulations (mean = 40.7 kb, s.e.m. = 0.7 kb) (Extended Data Fig. 1 and Supplementary Information 3.2). However, a higher false-positive rate for unknown fragments is suggested by a shorter mean coalescence time to sub-Saharan Africans (34,130 generations, s.e.m. = 136 generations) than for archaic fragments that contain DAV variants (38,318 generations, s.e.m. = 48 generations) (Wilcoxon signed-ranked test,  $P < 2.2 \times 10^{-16}$ ) (Extended Data Fig. 4). Because our method of detection is not dependent on existing archaic genomes, we are able to detect

mosaic fragments that switch similarity with different DAV genomes along their length. Figure 1d shows an example of a mosaic fragment, with multiple Denisovan-specific derived alleles in the first 250 kb, followed by Altai-specific variants (Supplementary Figs. 2.6.4–2.6.10). We find that 18.9% of fragments are mosaic.

Both our results and previous studies<sup>1</sup> indicate that the introgressing Neanderthal was closer to the Vindija Neanderthal than the Altai Neanderthal or Denisovan. However, our method yields many long fragments that are most closely related to the Altai Neanderthal and Denisovan genomes (Fig. 2c) and some previous studies have also reported such fragments in Europeans<sup>3,20,23</sup>. One possible explanation is introgression from groups with genomes that are related to the Denisovan and Altai genomes into introgressing Neanderthals before this group contributed to modern human genomes. Alternatively, the introgressing Neanderthals could have been a Vindija-like group that carried some anciently diverged haplotypes due to incomplete lineage sorting that now seem, by chance, to be more similar to the Altai and Denisovan genomes than the Vindija genome. To resolve the origin of these archaic fragments, we performed extensive simulations under different demographic and admixture models (Supplementary Information 3.3 and Supplementary Fig. 3.1.1). The results indicate that the observed characteristics of Denisovan-like fragments in Icelanders are not compatible with a simple introgression from a Vindija-like group without that population having had prior admixture with a Denisovan-like group (Supplementary Information 3.3.3 and Supplementary Fig. 3.1.1). An equally intriguing scenario that cannot be ruled out is direct admixture from a Denisovan-like group into the common ancestors of non-Africans before the main Neanderthal admixture event (Fig. 2a, b).

We estimated the effective population sizes ( $N_e$ ) of different archaic groups using pairwise differences between the archaic fragments identified in Icelanders and the high-coverage DAV genomes (Supplementary Information 2.7) and divergence times estimated in previous studies<sup>1,18</sup> (Supplementary Information 2.7.2). We find that Neanderthals had a relatively small  $N_e$  of 2,000–3,000 individuals, in agreement with previous pairwise sequential Markov chain analyses of Neanderthal genomes<sup>1</sup>.

Our greater sample size ( $n = 27,566$  compared with  $n = 502$  in the 1000 Genomes Project) enables a more fine-scale identification of genomic regions with very little or no archaic introgression (archaic deserts). Searching for 1-Mb windows with no fragments containing DAV variants, we found 282 distinct archaic deserts covering 570 Mb (23.3% of the callable genome) (Extended Data Fig. 5 and Supplementary Data 4). The X chromosome is particularly devoid of archaic introgression as previously reported<sup>15</sup> (Fig. 1d). The archaic deserts are slightly more gene dense ( $7.52 \pm 0.372$  genes per Mb) than non-deserts ( $6.87 \pm 0.179$  genes per Mb) (Wilcoxon signed-ranked test,  $P = 0.043$ ). They also have lower mean recombination rates (0.894 cM per Mb, s.e.m. = 0.031) than non-deserts (1.36 cM per Mb, s.e.m. = 0.02) (Wilcoxon signed-ranked test,  $P < 2.2 \times 10^{-16}$ ) (Extended Data Fig. 5), also after adjusting for gene density (Supplementary Fig. 4.2.1). Furthermore, the proportion of archaic fragments spanned by DAV-linked and DAV variants was higher in regions with a greater recombination rate (Spearman's  $\rho = 0.15$ ,  $P = 4.4 \times 10^{-54}$ ) and the nucleotide diversity of archaic fragments was also higher in these regions (Spearman's  $\rho = 0.18$ ,  $P = 1.3 \times 10^{-79}$ ) (Extended Data Fig. 6). Taken together, these observations indicate that non-deleterious archaic variants were more likely to be retained in the human gene pool when they could be uncoupled from deleterious archaic variants by recombination<sup>25</sup>.

## Mutation processes in archaic fragments

The rate and types of mutation are influenced by several factors that may differ between human populations and between great-ape species<sup>26</sup>. The fully phased Icelandic genomes allow us to compare syntenic archaic and non-archaic fragments from the same individual

# Article

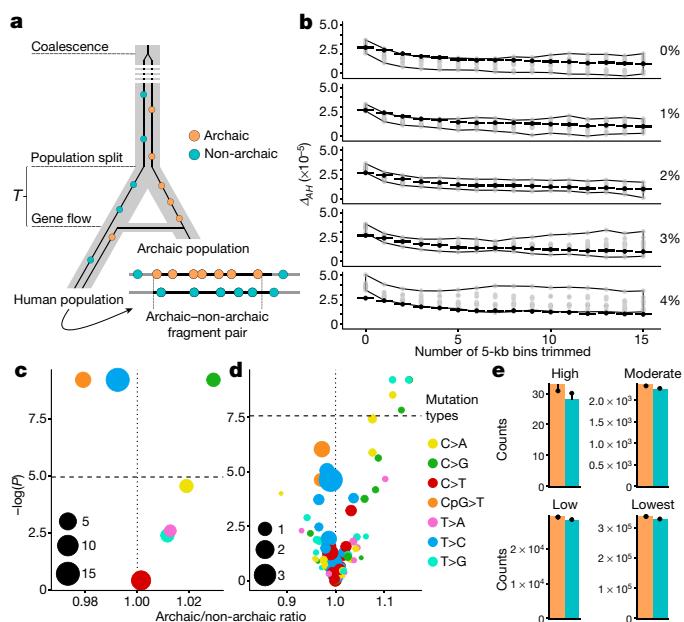
to determine whether differences in mutational processes existed between humans and Neanderthals during the approximately 500,000 years of divergence before archaic introgression ( $T$  in Fig. 3a; Supplementary Information 6.1). We calculated the mean difference in the number of derived alleles (including those shared with Africans) between the paired archaic and non-archaic fragments ( $\Delta_{AH}$ , Supplementary Information 6.2). We modelled the ascertainment bias of the hidden Markov model approach used to identify archaic fragments, which is predisposed towards SNP-dense fragments (Supplementary Information 3.4.1), thereby inflating the  $\Delta_{AH}$  particularly at the edges of fragments (Supplementary Information 3.4.2, 6.2.1). We compared the observed  $\Delta_{AH}$  in Icelandic genomes with multiple simulated scenarios of increasing differences in the density of archaic-to-human SNPs (Fig. 3b and Supplementary Information 3.4.1). Although the best-fitting scenario corresponds to a 2% higher SNP density in archaic fragments (Extended Data Fig. 7), given the uncertainties associated with the parameters used in the coalescent simulations and variance among replicates of the same scenario, we do not find a greater mutation rate in the Neanderthal lineage before the introgression event.

We next compared the mutation spectrum between the paired archaic and non-archaic fragments. A set of non-overlapping fragments was used to avoid counting the same derived allele multiple times (Supplementary Information 6.1) and we trimmed 15 kb from both ends of all fragments to avoid edge effects (Supplementary Information 6.3.1). We classified mutations as ancestral-to-derived alleles and collapsed the strand complement mutation counts into seven distinct mutational types, separating CpG>TpG from C>T, as the latter mutation type is thought to evolve in a more clock-like manner<sup>27</sup>. The archaic:non-archaic proportional ratio (Fig. 3c, Supplementary Data 7 and Supplementary Information 6.3.2) revealed more C>G and fewer T>C and CpG>T variants in archaic fragments than in their non-archaic counterparts. Notably, the most-extreme ratios are observed for C>G and CpG>T variants, which have the lowest (1:1) and highest (6:1) paternal-to-maternal age effect ratios for de novo mutations in contemporary humans (the mean ratio across all types is 3:1)<sup>28</sup>. These differences are consistent with mothers having been older and fathers younger in the Neanderthal lineage than in the human lineage, although other causes cannot be ruled out. The C>G enrichment in archaic fragments is not due to an excess of archaic fragments in the C>G-cluster-enriched regions<sup>28</sup>. Furthermore, the differences in mutational types are robust to several conservative filtering schemes and are therefore probably not due to false-positive archaic fragments (Supplementary Information 6.3.2 and Supplementary Data 7). We next expanded the classification of each mutation type, including their 5' and 3' base-pair context, resulting in 96 distinct mutational types (Fig. 3d and Supplementary Data 7). Five types (CCT>G, CCC>G, CTG>G, TCC>A and CCA>G) were overrepresented in archaic fragments after Bonferroni adjustment, three of which involved C>G mutations.

It has been suggested<sup>29–31</sup> that Neanderthals accumulated more deleterious mutations than humans because of less efficient selection due to a smaller  $N_e$ . We did not find an excess of deleterious variants in the archaic fragments identified in Icelanders ( $\chi^2$  test,  $P=0.71$ ) (Fig. 3e and Supplementary Table 6.4.1), after comparing their distribution across four functional impact categories (lowest, low, moderate and high) that were obtained from the Variant Effect Predictor<sup>32</sup> (Supplementary Information 6.4). Therefore, if introgressing Neanderthal fragments had more deleterious variants at the time of introgression, these have already been removed by purifying selection.

## Phenotypic effects of archaic fragments

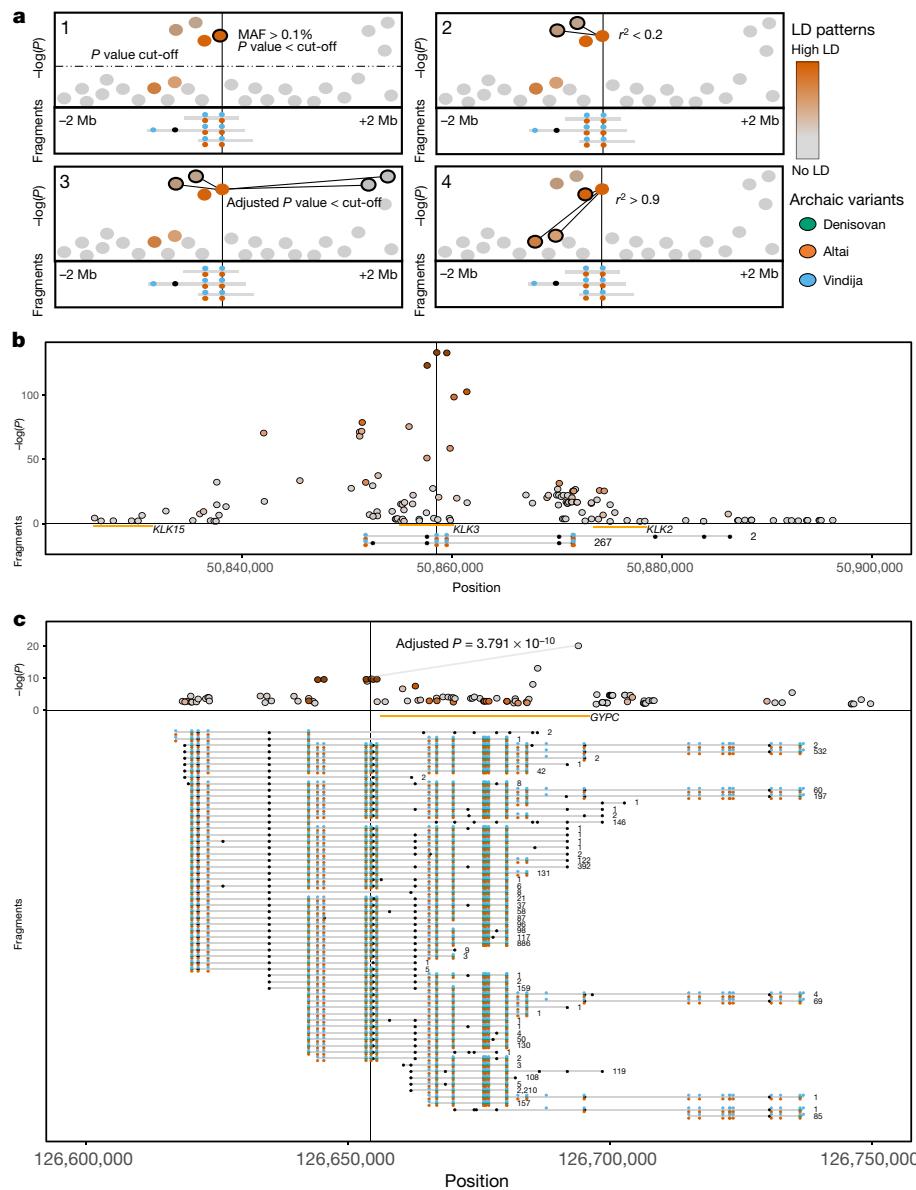
To study the influence of archaic variants on the phenotypic diversity of contemporary humans, we examined their association with 271 traits based on around 32 million variants called and imputed for 214,192 Icelanders (Supplementary Information 7.1). In the first step of filtering



**Fig. 3 | Comparison of the rate, spectrum and influence of mutations on archaic and non-archaic fragments.** **a**, The phylogenetic tree connecting human and archaic populations (grey bands). The black lines in the tree represent coalescence paths of an archaic and a non-archaic fragment pair in an Icelander. Circles represent mutations (derived alleles) that accumulated on non-archaic (blue) and archaic (orange) fragments.  $T$  refers to the time these two fragments were isolated in human and archaic ancestral groups. The two horizontal lines with mutations illustrate the paired fragments sampled. **b**, Comparison of the observed Icelandic  $\Delta_{AH}$  (black dots with 95% confidence intervals) with expectations based on 10 replicates (grey dots) of simulations of increasing Neanderthal-to-human mutation rate percentages (as indicated on the right). Wide bands show the maximum and minimum  $\Delta_{AH}$  among the 10 replicates per scenario. Results are based on trimming 0–15 5-kb bins from fragment ends (x axis). Only fragments  $\geq 100$  kb are used. **c**, The archaic/non-archaic proportional ratio (x axis) of the number of derived alleles against the  $P$  value (two-sided permutation test, y axis) for each of seven different mutation types (represented by different colours). The size of each point shows the mean archaic and non-archaic fragment sample size for each mutation type ( $\times 10^4$ , sample sizes are included in Supplementary Data 7). The horizontal dashed line shows the Bonferroni-adjusted significance threshold and the vertical dashed line indicates a ratio of 1 (no enrichment). **d**, The same as **c**, but subdividing mutation types according to the 96-mutation type spectrum. **e**, Observed counts of derived alleles from human (blue) and archaic (orange) fragments based on four categories of functional impact. The  $\chi^2$ -expected counts are denoted as black points and the black lines indicate the difference from the observed counts.

(Fig. 4a), we evaluated the association of the 395,304 archaic-derived alleles with each of 271 phenotypes using genome-wide significance thresholds based on functional effects (such as, exon, intron and intergenic regions)<sup>33</sup>. This yielded 4,361 archaic phenotype-associated variants, corresponding to 651 independent association signals after accounting for linkage disequilibrium between them.

In the second step of filtering, we tested whether any correlated ( $r^2 > 0.2$ ) non-archaic variant within a 2-Mb radius better accounted for an association than the strongest archaic candidate variant. For 550 association signals, we found a non-archaic variant in high linkage disequilibrium with the strongest archaic candidate. Of these, 431 involved a non-archaic variant with a substantially stronger association (>10-fold) with the phenotype in question. For 33 cases, the non-archaic variant had a stronger association (<10-fold) with the phenotype. In another 86 instances, archaic variants were disregarded because a highly correlated ( $r^2 > 0.9$ ) non-archaic variant with a slightly stronger association was identified. This left 101 archaic variants.



**Fig. 4 | Phenotype association with archaic variants.** **a**, Procedure to determine whether genotype–phenotype associations are due to archaic variants. Each subplot shows a 4-Mb region on the x axis and is divided into two frames. In the top frame, the y axis shows the  $-\log_{10}$ -transformed values of the association  $P$  value for variants in the region coloured by their linkage disequilibrium (LD) to the candidate variant, which is marked by a vertical line. All archaic variants are outlined. The bottom frame shows the archaic fragments (grey) punctuated by archaic variants (Altai, orange; Vindija, blue; Denisovan, green; other, black). (1) We selected all archaic variants with an association  $P$  value under the genome-wide significance threshold and a minor allele frequency (MAF) above 0.1%. (2) The candidate archaic variant is rejected as the cause of the association if there is a non-archaic variant with a stronger association to the phenotype with  $r^2 > 0.2$  to the candidate archaic variant.

(3) We reject the variant if the  $P$  value of the candidate variant drops below the genome-wide significance threshold after conditioning on a non-archaic variant with a lower  $P$  value within a 2-Mb radius. (4) We reject candidates that are not in strong linkage disequilibrium ( $r^2 > 0.9$ ) with at least one other archaic variant from the same fragment. **b**, Association of an archaic variant (rs17632542) on chromosome 19 with prostate-specific antigen. The top frame shows association  $P$  values for all variants in the region, coloured according to their linkage disequilibrium with the candidate archaic variant. The bottom frame shows genes (orange) and archaic fragments as described for a. **c**, Association of an archaic variant (rs28387074) on chromosome 2 with mean corpuscular haemoglobin concentration (as described for b). Although another non-archaic variant in the same region has a lower  $P$  value, the association of the candidate variant remains significant after adjustment.

In the third step, we tested whether the archaic variants remained genome-wide significant after conditioning on non-archaic variants within a 2-Mb radius with a stronger association to the phenotype. This analysis left 64 archaic variants.

Finally, we pruned variants that were not in high linkage disequilibrium with other archaic variants ( $r^2 > 0.9$ ), leaving only five independent archaic variants likely to be a true source of phenotypic association (Supplementary Data 8). One is an example of a strong association with a reduced level of prostate-specific antigen driven by rs17632542 (the

derived allele of which is found in Altai and Vindija Neanderthals), which has been reported to reduce the risk of prostate cancer; however, the introgressed status of the variants was not reported<sup>34</sup> (Fig. 4b). Another is rs28387074, which decreases the concentration of haemoglobin and survives adjustment for non-archaic variants that are responsible for another strong signal in the same region (Fig. 4c). The remaining three association signals due to archaic variants are for reduced height (rs3118914), decrease in mean corpuscular haemoglobin (rs72728264) and an increase in plasma prothrombin time (rs6013) (Supplementary

# Article

Data 8). Two of the association signals (rs3118914 and rs72728264) are DAV-unlinked variants that are in strong linkage disequilibrium ( $r^2 > 0.9$ ) with other DAV-unlinked variants.

We also assessed 30 previously reported phenotypic associations with archaic variants<sup>12,13,15</sup>. In four cases, we did not have a sufficiently similar phenotype to test. For the remaining 26 associations, 10 were not nominally significant ( $P > 0.05$ ) in our genome-wide association studies and 13 lost significance when conditioned on a nearby non-archaic variant that was more likely to be the true source of the association (Supplementary Data 8). These non-archaic variants are present in African populations, have a greater effect on the phenotype and a lower  $P$  value than the originally reported 13 archaic variants (Supplementary Figs. 7.1.6–7.1.10). We could therefore only validate the archaic origin of 3 out of 26 previously reported association findings attributed to archaic variants (Supplementary Data 8). This highlights the importance of considering flanking non-archaic variants when assigning an archaic origin to a phenotype association. To assess the genome-wide effect of archaic introgression on phenotype variation, we also counted the number of archaic-derived alleles carried by each contemporary Icelander and tested this polygenic score of archaic ancestry for association with each of the 271 phenotypes. In the case of height, for example, this test would reveal whether the surviving fragments from archaic introgression resulted in taller or shorter contemporary Icelanders. After adjusting for the number of tests, we found no evidence for association between the polygenic score and any of the 271 tested phenotypes (Supplementary Data 8), indicating that archaic ancestry has—at most—a modest directional impact on contemporary human phenotype variation.

## Discussion

We show that a large part of an archaic genome can be mined from contemporary descendants of populations that were recipients of introgression around 50–60 thousand years ago. The recovered archaic fragments are consistent with being descended from multiple archaic individuals, who belonged to an archaic population similar to the Vindija Neanderthal. However, the considerable proportion of archaic fragments that are closer to the Denisovan genome cannot be explained by incomplete lineage sorting (Supplementary Information 3.3.3.1). Rather, they require Denisovan introgression, either directly into humans or into Neanderthals who later mixed with humans, which must have occurred soon after they migrated out of Africa, because its signal is found in all contemporary non-African populations from the Simons Genome Diversity Project (Supplementary Information 5.2). This raises the possibility that there were Denisovan-like groups west of the Altai mountains, where such gene flow into humans must have occurred. This is consistent with growing evidence of extensive population movements towards the end of Neanderthal history both in Europe<sup>35</sup> and in the Altai mountains<sup>19</sup>. Hopefully, additional archaic genomes will shed light on this complex history of interbreeding between hominid groups.

The similar mutation rates in Neanderthals and modern humans imply that the apparent slowdown in the human mutation rate<sup>36</sup> is unlikely to have occurred between 500,000 and 55,000 years ago. However, differences in the mutation spectrum of archaic and non-archaic fragments raise the possibility of long-term differences in male and female generation intervals between the species. Finally, our most far-reaching conclusion is that archaic introgression has a relatively minor effect on phenotypic variation in contemporary humans. Given the non-random genomic distribution of archaic fragments in contemporary Icelanders, it follows that this influence must have been greater in the past.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information,

acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2225-9>.

- Prüfer, K. et al. A high-coverage Neandertal genome from Vindija Cave in Croatia. *Science* **358**, 655–658 (2017).
- Fu, Q. et al. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445–449 (2014).
- Browning, S. R., Browning, B. L., Zhou, Y., Tucci, S. & Akey, J. M. Analysis of human sequence data reveals two pulses of archaic Denisovan admixture. *Cell* **173**, 53–61 (2018).
- Wall, J. D. et al. Higher levels of Neanderthal ancestry in East Asians than in Europeans. *Genetics* **194**, 199–209 (2013).
- Kim, B. Y. & Lohmueller, K. E. Selection and reduced population size cannot explain higher amounts of Neandertal ancestry in East Asian than in European human populations. *Am. J. Hum. Genet.* **96**, 454–461 (2015).
- Vernot, B. & Akey, J. M. Complex history of admixture between modern humans and Neanderthals. *Am. J. Hum. Genet.* **96**, 448–453 (2015).
- Villanea, F. A. & Schraiber, J. G. Multiple episodes of interbreeding between Neanderthal and modern humans. *Nat. Ecol. Evol.* **3**, 39–44 (2019).
- Dannemann, M. & Kelso, J. The contribution of Neanderthals to phenotypic variation in modern humans. *Am. J. Hum. Genet.* **101**, 578–589 (2017).
- Gittelman, R. M. et al. Archaic hominin admixture facilitated adaptation to out-of-Africa environments. *Curr. Biol.* **26**, 3375–3382 (2016).
- Gregory, M. D. et al. Neanderthal-derived genetic variation shapes modern human cranium and brain. *Sci. Rep.* **7**, 6308 (2017).
- McCoy, R. C., Wakefield, J. & Akey, J. M. Impacts of Neanderthal-introgressed sequences on the landscape of human gene expression. *Cell* **168**, 916–927 (2017).
- Dannemann, M., Prüfer, K. & Kelso, J. Functional implications of Neanderthal introgression in modern humans. *Genome Biol.* **18**, 61 (2017).
- Simonti, C. N. et al. The phenotypic legacy of admixture between modern humans and Neanderthals. *Science* **351**, 737–741 (2016).
- Vernot, B. & Akey, J. M. Resurrecting surviving Neandertal lineages from modern human genomes. *Science* **343**, 1017–1021 (2014).
- Sankararaman, S. et al. The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* **507**, 354–357 (2014).
- Steinrücken, M., Spence, J. P., Kamm, J. A., Wieczorek, E. & Song, Y. S. Model-based detection and analysis of introgressed Neanderthal ancestry in modern humans. *Mol. Ecol.* **27**, 3873–3888 (2018).
- Meyer, M. et al. A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222–226 (2012).
- Prüfer, K. et al. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505**, 43–49 (2014).
- Slon, V. et al. The genome of the offspring of a Neanderthal mother and a Denisovan father. *Nature* **561**, 113–116 (2018).
- Skov, L. et al. Detecting archaic introgression using an unadmixed outgroup. *PLoS Genet.* **14**, e1007641 (2018).
- Kong, A. et al. Detection of sharing by descent, long-range phasing and haplotype imputation. *Nat. Genet.* **40**, 1068–1075 (2008).
- The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
- Sankararaman, S., Mallick, S., Patterson, N. & Reich, D. The combined landscape of Denisovan and Neanderthal ancestry in present-day humans. *Curr. Biol.* **26**, 1241–1247 (2016).
- Vernot, B. et al. Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals. *Science* **352**, 235–239 (2016).
- Schumer, M. et al. Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science* **360**, 656–660 (2018).
- Harris, K. & Pritchard, J. K. Rapid evolution of the human mutation spectrum. *eLife* **6**, e24284 (2017).
- Moorjani, P., Amorim, C. E. G., Arndt, P. F. & Przeworski, M. Variation in the molecular clock of primates. *Proc. Natl. Acad. Sci. USA* **113**, 10607–10612 (2016).
- Jónsson, H. et al. Parental influence on human germline de novo mutations in 1,548 trios from Iceland. *Nature* **549**, 519–522 (2017).
- Harris, K. & Nielsen, R. The genetic cost of Neanderthal introgression. *Genetics* **203**, 881–891 (2016).
- Juric, I., Aeschbacher, S. & Coop, G. The strength of selection against Neanderthal introgression. *PLoS Genet.* **12**, e1006340 (2016).
- Castellano, S. et al. Patterns of coding variation in the complete exomes of three Neandertals. *Proc. Natl. Acad. Sci. USA* **111**, 6666–6671 (2014).
- McLaren, W. et al. The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
- Sveinbjörnsson, G. et al. Weighting sequence variants based on their annotation increases power of whole-genome association studies. *Nat. Genet.* **48**, 314–317 (2016).
- Kote-Jarai, Z. et al. Identification of a novel prostate cancer susceptibility variant in the *KLK3* gene transcript. *Hum. Genet.* **129**, 687–694 (2011).
- Hajdinjak, M. et al. Reconstructing the genetic history of late Neanderthals. *Nature* **555**, 652–656 (2018).
- Besenbacher, S., Hvilsted, C., Marques-Bonet, T., Mailund, T. & Schierup, M. H. Direct estimation of mutations in great apes reconciles phylogenetic dating. *Nat. Ecol. Evol.* **3**, 286–292 (2019).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment. A detailed description of all analyses carried in this study out is included in the Supplementary Information.

### Ethics statement

We used all of the samples (whole-genome sequenced and imputed,  $n=27,566$ ) available at deCODE Genetics at the time the study started. Recruitment is based on volunteers providing samples for particular disease and/or control studies either directly or through clinical collaborators. The study was undertaken on the basis of approvals from the National Bioethics Committee and the Icelandic Data Protection Authority. Blood or buccal samples were taken from individuals who participated in various studies, after receiving informed consent from the participants or their guardians.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

### Data availability

All summary statistics such as archaic fragments and which SNPs they contain are available as Supplementary Data. Because Icelandic law and

the regulations of the Icelandic Data Protection Authority prohibit the release of individual level and personally identifying data, collaborators who want access to individual genotype level data have to access the data locally at our Icelandic facilities.

### Code availability

The Python script for performing simulations is available on Github (<https://github.com/LauritsSkov/ArchaicSimulations>).

**Acknowledgements** We thank K. Pruefer, B. Vernot, B. Peter, J. Kelso and S. Pääbo for comments on an earlier version of the manuscript. The study was supported by grant NNF18OC0031004 from the Novo Nordisk Foundation and grant 6108-00385 from the Research Council of Independent Research.

**Author contributions** L.S., M.C.M. and G.S. analysed the data with input from H.J., B.H., D.F.G., A.H. and M.H.S. L.S. and M.C.M. created the methods for analysing the data. L.S., M.C.M., A.H., F.M., K.S. and M.H.S. designed the study. L.S., M.C.M., A.H. and M.H.S. wrote the manuscript with input from all authors.

**Competing interests** All of the authors (except for L.S., M.C.M., F.M., E.A.L. and M.H.S.) are employees of deCODE Genetics and Amgen.

#### Additional information

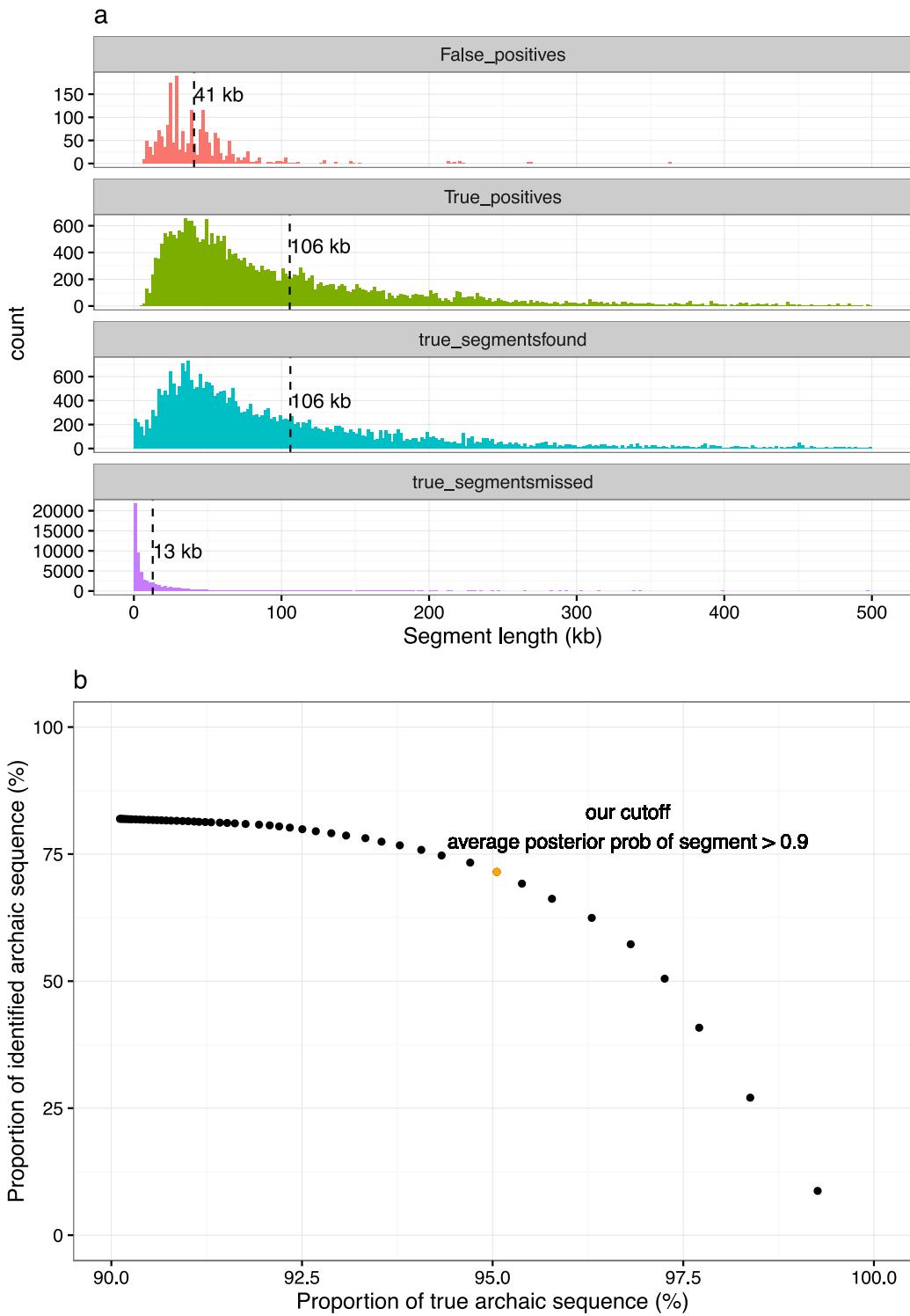
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2225-9>.

**Correspondence and requests for materials** should be addressed to L.S., M.H.S. or K.S.

**Peer review information** *Nature* thanks Stephan Schiffels and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

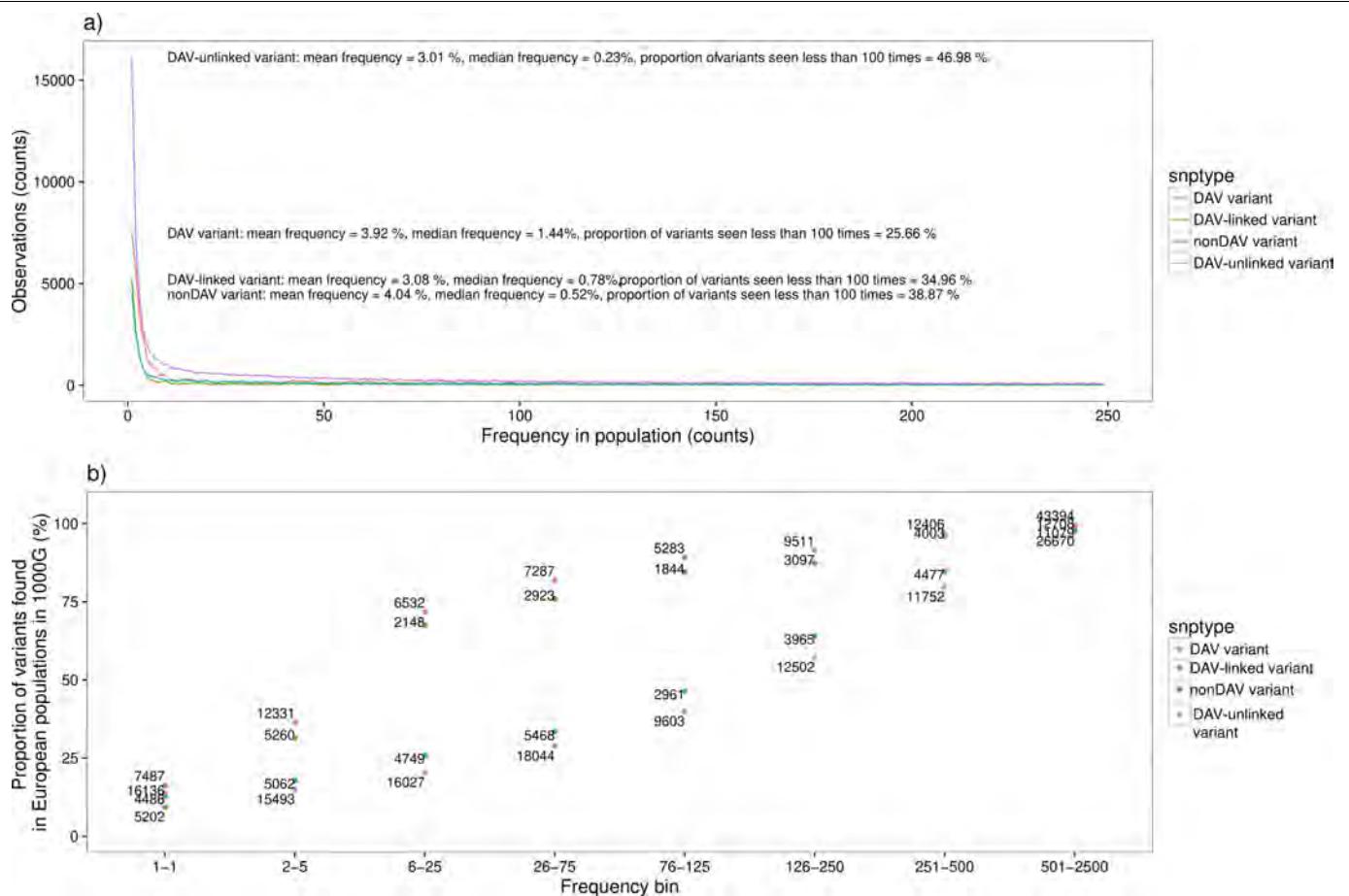
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

# Article



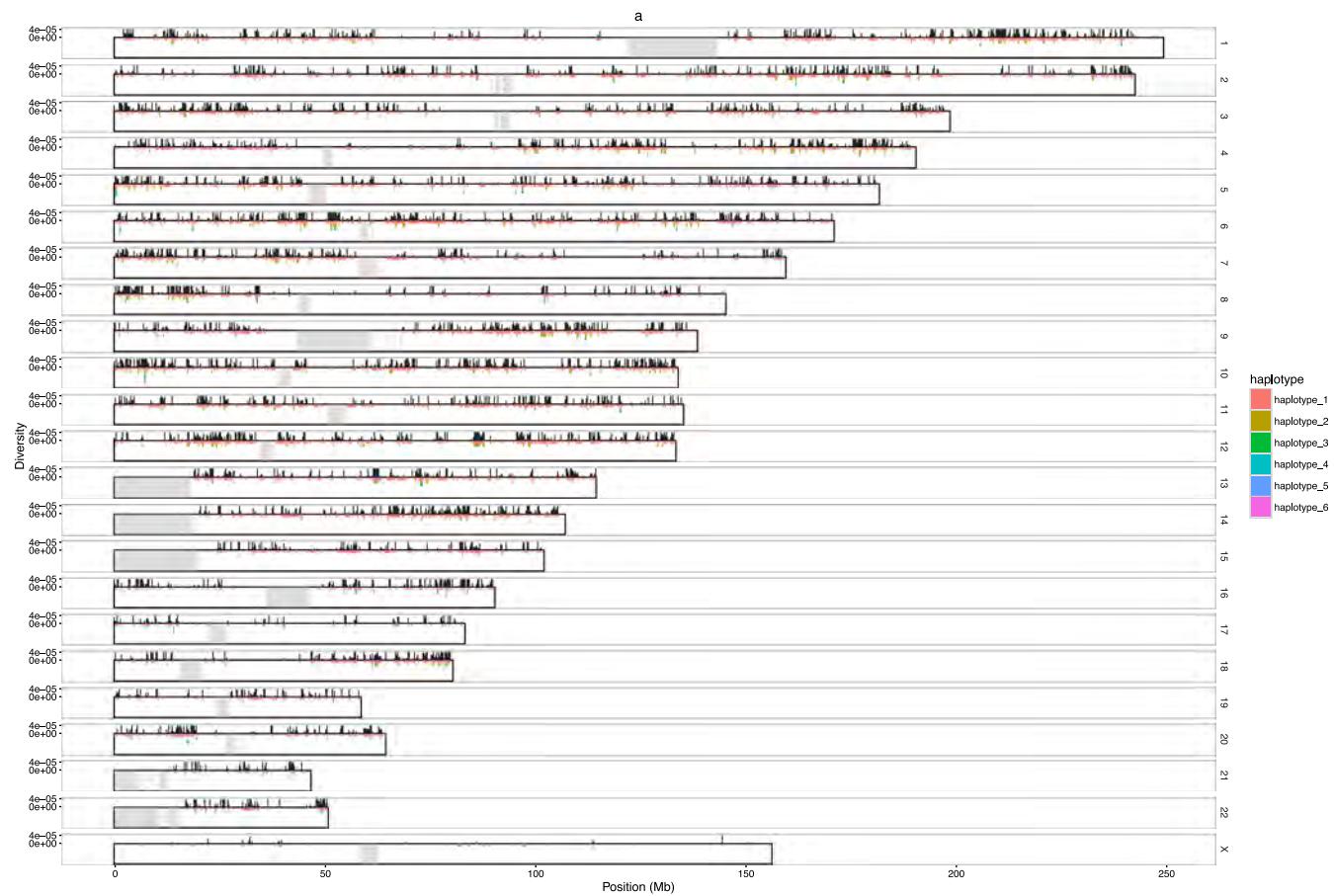
**Extended Data Fig. 1 | Estimates of the false-positive rate of archaic inference based on simulations.** **a**, Length distributions of false-positive and true-positive calls for fragments inferred to be archaic, and the length distributions for simulated archaic fragments that were found and those that

were missed. The dashed lines indicate the mean value of the distribution. **b**, False-positive rate as a function of the mean posterior probability of being archaic for each fragment.

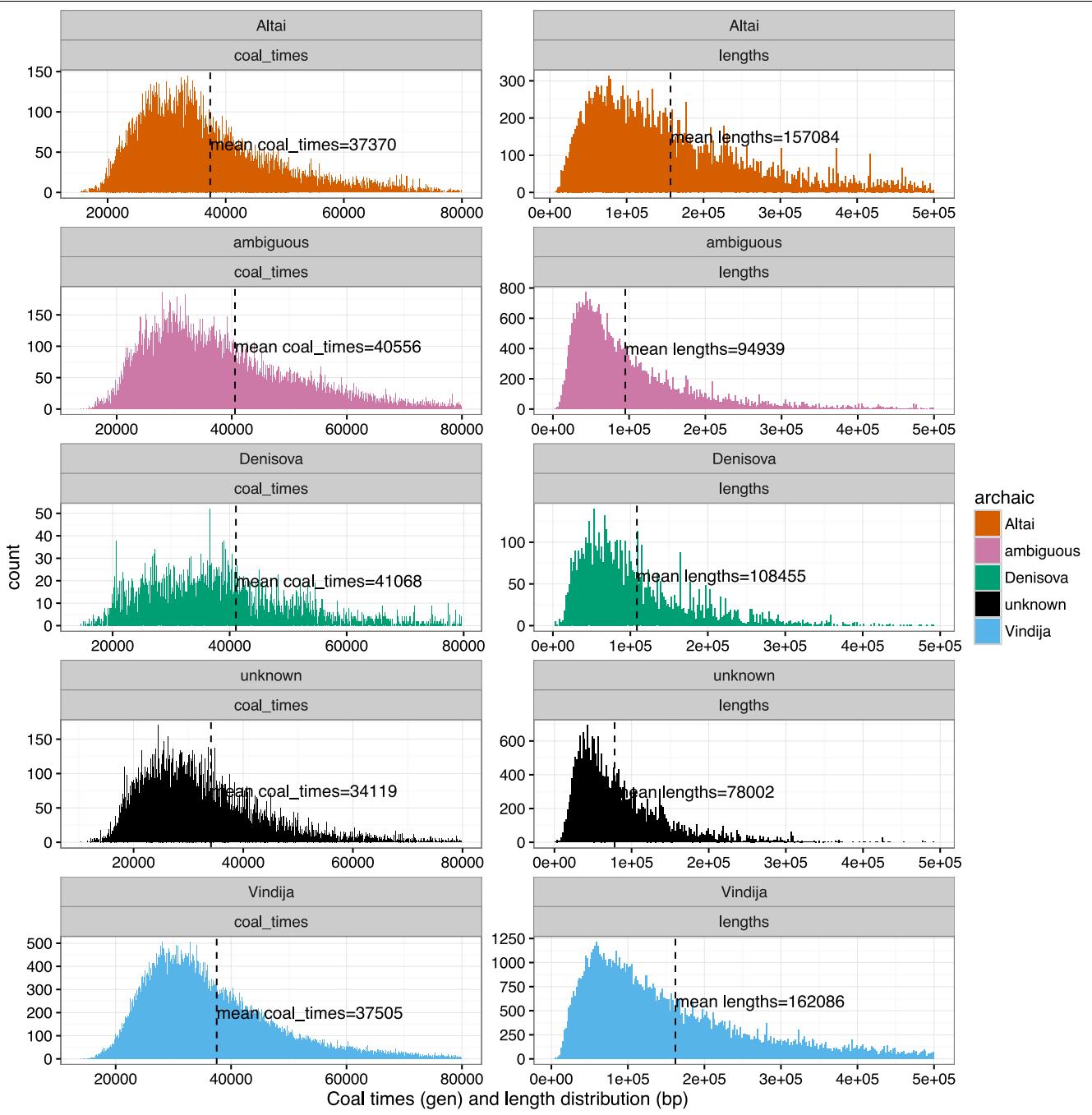


**Extended Data Fig. 2 | Properties of SNP classes in archaic fragments.** **a**, The site frequency spectrum for DAV variants, DAV-linked variants, DAV-unlinked variants and non-DAV variants. The x axis shows variants found on 1–250 chromosomes in our Icelandic population sample and the y axis shows the number of variants in each category. **b**, The number of times a variant is found in European populations (Utah residents (CEPH) with northern and western European ancestry (CEU), Toscani in Italy (TSI), Finnish in Finland (FIN), British

in England and Scotland (GBR) and Iberian population in Spain (IBS); codes as per the previous study) from the 1000 Genomes Project (1000G) as a function of the number of times it is observed in Iceland. The numerical labels represent the number of variants that belong to each category defined by the two axes and the type of archaic variant. The x axis is truncated at variants that were found 2,500 times, which corresponds to a frequency of around 5% in Iceland.

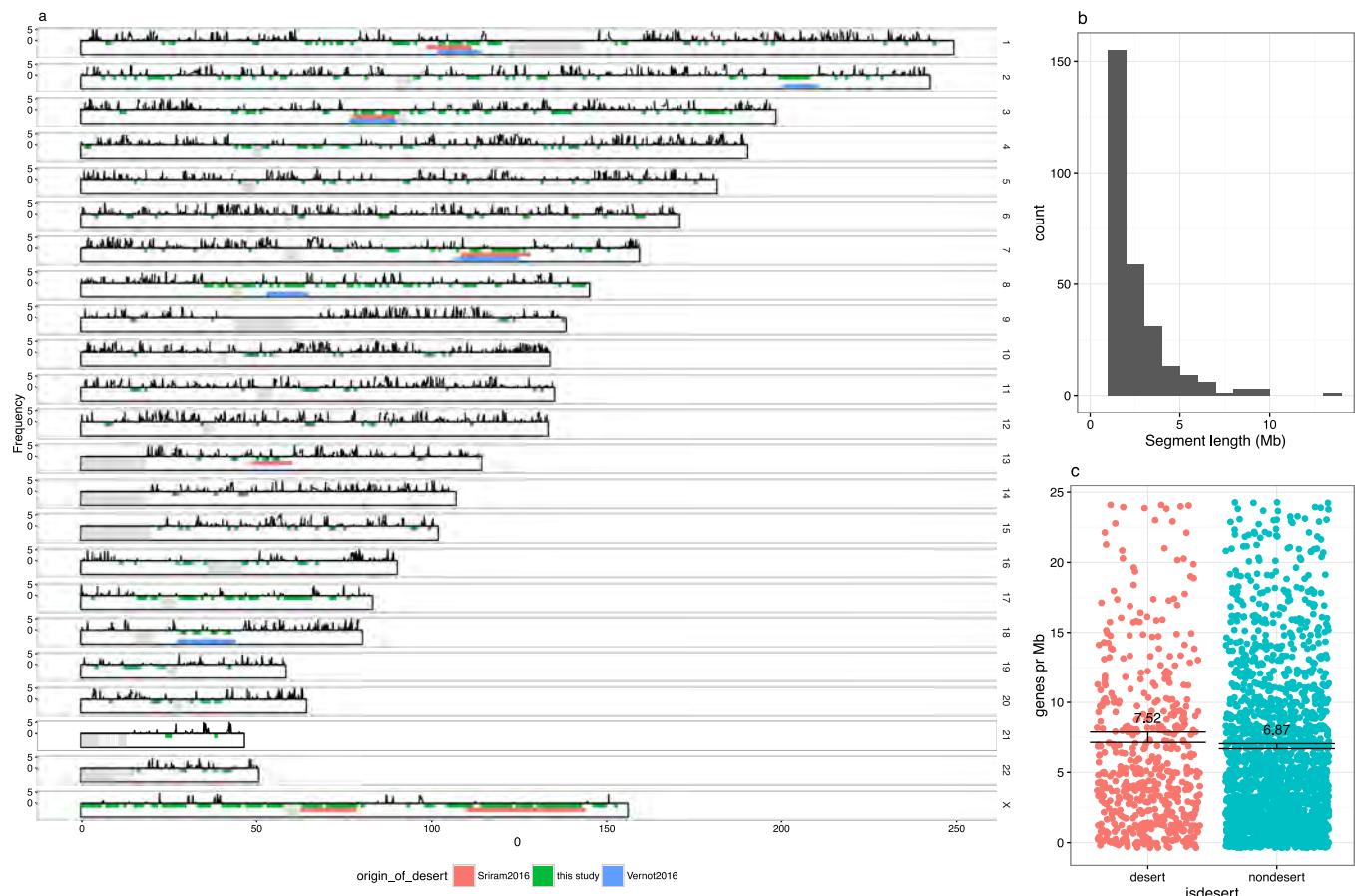


**Extended Data Fig. 3 | Nucleotide diversity based on DAV-linked and DAV variants from archaic fragments.** Using these variants, we identified a maximum of 6 subgroups of fragments, clustered by similarity (based on a mean difference of 1 mismatch per 10 kb per subgroup).



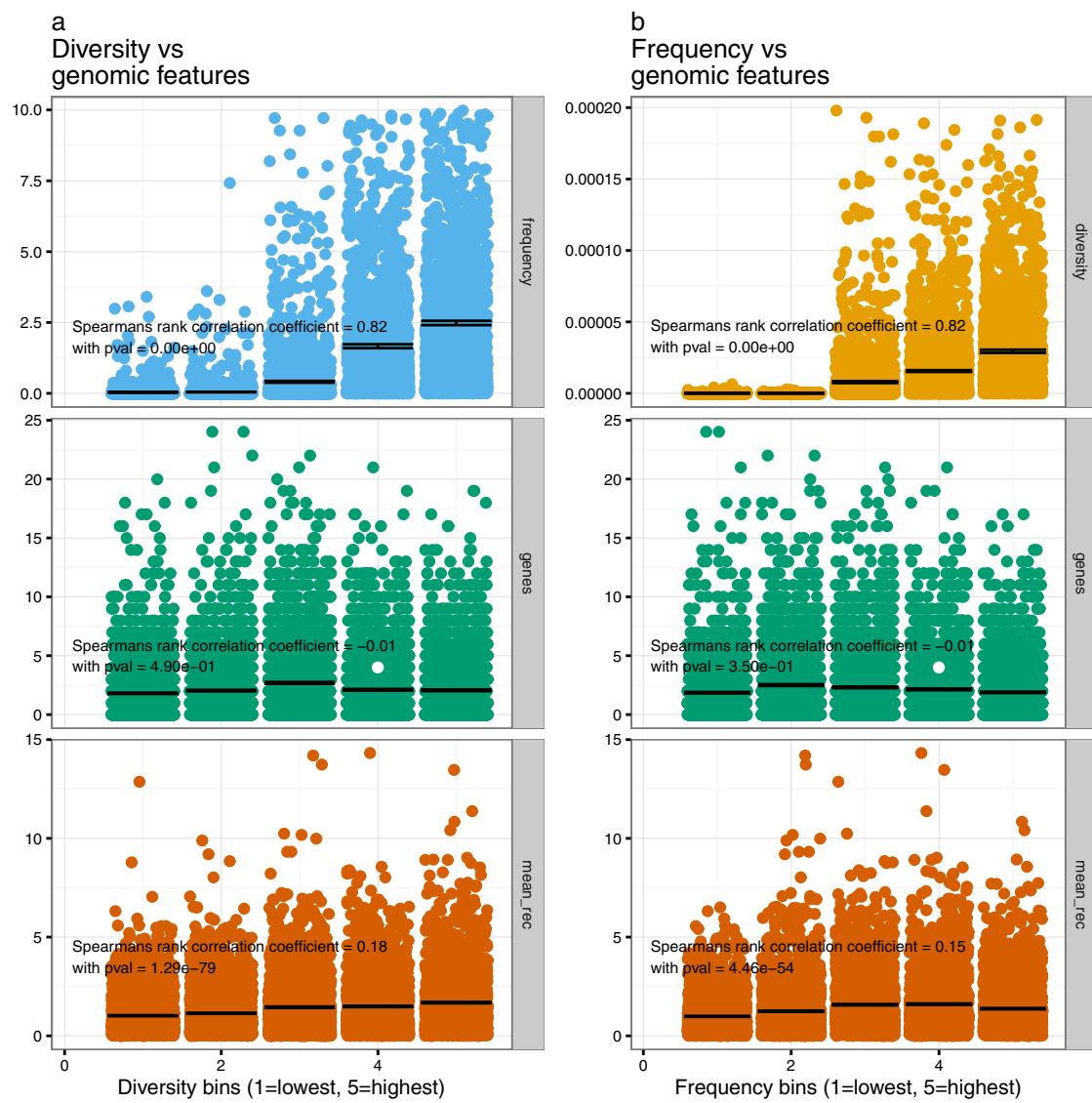
**Extended Data Fig. 4 | Length distribution and coalescence time for all archaic fragments.** The distribution of length and coalescence times (with present day sub-Saharan Africans) for all archaic fragments (mosaic and non-mosaic) are shown according to the archaic genome that they are most

closely related to (Altai, Denisova, Vindija, multiple archaics (ambiguous) and unknown). The coalescence time estimates are described in Supplementary Information 2.7.


**Extended Data Fig. 5 | A genomic map of archaic introgression deserts.**

**a**, The mean frequency of archaic fragments in each 100-kb bin is reported along the genome (ideogram). The y axis is truncated at 5%. Regions with no archaic fragments (only counting fragments with DAV variants and DAV-linked variants) in regions larger than 1 Mb (and a minimum of 100 kb could be called)

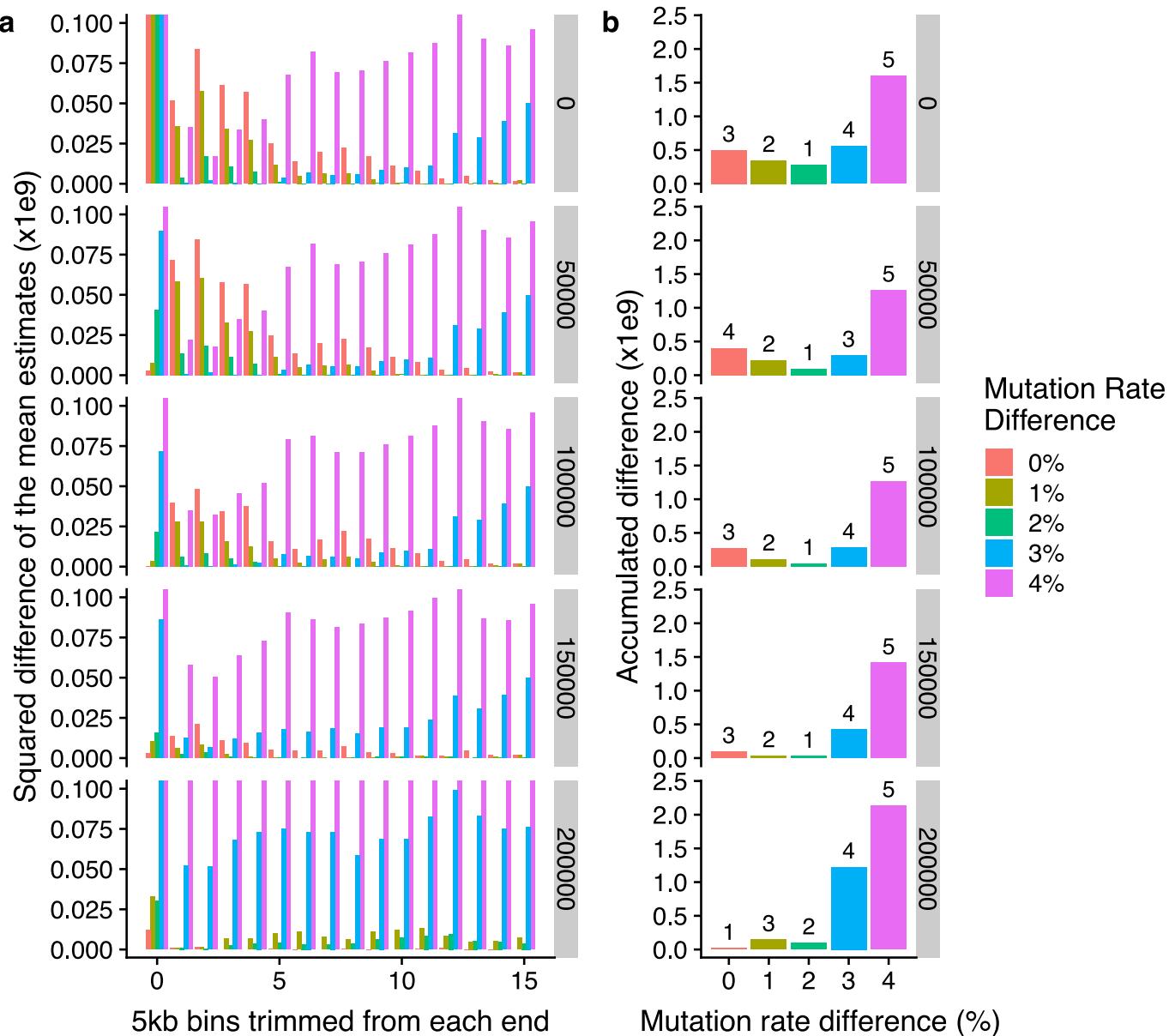
are marked in green, along with previously reported deserts in red<sup>23</sup> and blue<sup>24</sup>. **b**, The size distribution of archaic introgression deserts. **c**, Gene density in deserts ( $n=570$ ) compared with non-deserts ( $n=2,224$ ). The numbers indicate the mean number of genes per Mb and the error bars are 95% confidence intervals. The y axis is truncated at 25 genes per Mb.



**Extended Data Fig. 6 | Archaic introgression and genomic features in 250-kb windows for which more than 1% of the bases in the window could be called.**  $n=10,466$ . **a, b**, The relationship between the nucleotide diversity of archaic fragments (**a**) or their frequency (%) (**b**) and the recombination rate (cM per Mb) and gene density (number of genes per 250 kb). Data were analysed by Spearman's rank correlation coefficient and  $P$  values for all data points (two-sided test) are shown. For visualization, we group the data into five

equally sized bins, sorted by diversity (**a**) or frequency (**b**). Each bin contains 2,093 data points. In this analysis, we considered only DAV-linked variants and DAV variants. The data are coloured according to which feature is being tested (blue for frequency, orange for genes, green for recombination rate and yellow for diversity). The midpoint of the bar is the mean of the measure and error bars are 95% confidence intervals. They axis is truncated at five times the mean value.

# Article



## Extended Data Fig. 7 | Difference between observed and simulated fragments based on increasing Neanderthal-to-human mutation rates.

**a**, The square-root difference of the 10-replicate mean simulated scenarios for  $\Delta_{AH}$  (difference scenarios have mutation rates of 0, 1, 2, 3 and 4%) and the observed values of  $\Delta_{AH}$  for the Icelandic fragments, after trimming different

numbers of 5-kb bins from the ends of each fragment. The analysis was repeated applying 0-, 50-, 100-, 150- and 200-kb minimum fragment length filters (facets). **b**, As in **a**, but adding up the square-root difference value for all the bins trimmed and ranking each bar by the minimum difference (number on top of each bar: 1, minimum difference; 5, maximum difference).

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

## Software and code

Policy information about [availability of computer code](#)

Data collection

The description of the data acquisition and processing is described in the Data Descriptor (Jónsson, H. et al. Whole genome characterization of sequence diversity of 15,220 Icelanders. *Sci. Data* 4:170115 doi: 10.1038/sdata.2017.115 (2017).)

Data analysis

Custom scripts written in python 2.7  
Bedtools v2.25.0-76-g5e7c696z, <https://github.com/arq5x/bedtools2/>  
Msprime v0.4.0  
Scipy v1.3.1

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Icelandic law and the regulations of the Icelandic Data Protection Authority prohibit the release of individual level and personally identifying data. We are actively participating in multiple meta-analysis based on our data and are in collaboration with groups at over 100 international universities and institutions. Therefore, collaborations based on our sequencing data are based on the release of summary level statistics, such as effect sizes and P-values for meta-analysis, or the collaborators travelling to our Icelandic facilities for local data access.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](http://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	27566
Data exclusions	We excluded 268 individuals with low callability (less than 7,000,000 genotyped variants per genome). These variants were excluded because no introgressing archaic fragments could be identified.
Replication	We compare to other methods used for detecting archaic introgression (Supplementary section 2.4). We compare introgressed archaic tracks found in parent-offspring pairs and found a mean concordance between fragments of 99.63 (median=100)
Randomization	The experiments were not randomized
Blinding	The investigators were not blinded to allocation during experiments and outcome assessment.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input checked="" type="checkbox"/>	Antibodies	<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Eukaryotic cell lines	<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	Palaeontology	<input checked="" type="checkbox"/>	MRI-based neuroimaging
<input checked="" type="checkbox"/>	Animals and other organisms		
<input checked="" type="checkbox"/>	Human research participants		
<input checked="" type="checkbox"/>	Clinical data		

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	We used all of the available samples 27566 (whole genome sequenced and imputed) at deCODE Genetics at the time the study was started. There are no covariate-relevant population characteristics of the human research participants.
Recruitment	Recruitment is based on volunteers providing samples for particular disease and/or control studies either directly or through clinical collaborators. We are not aware of any recruitment biases that would impact the results of this study
Ethics oversight	The study is undertaken on the basis of approvals from the National Bioethics Committee and the Icelandic Data Protection Authority. Blood or buccal samples were taken from individuals participating in various studies, after receiving informed consent from them or their guardians.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Article

# Variability in the analysis of a single neuroimaging dataset by many teams

<https://doi.org/10.1038/s41586-020-2314-9>

Received: 14 November 2019

Accepted: 7 April 2020

Published online: 20 May 2020

 Check for updates

A list of authors and affiliations appears in the online version of the paper.

Data analysis workflows in many scientific domains have become increasingly complex and flexible. Here we assess the effect of this flexibility on the results of functional magnetic resonance imaging by asking 70 independent teams to analyse the same dataset, testing the same 9 ex-ante hypotheses<sup>1</sup>. The flexibility of analytical approaches is exemplified by the fact that no two teams chose identical workflows to analyse the data. This flexibility resulted in sizeable variation in the results of hypothesis tests, even for teams whose statistical maps were highly correlated at intermediate stages of the analysis pipeline. Variation in reported results was related to several aspects of analysis methodology. Notably, a meta-analytical approach that aggregated information across teams yielded a significant consensus in activated regions. Furthermore, prediction markets of researchers in the field revealed an overestimation of the likelihood of significant findings, even by researchers with direct knowledge of the dataset<sup>2–5</sup>. Our findings show that analytical flexibility can have substantial effects on scientific conclusions, and identify factors that may be related to variability in the analysis of functional magnetic resonance imaging. The results emphasize the importance of validating and sharing complex analysis workflows, and demonstrate the need for performing and reporting multiple analyses of the same data. Potential approaches that could be used to mitigate issues related to analytical variability are discussed.

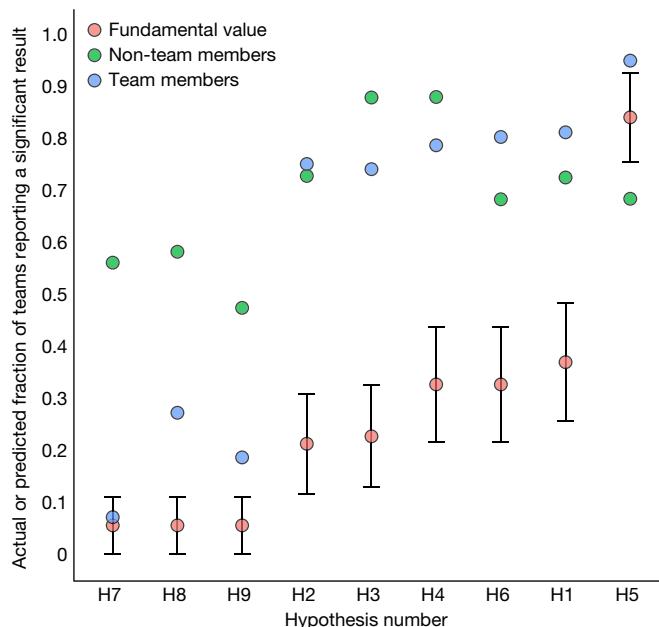
Data analysis workflows in many areas of science have a large number of analysis steps that involve many possible choices (that is, “researcher degrees of freedom”<sup>6,7</sup>). Simulation studies show that variability in analytical choices can have substantial effects on results<sup>8</sup>, but its degree and effect in practice is unclear. Recent work in psychology addressed this through a “many analysts” approach<sup>9</sup>, in which the same dataset was analysed by a large number of groups, uncovering substantial variability in behavioural results across analysis teams. In the Neuroimaging Analysis Replication and Prediction Study (NARPS), we applied a similar approach to the domain of functional magnetic resonance imaging (fMRI), the analysis workflows of which are complex and highly variable. Our goal was to assess—with the highest possible ecological validity—the degree and effect of analytical flexibility on fMRI results in practice. In addition, we estimated the beliefs of researchers in the field regarding the degree of variability in analysis outcomes using prediction markets to test whether peers in the field could predict the results<sup>2–5</sup>.

## Variability of results across teams

The first aim of NARPS was to assess the real-world variability of results across independent teams analysing the same dataset. The dataset included fMRI data from 108 individuals, each performing one of two versions of a task that was previously used to study decision-making under risk<sup>10</sup>. The two versions were designed to address a debate on the effect of gain and loss distributions on neural activity in this task<sup>10–12</sup>. A full description of the dataset is available in a Data Descriptor<sup>1</sup>; the dataset is openly available at <https://doi.org/10.18112/openneuro.ds001734.v1.0.4>.

Seventy teams (69 of whom had previous fMRI publications) were provided with the raw data, and an optional preprocessed version of the dataset (with fMRIprep<sup>13</sup>). They were asked to analyse the data to test nine ex-ante hypotheses (Extended Data Table 1), each consisting of a description of activity in a specific brain region in relation to a particular feature of the task. They were given up to 100 days to report whether each hypothesis was supported on the basis of a whole-brain-corrected analysis (yes or no). In addition, each team submitted a detailed report of the methods of analysis that they had used, together with unthresholded and thresholded statistical maps supporting each hypothesis test (Extended Data Tables 2, 3a). To perform an ecologically valid study testing the sources of variability that contribute to published literature ‘in the wild’, the instructions to the teams were as minimal as possible. The only instructions were to perform the analysis as they usually would in their own research laboratory and report the binary decision on the basis of their own criteria for a whole-brain-corrected result for the specific region described in the hypothesis. The dataset, reports and collections were kept private until after the prediction markets were closed.

Overall, the rates of reported significant findings varied across hypotheses (Fig. 1, Extended Data Table 1). Only one hypothesis (hypothesis 5) showed a high rate of significant findings (84.3%), whereas three other hypotheses showed consistent non-significant findings across teams (5.7% significant findings). For the remaining five hypotheses, the results were variable, with 21.4% to 37.1% of teams reporting a significant result. The extent of the variation in results across teams was quantified by the fraction of teams that reported a result different from the majority of teams (that is, the absolute distance from consensus). On average across the 9 hypotheses, 20% of teams



**Fig. 1 | Fraction of teams reporting a significant result and prediction market beliefs.** The observed fraction of teams reporting significant results (fundamental value, pink dots;  $n = 70$  analysis teams), as well as final market prices for the team members markets (blue dots;  $n = 83$  active traders) and the non-team members markets (green dots;  $n = 65$  active traders). The corresponding 95% confidence intervals are shown for each of the nine hypotheses (note that hypotheses are sorted on the basis of the fundamental value). Confidence intervals were constructed by assuming convergence of the binomial distribution towards the normal.

reported a result that differed from the majority of teams. Given that the maximum possible variation is 50%, the observed fraction of 20% divergent results thus falls midway between complete consistency across teams and completely random results, demonstrating that analytical choices have a major effect on reported results.

## Factors related to analytical variability

To examine the sources of the analytical variability in the reported binary results, we analysed the pipelines used by the teams as well as the unthresholded and thresholded statistical maps they provided. There were no two teams with identical analysis pipelines. After exclusions (Extended Data Table 3b), thresholded maps of 65 teams and unthresholded ( $z$ - or  $t$ -statistic) maps of 64 teams were included in the analyses. Fully reproducible code for all analyses of the data reported here is available at <https://doi.org/10.5281/zenodo.3709273>.

## Variability of reported results

A set of mixed-effects logistic regression models identified several analytical variables and image features that were associated with reported outcomes (Extended Data Table 3c). The strongest factor was spatial smoothness; higher estimated smoothness of the unthresholded statistical maps (estimated using the FMRIB Software Library (FSL) smoothest function) was associated with a greater likelihood of significant outcomes ( $P < 0.001$ , delta pseudo- $R^2 = 0.04$ ; mean full width at half-maximum, 9.69 mm, range 2.50–21.28 mm across teams). Notably, although the estimated smoothness was related to the width of the applied smoothing kernel ( $r = 0.71$ ; median applied smoothing 5 mm, range 0–9 mm across teams), the applied smoothing value itself was not significantly related to positive outcomes in a separate analysis, suggesting that the relevant smoothness arose from analytical

steps beyond explicit smoothing (such as modelling of head motion;  $P = 0.014$ ). An effect on outcomes was also found for the software package used ( $P = 0.004$ , delta pseudo- $R^2 = 0.04$ ;  $n = 23$  (SPM),  $n = 21$  (FSL),  $n = 7$  (AFNI) and  $n = 13$  (other software package))—with FSL being associated with a higher likelihood of significant results across all hypotheses compared to SPM; odds ratio = 6.69—and for the effect of different methods of multiple test correction ( $P = 0.024$ , delta pseudo- $R^2 = 0.02$ ;  $n = 48$  (parametric),  $n = 14$  (nonparametric),  $n = 2$  (other)), with parametric correction methods resulting in higher rates of detection than nonparametric methods. No significant effect was detected for the use of standardized preprocessed data versus custom preprocessing pipelines (48% of included teams used fMRIprep;  $P = 0.132$ ) or for the modelling of head motion parameters (used by 73% of the teams;  $P = 0.281$ ). Nonparametric bootstrap analyses confirmed the significant effect of spatial smoothness, but provided inconsistent support for the effects of multiple testing and software package; because of low power, these results should be interpreted with caution.

## Variability of thresholded statistical maps

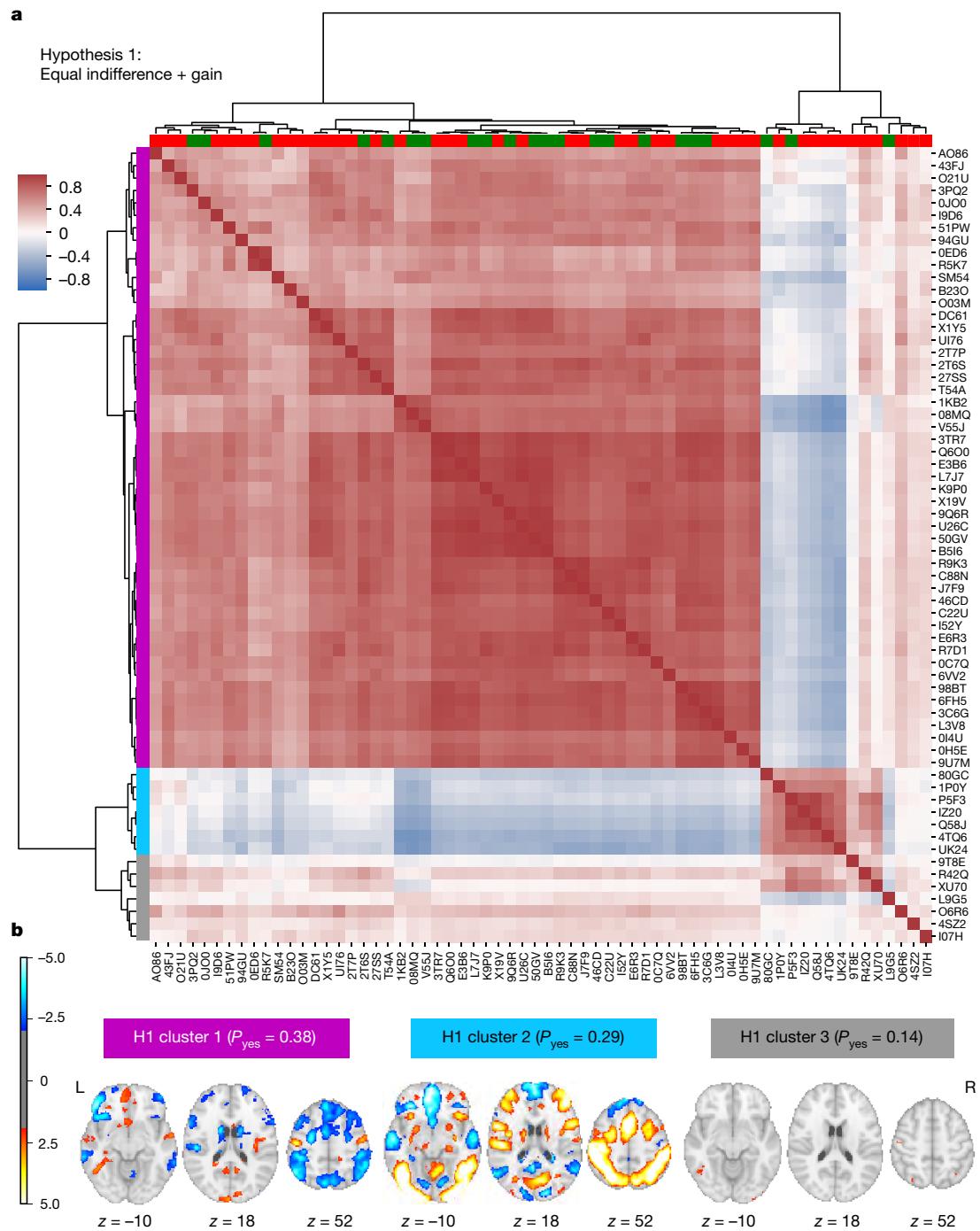
The nature of analytical variability was further explored by analysing the statistical maps that were submitted by the research teams. The thresholded maps were highly sparse. Binary agreement between thresholded maps over all voxels was relatively high (median per cent agreement ranged from 93% to 99% across hypotheses), largely reflecting agreement on which voxels were not active. However, when restricted to voxels showing activation for any team, the overlap was very low (median similarity ranging from 0.00 to 0.06 across hypotheses). This may reflect variability in the number of activated voxels found by each team; for every hypothesis, the number of active voxels ranged across teams from zero to tens of thousands (Extended Data Table 4a). Analysis of the overlap between activated voxels showed that the proportion of teams with activation in the most frequently activated voxel for a given hypothesis ranged between 0.23 and 0.77 (Extended Data Fig. 1).

## Variability of unthresholded statistical maps

Analysis of the correlation between unthresholded  $z$ -statistic maps across teams showed that for each hypothesis, a large cluster of teams had statistical maps that were strongly positively correlated with one another (Fig. 2, Extended Data Fig. 2). The mean Spearman correlation between all pairs of unthresholded maps (Extended Data Table 4b) was moderate (mean correlation range 0.18–0.52 across hypotheses), with higher correlations within the main cluster of analysis teams (range 0.44–0.85 across hypotheses). An analysis of voxelwise heterogeneity across unthresholded maps (equivalent to tau-squared) demonstrated that inter-team variability was large—in many cases several times the variability expected across different datasets (Extended Data Fig. 3a).

For hypotheses 1 and 3, there was a subset of seven teams whose unthresholded maps were anticorrelated with those of the main cluster of teams. A comparison of the average map for the anticorrelated cluster for hypotheses 1 and 3 confirmed that this map was highly correlated ( $r = 0.87$ ) with the overall task-activation map, as previously reported<sup>1</sup>. Further analysis showed that four of these teams used models that did not properly separate the parametric effect of gain from overall task activation; because of the anticorrelation of value-system activations with task activations<sup>14</sup>, this model mis-specification led to an anticorrelation with the parametric effects of gain. In two cases, the model included multiple regressors that were correlated with the gain parameter, which modified the interpretation of the primary gains regressor, and for one additional team, modelling details were not available.

The discrepancy between the overall correlations of unthresholded maps and the divergence of reported binary results (even within the highly correlated cluster) suggested that the variability in regional results might be due to procedures related to statistical correction for multiple comparisons and the subjective decision of teams on the



**Fig. 2 | Analytical variability in whole-brain statistical results for hypothesis 1 (and hypothesis 3).** **a**, Spearman correlation values between whole-brain unthresholded statistical maps for each team ( $n=64$ ) were computed and clustered according to their similarity (using Ward clustering on Euclidean distances). Row colours (left) denote cluster membership (purple, cluster 1; blue, cluster 2; grey, cluster 3); column colours (top) represent hypothesis decisions (green, yes; red, no). Brackets represent clustering.

**b**, Average statistical maps (thresholded at uncorrected  $z>2.0$ ) for each of the three clusters shown on the left in **a**. The probability of reporting a positive hypothesis outcome ( $P_{\text{yes}}$ ) is presented for each cluster. L, left; R, right. Unthresholded maps for hypotheses 1 and 3 are identical (as they both relate to the same contrast and group but different regions), and the colours represent reported results for hypothesis 1. Images can be viewed at <https://identifiers.org/neurovault.collection:6048>.

anatomical specification of regions of interest (ROIs). To test this, we applied a consistent thresholding method and ROI specification on the unthresholded maps across all teams for each hypothesis. This showed that even using a correction method known to be liberal and a standard anatomical definition for all regions, the degree of variability across results was qualitatively similar to that of the actual reported decisions (Extended Data Fig. 4).

We assessed the consistency across teams using an image-based meta-analysis (accounting for correlations due to common data), which demonstrated significant active voxels for all hypotheses except for hypothesis 9 after false discovery rate (FDR) correction (Extended Data Fig. 3b) and confirmatory evidence for hypotheses 2, 4, 5 and 6. These results show that inconsistent results at the individual team level underlie consistent results when the results of teams are combined.

## Prediction markets

The second aim of NARPS was to test whether peers in the field could predict the results, using prediction markets in which researchers trade on the outcomes of scientific analyses and receive monetary payouts based on performance. Prediction markets have been used to assess the replicability of scientific hypotheses in the social sciences, and have revealed correlations between market prices and actual scientific outcomes<sup>2–5</sup>. We performed two separate prediction markets: one involving members from analysis teams ('team members' market) and another independent market for researchers who had not participated in the analysis ('non-team members' market). The markets were open for 10 consecutive days approximately 1.5 months after all analysis teams had submitted their results (which were kept confidential). On each market, traders were provided with tokens worth US\$50, and traded via an online market platform on the fraction of teams that reported a significant result for each hypothesis (that is, the fundamental values). The market prices serve as measures of the aggregate beliefs of traders for the fraction of teams reporting a significant result for each hypothesis. Overall,  $n = 65$  traders actively traded in the non-team members market and  $n = 83$  traded in the team members market. After the markets closed, traders were paid on the basis of their performance in the markets. The analysis of the markets was preregistered on the Open Science Framework (OSF) (<https://osf.io/59ksz/>). Note that because some analyses were performed on the final market prices (that is, the predictions of the markets), for which there is one value per hypothesis per market, the number of observations for each of the markets was low ( $n = 9$ ), leading to limited statistical power. Therefore, the results should be interpreted with caution.

The predictions of the markets ranged from 0.073 to 0.952 ( $m = 0.599$ , s.d. = 0.325) in the team members market and from 0.476 to 0.882 ( $m = 0.690$ , s.d. = 0.137) in the non-team members market. Except for hypothesis 7 in the team members market, all predictions were outside the 95% confidence intervals of the fundamental values (Fig. 1, Extended Data Table 5a). The Spearman correlation between the fundamental values and the predictions of the markets was significant for the team members market ( $r = 0.962$ ,  $P < 0.001$ ,  $n = 9$ ) but not for the non-team members market ( $r = 0.553$ ,  $P = 0.122$ ,  $n = 9$ ), nor between the predictions of both markets ( $r = 0.500$ ,  $P = 0.170$ ,  $n = 9$ ).

Wilcoxon signed-rank tests suggested that traders in both markets systematically overestimated the fundamental values (team members:  $z = 2.886$ ,  $P = 0.004$ ,  $n = 9$ ; non-team members:  $z = 2.660$ ,  $P = 0.008$ ,  $n = 9$ ). The result in the team members market was not driven by an overrepresentation of teams who reported significant results (Supplementary Methods and Supplementary Results). Predictions in the team members market did not significantly differ from those in the non-team members market (Wilcoxon signed-rank test,  $z = 1.035$ ,  $P = 0.301$ ,  $n = 9$ ), but as mentioned above, statistical power for this test was limited. Team members generally traded in the direction consistent with the results of their own team (Extended Data Table 5b), which may explain why their collective predictions were more accurate than those of non-team members (Fig. 1). Additional results are presented in the Supplementary Information (see also Extended Data Fig. 5, Extended Data Table 5).

## Discussion

The analysis of a single fMRI dataset by 70 independent analysis teams, all of whom used different analysis pipelines, revealed substantial variability in reported binary results, with high levels of disagreement across teams for most of the tested hypotheses. For every hypothesis, at least four different analysis pipelines could be found that were used in practice by research groups in the field and resulted in a significant outcome. Our findings highlight the fact that it is hard to estimate the reproducibility of single studies that are performed using a single analysis pipeline. Notably, analyses of the underlying statistical

parametric maps on which the hypothesis tests were based revealed greater consistency than would be expected from those inferences, and significant consensus in activated regions across teams was observed using meta-analysis. Teams with highly correlated underlying unthresholded statistical maps nonetheless reported different hypothesis outcomes (Fig. 2). Detailed analysis of the workflow descriptions and statistical results that were submitted by the analysis teams identified several common analytical variables that were related to differential reporting of significant outcomes, including the spatial smoothness of the data (a result of multiple factors beyond the applied smoothing kernel), the choice of analysis software and the correction method; however, the last two were not consistently supported by nonparametric bootstrap analyses. In addition, we identified model-specification errors for several analysis teams, which led to statistical maps that were anticorrelated with the majority for some of the hypotheses. Prediction markets that were performed on the outcomes of analyses demonstrated a general overestimation by researchers of the likelihood of significant results across hypotheses—even by those researchers who had analysed the data themselves—reflecting a marked optimism bias by researchers in the field.

The substantial amount of analytical variability, and the subsequent variability of reported hypothesis results with the same data, demonstrates that steps need to be taken to improve the reproducibility of data analysis outcomes. First, we suggest that unthresholded statistical maps should be shared as a standard practice alongside thresholded statistical maps using tools such as NeuroVault<sup>15</sup>. In the long run, the shared maps will allow the use of image-based meta-analysis, which we found to provide converging results across laboratories. Second, public sharing of data and analysis code should become common practice, to enable others to run their own analysis with the same data or to validate the code used. These practices, combined with the use of preregistration<sup>16</sup> or registered reports<sup>17</sup>, will reduce researcher degrees of freedom but would not prevent analytical variability, as demonstrated here; however, they would ensure that the effects of variability can be assessed. All of the data and code used in the current study are publicly available with a fully reproducible execution environment for all figures and results. We believe that this can serve as an example for future studies.

Foremost, we propose that complex datasets should be analysed using several analysis pipelines, and preferably by more than one research team. Achieving such 'multiverse analysis' on a large scale will require the development of automated statistical analysis tools (for example, FitLins<sup>18</sup>) that can run a broad range of pipelines and assess their convergence. Different versions of such multiverse analysis have been suggested in other fields<sup>19–21</sup>, but are not widely used. Analysis pipelines should also be validated using simulated data to assess their validity with regard to ground truth, and assessed for their effects on predictions with new data<sup>22</sup>.

Our findings emphasize the urgent need to develop new practices and tools to overcome the challenge of variability across analysis pipelines and its effect on analytical results. Nonetheless, we maintain that fMRI can provide reliable answers to scientific questions, as strongly demonstrated in the meta-analytical results across teams along with numerous large-scale studies in the literature and replication of many findings using fMRI. Moreover, although the present investigation was limited to the analysis of a single fMRI dataset, it seems highly likely that similar variability will be present for other fields of research in which the data are high dimensional and the analysis workflows are complex and varied. The multiverse approach combined with meta-analysis is suggested as a promising solution. Notably, transparent scientific projects that involve community-wide self-assessment—such as this one—are definitive evidence of the awareness of researchers of reproducibility concerns, and the desire to assess their effect and improve practices accordingly (for additional discussion see Supplementary Discussion).

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2314-9>.

1. Botvinik-Nezer, R. et al. fMRI data of mixed gambles from the Neuroimaging Analysis Replication and Prediction Study. *Sci. Data* **6**, 106 (2019).
2. Dreber, A. et al. Using prediction markets to estimate the reproducibility of scientific research. *Proc. Natl Acad. Sci. USA* **112**, 15343–15347 (2015).
3. Camerer, C. F. et al. Evaluating replicability of laboratory experiments in economics. *Science* **351**, 1433–1436 (2016).
4. Camerer, C. F. et al. Evaluating the replicability of social science experiments in *Nature* and *Science* between 2010 and 2015. *Nat. Hum. Behav.* **2**, 637–644 (2018).
5. Forstell, E. et al. Predicting replication outcomes in the Many Labs 2 study. *J. Econ. Psychol.* **75**, 102117 (2019).
6. Wicherts, J. M. et al. Degrees of freedom in planning, running, analyzing, and reporting psychological studies: a checklist to avoid P-hacking. *Front. Psychol.* **7**, 1832 (2016).
7. Simmons, J. P., Nelson, L. D. & Simonsohn, U. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol. Sci.* **22**, 1359–1366 (2011).
8. Carp, J. On the plurality of (methodological) worlds: estimating the analytic flexibility of fMRI experiments. *Front. Neurosci.* **6**, 149 (2012).
9. Silberzahn, R. et al. Many analysts, one data set: making transparent how variations in analytic choices affect results. *Adv. Methods Pract. Psychol. Sci.* **1**, 337–356 (2018).
10. Tom, S. M., Fox, C. R., Trepel, C. & Poldrack, R. A. The neural basis of loss aversion in decision-making under risk. *Science* **315**, 515–518 (2007).
11. De Martino, B., Camerer, C. F. & Adolphs, R. Amygdala damage eliminates monetary loss aversion. *Proc. Natl Acad. Sci. USA* **107**, 3788–3792 (2010).
12. Canessa, N. et al. The functional and structural neural basis of individual differences in loss aversion. *J. Neurosci.* **33**, 14307–14317 (2013).
13. Esteban, O. et al. fMRIprep: a robust preprocessing pipeline for functional MRI. *Nat. Methods* **16**, 111–116 (2019).
14. Acikalin, M. Y., Gorgolewski, K. J. & Poldrack, R. A. A coordinate-based meta-analysis of overlaps in regional specialization and functional connectivity across subjective value and default mode networks. *Front. Neurosci.* **11**, 1 (2017).
15. Gorgolewski, K. J. et al. NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Front. Neuroinform.* **9**, 8 (2015).
16. Nosek, B. A., Ebersole, C. R., DeHaven, A. C. & Mellor, D. T. The preregistration revolution. *Proc. Natl Acad. Sci. USA* **115**, 2600–2606 (2018).
17. Nosek, B. A. & Lakens, D. Registered reports: a method to increase the credibility of published results. *Soc. Psychol.* **45**, 137–141 (2014).
18. Markiewicz, C., De La Vega, A., Yarkoni, T., Poldrack, R. & Gorgolewski, K. FitLins: reproducible model estimation for fMRI. Poster W621 in *25th Annual Meeting of the Organization for Human Brain Mapping* (OHBM, 2019).
19. Simonsohn, U., Simmons, J. P. & Nelson, L. D. Specification curve: descriptive and inferential statistics on all reasonable specifications. <https://doi.org/10.2139/ssrn.2694998> (2015).
20. Patel, C. J., Burford, B. & Ioannidis, J. P. A. Assessment of vibration of effects due to model specification can demonstrate the instability of observational associations. *J. Clin. Epidemiol.* **68**, 1046–1058 (2015).
21. Steegen, S., Tuerlinckx, F., Gelman, A. & Vanpaemel, W. Increasing transparency through a multiverse analysis. *Perspect. Psychol. Sci.* **11**, 702–712 (2016).
22. LaConte, S. et al. The evaluation of preprocessing choices in single-subject BOLD fMRI using NPAIRS performance metrics. *Neuroimage* **18**, 10–27 (2003).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

# Article

- Rotem Botvinik-Nezer<sup>1,2,3</sup>, Felix Holzmeister<sup>4</sup>, Colin F. Camerer<sup>5</sup>, Anna Dreber<sup>6,7</sup>, Juergen Huber<sup>8</sup>, Magnus Johannesson<sup>9</sup>, Michael Kirchler<sup>10</sup>, Roni Iwanir<sup>11</sup>, Jeanette A. Mumford<sup>9</sup>, R. Alison Adcock<sup>11,10</sup>, Paolo Avesani<sup>11,12</sup>, Blazej M. Baczkowski<sup>13</sup>, Aahana Bjacharya<sup>14</sup>, Leah Bakst<sup>15,16</sup>, Sheryl Ball<sup>17,18</sup>, Marco Barilaro<sup>19</sup>, Nadège Bault<sup>20</sup>, Derek Beaton<sup>21</sup>, Julia Beitner<sup>22,23</sup>, Roland G. Benoit<sup>24</sup>, Ruud M. W. J. Berkers<sup>24</sup>, Jamil P. Bhanji<sup>25</sup>, Bharat B. Biswal<sup>26,27</sup>, Sebastian Bobadilla-Suarez<sup>28</sup>, Tiago Bortolini<sup>29</sup>, Katherine L. Bottenhorn<sup>30</sup>, Alexander Bowring<sup>31</sup>, Senne Braem<sup>32,33</sup>, Hayley R. Brooks<sup>34</sup>, Emily G. Brudner<sup>25</sup>, Cristian B. Calderon<sup>32</sup>, Julia A. Camilleri<sup>35,36</sup>, Jaime J. Castrellon<sup>9,37</sup>, Luca Cecchetti<sup>38</sup>, Edna C. Cieslik<sup>35,36</sup>, Zachary J. Cole<sup>39</sup>, Olivier Collignon<sup>12,19</sup>, Robert W. Cox<sup>40</sup>, William A. Cunningham<sup>41</sup>, Stefan Czoschke<sup>42</sup>, Kamalaker Dadi<sup>43</sup>, Charles P. Davis<sup>44,45,46</sup>, Alberto De Luca<sup>47</sup>, Mauricio R. Delgado<sup>25</sup>, Lysia Demetriou<sup>48,49</sup>, Jeffrey B. Dennison<sup>50</sup>, Xin Di<sup>26,27</sup>, Erin W. Dickie<sup>51,52</sup>, Ekaterina Dobryakova<sup>53</sup>, Claire L. Donnat<sup>54</sup>, Juergen Dukart<sup>35,36</sup>, Niall W. Duncan<sup>55,56</sup>, Joke Durnez<sup>7</sup>, Ann Eed<sup>58</sup>, Simon B. Eickhoff<sup>35,36</sup>, Andrew Erhart<sup>34</sup>, Laura Fontanesi<sup>59</sup>, G. Matthew Fricke<sup>60</sup>, Shiguang Fu<sup>61,62</sup>, Adriana Galván<sup>63</sup>, Remi Gau<sup>19</sup>, Sarah Genon<sup>35,36</sup>, Tristan Glatard<sup>64</sup>, Enrico Gleerean<sup>65</sup>, Jelle J. Goeman<sup>66</sup>, Sergej A. E. Golowin<sup>55</sup>, Carlos González-García<sup>32</sup>, Krzysztof J. Gorgolewski<sup>67</sup>, Cheryl L. Grady<sup>21</sup>, Mikella A. Green<sup>9,37</sup>, João F. Guassi Moreira<sup>68</sup>, Olivia Guest<sup>28,69</sup>, Shabnam Hakimi<sup>19</sup>, J. Paul Hamilton<sup>69</sup>, Roeland Hancock<sup>45,46</sup>, Giacomo Handjara<sup>38</sup>, Bronson B. Harry<sup>70</sup>, Colin Hawco<sup>71</sup>, Peer Herholz<sup>72</sup>, Gabrielle Herman<sup>71</sup>, Stephan Heinus<sup>37,74</sup>, Felix Hoffstaedter<sup>35,36</sup>, Jeremy Hogeveen<sup>75,76</sup>, Susan Holmes<sup>54</sup>, Chuan-Peng Hu<sup>77</sup>, Scott A. Huettel<sup>37</sup>, Matthew E. Hughes<sup>78</sup>, Vittorio Iacobella<sup>12</sup>, Alexandru D. Iordan<sup>79</sup>, Peder M. Isager<sup>80</sup>, Ayse I. Isik<sup>81</sup>, Andrew Jahn<sup>82</sup>, Matthew R. Johnson<sup>39,83</sup>, Tom Johnstone<sup>78</sup>, Michael J. E. Joseph<sup>71</sup>, Anthony C. Juliano<sup>84</sup>, Joseph W. Kable<sup>55,86</sup>, Michalis Kassinopoulos<sup>87</sup>, Cemal Koba<sup>38</sup>, Xiang-Zhen Kong<sup>88</sup>, Timothy R. Koscielny<sup>89</sup>, Nuri Ertuk Kucukboyaci<sup>53,90</sup>, Brice A. Kuh<sup>91</sup>, Sebastian Kupek<sup>92</sup>, Angela R. Laird<sup>93</sup>, Claus Lamm<sup>94,95</sup>, Robert Langner<sup>35,36</sup>, Nina Lauharatanahirun<sup>96,97</sup>, Hongmi Lee<sup>98</sup>, Sangil Lee<sup>85</sup>, Alexander Leemans<sup>47</sup>, Andrea Leo<sup>38</sup>, Elise Lesage<sup>32</sup>, Flora Li<sup>99,100</sup>, Monica Y. C. Li<sup>44,45,46,101</sup>, Phui Cheng Lim<sup>39,83</sup>, Evan N. Lintz<sup>39</sup>, Schuyler W. Liphardt<sup>102</sup>, Annabel B. Losecaat Vermeer<sup>94</sup>, Bradley C. Love<sup>28,103</sup>, Michael L. Mack<sup>41</sup>, Norberto Malpica<sup>104</sup>, Theo Marins<sup>29</sup>, Camille Maumet<sup>105</sup>, Kelsey McDonald<sup>37</sup>, Joseph T. McGuire<sup>15,16</sup>, Helena Melero<sup>104,106,107</sup>, Adriana S. Méndez Leal<sup>63</sup>, Benjamin Meyer<sup>77,108</sup>, Kristin N. Meyer<sup>109</sup>, Glad Mihi<sup>110,111</sup>, Georgios D. Mitsis<sup>112</sup>, Jorge Mol<sup>29,67</sup>, Dylan M. Nielson<sup>113</sup>, Gustav Nilsson<sup>114,115</sup>, Michael P. Notter<sup>116</sup>, Emanuele Olivetti<sup>11,12</sup>, Adrian I. Onicas<sup>38</sup>, Paolo Papale<sup>38,117</sup>, Kaustubh R. Patil<sup>35,36</sup>, Jonathan E. Peelle<sup>14</sup>, Alexandre Pérez<sup>72</sup>, Doris Pischchedda<sup>118,119,120</sup>, Jean-Baptiste Poline<sup>72,121</sup>, Yanina Prystaika<sup>44,45,46</sup>, Shruti Ray<sup>26</sup>, Patricia A. Reuter-Lorenz<sup>79</sup>, Richard C. Reynolds<sup>122</sup>, Emiliano Ricciardi<sup>38</sup>, Jenny R. Rieck<sup>21</sup>, Anais M. Rodriguez-Thompson<sup>109</sup>, Anthony Romyn<sup>41</sup>, Taylor Salo<sup>30</sup>, Gregory R. Samanez-Larkin<sup>9,37</sup>, Emilio Sanz-Morales<sup>104</sup>, Margaret L. Schlichting<sup>41</sup>, Douglas H. Schultz<sup>39,83</sup>, Qiang Shen<sup>61,62</sup>, Margaret A. Sheridan<sup>109</sup>, Jennifer A. Silvers<sup>63</sup>, Kenny Skagerlund<sup>123,124</sup>, Alec Smith<sup>718</sup>, David V. Smith<sup>50</sup>, Peter Sokol-Hessner<sup>34</sup>, Simon R. Steinkamp<sup>125</sup>, Sarah M. Tashjian<sup>63</sup>, Bertrand Thirion<sup>43</sup>, John N. Thorp<sup>12</sup>, Gustav Tinghög<sup>127,128</sup>, Loreen Tisdall<sup>67,129</sup>, Steven H. Tompson<sup>96</sup>, Claudio Toro-Serey<sup>15,16</sup>, Juan Jesus Torre Tresols<sup>43</sup>, Leonardo Tozzi<sup>130</sup>, Vuong Truong<sup>55,56</sup>, Luca Turella<sup>12</sup>, Anna E. van 't Veer<sup>131</sup>, Tom Verguts<sup>32</sup>, Jean M. Vettel<sup>132,133,134</sup>, Sagana Vijayarajah<sup>41</sup>, Khoi Vo<sup>9,37</sup>, Matthew B. Wall<sup>135,136,137</sup>, Wouter D. Weeda<sup>13</sup>, Susanne Weis<sup>35,36</sup>, David J. White<sup>138</sup>, David Wisniewski<sup>32</sup>, Alba Xifra-Porras<sup>87</sup>, Emily A. Yearling<sup>44,45,46</sup>, Sangsuk Yoon<sup>139</sup>, Rui Yuan<sup>130</sup>, Kenneth S. L. Yuen<sup>77,108</sup>, Lei Zhang<sup>94</sup>, Xu Zhang<sup>45,46,140</sup>, Joshua E. Zosky<sup>39,83</sup>, Thomas E. Nichols<sup>31</sup>, Russell A. Poldrack<sup>67,53</sup> & Tom Schonberg<sup>1,2</sup>.
- <sup>1</sup>Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel. <sup>2</sup>Department of Neurobiology, The George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel. <sup>3</sup>Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA. <sup>4</sup>Department of Banking and Finance, University of Innsbruck, Innsbruck, Austria. <sup>5</sup>HSS and CNS, California Institute of Technology, Pasadena, CA, USA. <sup>6</sup>Department of Economics, Stockholm School of Economics, Stockholm, Sweden. <sup>7</sup>Department of Economics, University of Innsbruck, Innsbruck, Austria. <sup>8</sup>Center for Healthy Minds, University of Wisconsin–Madison, Madison, WI, USA. <sup>9</sup>Center for Cognitive Neuroscience, Duke University, Durham, NC, USA. <sup>10</sup>Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC, USA. <sup>11</sup>Neuroinformatics Laboratory, Fondazione Bruno Kessler, Trento, Italy. <sup>12</sup>Center for Mind/Brain Sciences - CIMEC, University of Trento, Rovereto, Italy. <sup>13</sup>Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. <sup>14</sup>Department of Otolaryngology, Washington University in St. Louis, St. Louis, MO, USA. <sup>15</sup>Department of Psychological and Brain Sciences, Boston University, Boston, MA, USA. <sup>16</sup>Center for Systems Neuroscience, Boston University, Boston, MA, USA. <sup>17</sup>Department of Economics, Virginia Tech, Blacksburg, VA, USA. <sup>18</sup>School of Neuroscience, Virginia Tech, Blacksburg, VA, USA. <sup>19</sup>Crossmodal Perception and Plasticity Laboratory, Institutes for Research in Psychology (IPSY) and Neurosciences (IoNS), UCLouvain, Louvain-la-Neuve, Belgium. <sup>20</sup>School of Psychology, University of Plymouth, Plymouth, UK. <sup>21</sup>Rotman Research Institute, Baycrest Health Sciences Centre, Toronto, Ontario, Canada. <sup>22</sup>Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands. <sup>23</sup>Department of Psychology, Goethe University, Frankfurt am Main, Germany. <sup>24</sup>Max Planck Research Group: Adaptive Memory, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. <sup>25</sup>Department of Psychology, Rutgers University-Newark, Newark, NJ, USA. <sup>26</sup>Department of Biomedical Engineering, New Jersey Institute of Technology, Newark, NJ, USA. <sup>27</sup>School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, China. <sup>28</sup>Department of Experimental Psychology, University College London, London, UK. <sup>29</sup>D'Or Institute for Research and Education (IDOR), Rio de Janeiro, Brazil. <sup>30</sup>Department of Psychology, Florida International University, Miami, FL, USA. <sup>31</sup>Oxford Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, Oxford, UK. <sup>32</sup>Department of Experimental Psychology, Ghent University, Ghent, Belgium. <sup>33</sup>Department of Psychology, Vrije Universiteit Brussel, Brussels, Belgium. <sup>34</sup>Department of Psychology, University of Denver, Denver, CO, USA. <sup>35</sup>Institute of Neuroscience and Medicine, Brain and Behaviour (INM-7), Research Centre Jülich, Jülich, Germany. <sup>36</sup>Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. <sup>37</sup>Department of Psychology and Neuroscience, Duke University, Durham, NC, USA. <sup>38</sup>MoMiLab Research Unit, IMT School for Advanced Studies Lucca, Lucca, Italy. <sup>39</sup>Department of Psychology, University of Nebraska-Lincoln, Lincoln, NE, USA. <sup>40</sup>National Institute of Mental Health (NIMH), National Institutes of Health, Bethesda, MD, USA. <sup>41</sup>Department of Psychology, University of Toronto, Toronto, Ontario, Canada. <sup>42</sup>Institute of Medical Psychology, Goethe University, Frankfurt am Main, Germany. <sup>43</sup>Inria, CEA, Université Paris-Saclay, Palaiseau, France. <sup>44</sup>Department of Psychological Sciences, University of Connecticut, Storrs, CT, USA. <sup>45</sup>Brain Imaging Research Center, University of Connecticut, Storrs, CT, USA. <sup>46</sup>Connecticut Institute for the Brain and Cognitive Sciences, University of Connecticut, Storrs, CT, USA. <sup>47</sup>PROVIDI Lab, Image Sciences Institute, University Medical Center Utrecht, Utrecht, The Netherlands. <sup>48</sup>Section of Endocrinology and Investigative Medicine, Faculty of Medicine, Imperial College London, London, UK. <sup>49</sup>Nuffield Department of Women's and Reproductive Health, University of Oxford, Oxford, UK. <sup>50</sup>Department of Psychology, Temple University, Philadelphia, PA, USA. <sup>51</sup>Krembil Centre for Neuroinformatics, Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Ontario, Canada. <sup>52</sup>Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada. <sup>53</sup>Center for Traumatic Brain Injury Research, Kessler Foundation, East Hanover, NJ, USA. <sup>54</sup>Department of Statistics, Stanford University, Stanford, CA, USA. <sup>55</sup>Graduate Institute of Mind, Brain and Consciousness, Taipei Medical University, Taipei, Taiwan. <sup>56</sup>Brain and Consciousness Research Centre, TMU-ShuangHo Hospital, New Taipei City, Taiwan. <sup>57</sup>Department of Psychology and Stanford Center for Reproducible Neuroscience, Stanford University, Stanford, CA, USA. <sup>58</sup>Instituto de Neurociencias, CSIC-UMH, Alicante, Spain. <sup>59</sup>Faculty of Psychology, University of Basel, Basel, Switzerland. <sup>60</sup>Computer Science Department, University of New Mexico, Albuquerque, NM, USA. <sup>61</sup>School of Management, Zhejiang University of Technology, Hangzhou, China. <sup>62</sup>Institute of Neuromanagement, Zhejiang University of Technology, Hangzhou, China. <sup>63</sup>Department of Psychology, University of California Los Angeles, Los Angeles, CA, USA. <sup>64</sup>Department of Computer Science and Software Engineering, Concordia University, Montreal, Quebec, Canada. <sup>65</sup>Department of Neuroscience and Biomedical Engineering, Aalto University, Espoo, Finland. <sup>66</sup>Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands. <sup>67</sup>Department of Psychology, Stanford University, Stanford, CA, USA. <sup>68</sup>Research Centre on Interactive Media, Smart Systems and Emerging Technologies - RISE, Nicosia, Cyprus. <sup>69</sup>Center for Social and Affective Neuroscience, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden. <sup>70</sup>The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Sydney, New South Wales, Australia. <sup>71</sup>Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Ontario, Canada. <sup>72</sup>McConnell Brain Imaging Centre, The Neuro (Montreal Neurological Institute-Hospital), Faculty of Medicine, McGill University, Montreal, Quebec, Canada. <sup>73</sup>Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands. <sup>74</sup>Department of Research and Development, Epilepsy Centre Kempenhaeghe, Heeze, The Netherlands. <sup>75</sup>Department of Psychology, University of New Mexico, Albuquerque, NM, USA. <sup>76</sup>Psychology Clinical Neuroscience Center, University of New Mexico, Albuquerque, NM, USA. <sup>77</sup>Leibniz-Institut für Resilienzforschung (LIR), Mainz, Germany. <sup>78</sup>School of Health Sciences, Swinburne University of Technology, Hawthorn, Victoria, Australia. <sup>79</sup>Department of Psychology, University of Michigan, Ann Arbor, MI, USA. <sup>80</sup>Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, Eindhoven, The Netherlands. <sup>81</sup>Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany. <sup>82</sup>fMRI Laboratory, University of Michigan, Ann Arbor, MI, USA. <sup>83</sup>Center for Brain, Biology and Behavior, University of Nebraska-Lincoln, Lincoln, NE, USA. <sup>84</sup>Center for Neuropsychology and Neuroscience Research, Kessler Foundation, East Hanover, NJ, USA. <sup>85</sup>Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA. <sup>86</sup>MindCORE, University of Pennsylvania, Philadelphia, PA, USA. <sup>87</sup>Graduate Program in Biological and Biomedical Engineering, McGill University, Montreal, Quebec, Canada. <sup>88</sup>Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands. <sup>89</sup>Department of Psychiatry, University of Iowa Carver College of Medicine, Iowa City, IA, USA. <sup>90</sup>Department of Physical Medicine and Rehabilitation, Rutgers New Jersey Medical School, Newark, NJ, USA. <sup>91</sup>Department of Psychology, University of Oregon, Eugene, OR, USA. <sup>92</sup>Faculty of Economics and Statistics, University of Innsbruck, Innsbruck, Austria. <sup>93</sup>Department of Physics, Florida International University, Miami, Florida, USA. <sup>94</sup>Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria. <sup>95</sup>Vienna Cognitive Science Hub, University of Vienna, Vienna, Austria. <sup>96</sup>US CCDC Army Research Laboratory, Human Research and Engineering Directorate, Aberdeen Proving Ground, MD, USA. <sup>97</sup>Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA, USA. <sup>98</sup>Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD, USA. <sup>99</sup>Fralin Biomedical Research Institute, Roanoke, VA, USA. <sup>100</sup>Economics Experimental Lab, Nanjing Audit University, Nanjing, China. <sup>101</sup>Haskins Laboratories, New Haven, CT, USA. <sup>102</sup>Biology Department, University of New Mexico, Albuquerque, NM, USA. <sup>103</sup>The Alan Turing Institute, London, UK. <sup>104</sup>Laboratorio de Análisis de Imagen Médica y Biométrica (LAIMBIO), Universidad Rey Juan Carlos, Madrid, Spain. <sup>105</sup>Inria, Univ Rennes, CNRS, Inserm, IRISA UMR 6074, Empenn ERL U 1228, Rennes, France. <sup>106</sup>Departamento de Psicobiología, División de

<sup>1</sup>Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel. <sup>2</sup>Department of Neurobiology, The George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel. <sup>3</sup>Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA. <sup>4</sup>Department of Banking and Finance, University of Innsbruck, Innsbruck, Austria. <sup>5</sup>HSS and CNS, California Institute of Technology, Pasadena, CA, USA. <sup>6</sup>Department of Economics, Stockholm School of Economics, Stockholm, Sweden. <sup>7</sup>Department of Economics, University of Innsbruck, Innsbruck, Austria. <sup>8</sup>Center for Healthy Minds, University of Wisconsin–Madison, Madison, WI, USA. <sup>9</sup>Center for Cognitive Neuroscience, Duke University, Durham, NC, USA. <sup>10</sup>Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC, USA. <sup>11</sup>Neuroinformatics Laboratory, Fondazione Bruno Kessler, Trento, Italy. <sup>12</sup>Center for Mind/Brain Sciences - CIMEC, University of Trento, Rovereto, Italy. <sup>13</sup>Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. <sup>14</sup>Department of Otolaryngology, Washington University in St. Louis, St. Louis, MO, USA. <sup>15</sup>Department of Psychological and Brain Sciences, Boston University, Boston, MA, USA. <sup>16</sup>Center for Systems Neuroscience, Boston University, Boston, MA, USA. <sup>17</sup>Department of Economics, Virginia Tech, Blacksburg, VA, USA. <sup>18</sup>School of Neuroscience, Virginia Tech, Blacksburg, VA, USA. <sup>19</sup>Crossmodal Perception and Plasticity Laboratory, Institutes for Research in Psychology (IPSY) and Neurosciences (IoNS), UCLouvain, Louvain-la-Neuve, Belgium. <sup>20</sup>School of Psychology, University of Plymouth, Plymouth, UK. <sup>21</sup>Rotman Research Institute, Baycrest Health Sciences Centre, Toronto, Ontario, Canada. <sup>22</sup>Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands. <sup>23</sup>Department of Psychology, Goethe University, Frankfurt am Main, Germany. <sup>24</sup>Max Planck Research Group: Adaptive Memory, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. <sup>25</sup>Department of Psychology, Rutgers University-Newark, Newark, NJ, USA. <sup>26</sup>Department of Biomedical Engineering, New Jersey Institute of Technology, Newark, NJ, USA. <sup>27</sup>School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, China. <sup>28</sup>Department of Experimental Psychology, University College London, London, UK. <sup>29</sup>D'Or Institute for Research and Education (IDOR), Rio de Janeiro, Brazil. <sup>30</sup>Department of Psychology, Florida International University, Miami, FL, USA. <sup>31</sup>Oxford Big Data Institute, Li Ka

Psicología, CES Cardenal Cisneros, Madrid, Spain. <sup>107</sup>Northeastern University Biomedical Imaging Center, Northeastern University, Boston, MA, USA. <sup>108</sup>Neuroimaging Center (NIC), Focus Program Translational Neurosciences (FTN), Johannes Gutenberg University Medical Center Mainz, Mainz, Germany. <sup>109</sup>Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. <sup>110</sup>Max Planck Research Group: Neural Mechanisms of Human Communication, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. <sup>111</sup>Chair of Cognitive and Clinical Neuroscience, Faculty of Psychology, Technische Universität Dresden, Dresden, Germany. <sup>112</sup>Department of Bioengineering, McGill University, Montreal, Quebec, Canada. <sup>113</sup>Data Science and Sharing Team, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA. <sup>114</sup>Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden. <sup>115</sup>Department of Psychology, Stockholm University, Stockholm, Sweden. <sup>116</sup>The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology, University Hospital Center and University of Lausanne, Lausanne, Switzerland. <sup>117</sup>Department of Vision and Cognition, Netherlands Institute for Neuroscience, Amsterdam, The Netherlands. <sup>118</sup>Bernstein Center for Computational Neuroscience and Berlin Center for Advanced Neuroimaging and Clinic for Neurology, Charité Universitätsmedizin, corporate member of Freie Universität Berlin, Humboldt Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany. <sup>119</sup>Cluster of Excellence Science of Intelligence, Technische Universität Berlin and Humboldt Universität zu Berlin, Berlin, Germany. <sup>120</sup>NeuroMI - Milan Center for Neuroscience, Milan, Italy. <sup>121</sup>Henry H. Wheeler, Jr. Brain Imaging Center, Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley, CA, USA. <sup>122</sup>Scientific and Statistical Computing Core, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA. <sup>123</sup>Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden. <sup>124</sup>Center for Social and Affective Neuroscience, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden. <sup>125</sup>Institute of Neuroscience and Medicine, Cognitive Neuroscience (INM-3), Research Centre Jülich, Jülich, Germany. <sup>126</sup>Department of Psychology, Columbia University, New York, NY, USA. <sup>127</sup>Department of Management and Engineering, Linköping University, Linköping, Sweden. <sup>128</sup>Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden. <sup>129</sup>Center for Cognitive and Decision Sciences, University of Basel, Basel, Switzerland. <sup>130</sup>Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA. <sup>131</sup>Methodology and Statistics Unit, Institute of Psychology, Leiden University, Leiden, The Netherlands. <sup>132</sup>US Combat Capabilities Development Command Army Research Laboratory, Aberdeen, MD, USA. <sup>133</sup>University of California Santa Barbara, Santa Barbara, CA, USA. <sup>134</sup>University of Pennsylvania, Philadelphia, PA, USA. <sup>135</sup>Invicro, London, UK. <sup>136</sup>Faculty of Medicine, Imperial College London, London, UK. <sup>137</sup>Clinical Psychopharmacology Unit, University College London, London, UK. <sup>138</sup>Centre for Human Psychopharmacology, Swinburne University, Hawthorn, Victoria, Australia. <sup>139</sup>Department of Management and Marketing, School of Business, University of Dayton, Dayton, OH, USA. <sup>140</sup>Biomedical Engineering Department, University of Connecticut, Storrs, CT, USA. <sup>✉</sup>e-mail: thomas.nichols@bdi.ox.ac.uk; poldrack@stanford.edu; schonberg@tauex.tau.ac.il

# Article

## Methods

### Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

### fMRI dataset

To test the variability of neuroimaging results across analysis pipelines used in practice in research laboratories, we distributed a single fMRI dataset to independent analysis groups from around the world, requesting them to test nine predefined hypotheses. The full dataset is publicly available in the Brain Imaging Data Structure (BIDS)<sup>23</sup> on OpenNeuro (<https://doi.org/10.18112/openneuro.ds001734.v1.0.4>) and is described in detail in a Data Descriptor<sup>1</sup>.

In brief, the fMRI dataset consisted of data from 108 participants who performed a mixed gambles task, which is often used to study decision-making under risk. In this task, participants are asked on each trial to accept or reject a presented prospect. The prospects consist of an equal 50% chance of either gaining a given amount of money or losing another, similar or different, amount of money. Participants were divided into two groups: in the ‘equal indifference’ group ( $n = 54$ ) the potential losses were half the size of the potential gains<sup>10</sup> (reflecting the ‘loss aversion’ phenomenon, in which people tend to be more sensitive to losses than to equal-sized gains<sup>24</sup>); and in the ‘equal range’ group ( $n = 54$ ) the potential losses and the potential gains were taken from the same scale<sup>11,12</sup>. The two groups were used to resolve inconsistencies of previous published results.

The dataset was distributed to the teams via Globus (<https://www.globus.org/>). The distributed dataset included raw data of 108 participants ( $n = 54$  for each experimental group), as well as the same data after preprocessing with fMRIprep v.1.1.4 (RRID: SCR\_016216)<sup>13</sup>. The fMRIprep preprocessing mainly included brain extraction, spatial normalization, surface reconstruction, head motion estimation and susceptibility distortion correction. Both the raw and the preprocessed datasets underwent quality assurance (described in detail in the Data Descriptor<sup>1</sup>).

MRI data collection was approved by the Helsinki committee at Sheba Tel Hashomer Medical Center and the ethics committee at Tel Aviv University, and all participants gave written informed consent (as described in the Data Descriptor of this dataset<sup>1</sup>). The Board for Ethical Questions in Science at the University of Innsbruck approved the data collection in the prediction markets, and certified that the project complied with all requirements of the ethical principles and guidelines of good scientific practice. The Stanford University Institutional Review Board (IRB) determined that the analysis of the submitted team results did not meet the definition of human subject research, and thus no further IRB review was required. We have complied with all relevant ethical regulations.

### Predefined hypotheses

Previous studies with the mixed gambles task suggested that activity in the ventromedial prefrontal cortex and ventral striatum, among other brain regions, is related to the magnitude of the potential gain<sup>10</sup>. A fundamental open question in the field of decision-making under risk is whether the magnitude of the potential loss is coded by the same brain regions (through negative activation), or by regions related to negative emotions, such as the amygdala<sup>10–12</sup>. The specific hypotheses included in NARPS were chosen to address this open question, using two different designs that were used in those previous studies (that is, equal indifference versus equal range). Each analysis team tested the same nine predefined hypotheses (Extended Data Table 1). Each hypothesis predicted fMRI activations in a specific brain region, in relation to a specific aspect of the task (gain or loss amount) and a specific group (equal indifference or equal range, or a comparison between the two

groups). Therefore, for each hypothesis, the maximum sample size was 54 participants (hypotheses 1–8) or 54 participants per group in the group comparison (hypothesis 9). Although the hypotheses referred to specific brain regions, analysis teams were instructed to report their results on the basis of a whole-brain analysis (not an ROI-based analysis, as is sometimes used in fMRI studies).

### Recruitment of and instructions to analysis teams

We recruited analysis teams via social media, mainly Twitter and Facebook, as well as during the 2018 annual meeting of the Society for Neuroeconomics. Ninety-seven teams registered to participate in the study. Each team consisted of up to three members. To ensure independent analyses across teams, and to prevent influencing the subsequent prediction markets, all team members signed an electronic nondisclosure agreement that they would not release, publicize or discuss their results with anyone until the end of the study. All team members of 82 teams signed the nondisclosure form. They were offered co-authorship on the present publication in return for their participation.

Analysis teams were provided with access to the full dataset. They were asked to freely analyse the data with their usual analysis pipeline to test the nine hypotheses and report a binary decision for each hypothesis on whether it was significantly supported on the basis of a whole-brain analysis. Although the hypotheses were region-specific, we clearly requested a whole-brain analysis result to avoid the need of teams to create and share masks of ROIs. Each team also filled in a full report of the analysis methods used (following the guidelines of the Committee on Best Practices in Data Analysis and Sharing; COBIDAS<sup>25</sup>) and created a collection on NeuroVault<sup>15</sup> (RRID: SCR\_003806) with one unthresholded and one thresholded statistical map for each hypothesis, on which their decisions were based (teams could optionally include additional maps in their collection; see Extended Data Table 3a for links for collections). For each result (that is, the binary decision on whether a given hypothesis was supported by the data or not), teams further reported how confident they were in this result and how similar they thought their result was to the results of the other teams (each measure was an integer between 1 (not at all) to 10 (extremely)). These measures are presented in Extended Data Tables 1, 2. To measure the variability of results in an ecological manner, instructions to the analysis teams were minimized and the teams were asked to perform the analysis as they usually would in their own laboratory and to report the binary decision on the basis of their own criteria.

Seventy of the 82 teams submitted their results and reports by the final deadline (15 March 2019; overall teams were given up to 100 days, varying based on the date they joined, to complete and report their analysis). The dataset, reports and collections were kept private until the end of the study and closure of the prediction markets. To avoid identification of the teams, each team was provided with a unique random four-character ID.

Overall, 180 participants were part of NARPS analysis teams. Out of 70 analysis teams, 5 teams consisted of 1 member, 20 teams consisted of 2 members and 45 teams consisted of 3 members. Out of the 180 team members, there were 62 principal investigators, 43 post-doctoral researchers, 53 graduate students and 22 members from other positions (for example, data scientists or research analysts).

### Factors related to analytical variability

To explore the factors related to the variability in results across teams, the reports of all teams were manually annotated to create a table describing the methods used by each team. Code for all analyses of the reports and statistical maps submitted by the analysis teams is openly shared in GitHub (<https://github.com/poldrack/narps>). Analyses reported in this manuscript were performed using code release v.2.0.3 (<https://doi.org/10.5281/zenodo.3709273>). We performed exploratory analyses of the relation between the reported hypothesis outcomes and several analytical choices and image features using mixed-effects

logistic regression models implemented in R, with the lme4 package<sup>26</sup>. The factors included in the model were: hypothesis number, estimated smoothness (based on the smoothest function in FSL), use of standar-dized preprocessing, software package, method of correction for multiple comparisons and modelling of head movement. The teams were modelled as a random effect. One team submitted results that were not based on a whole-brain analysis as requested, and therefore their data were excluded from all analyses.

Inferences using logistic regression models were confirmed using nonparametric bootstrap analysis, resampling data team-wise to maintain random effect structure. For the continuous or binary regressors (smoothness, movement modelling and use of fMRIPrep data), we computed bootstrap confidence intervals and, as an approximate hypothesis test, tested whether the confidence interval includes zero. For the factorial variables (hypothesis, software package and multiple testing method), this was not possible because there is not a single coefficient for the factor; in addition, for software package and multiple testing methods, some bootstrap samples did not contain all values of the factor. For these variables we instead performed model comparison between the full model and a reduced model excluding each factor, and computed the proportion of times the full model was selected on the basis of the model selection criterion (using both Bayesian information criterion and Akaike information criterion) being numerically lower in the full model<sup>27</sup>.

In addition, we performed exploratory analyses to examine the variability across statistical maps submitted by the teams. The unthresholded and thresholded statistical maps of all teams were resampled to common space (FSL MNI space, 91 × 109 × 91, 2 mm isotropic) using nilearn<sup>28</sup> (RRID: SCR\_001362). For unthresholded maps, we used third-order spline interpolation; for thresholded maps, we used linear interpolation and then thresholded at 0.5, to prevent artefacts that appeared when using nearest neighbour interpolation. Of the 69 teams included in the analyses, unthresholded maps of 5 teams and thresholded maps of 4 teams were excluded from the image-based analyses (see Extended Data Table 3b for details). As some of the hypotheses reflected negative activations—which can be represented by either positive or negative values in the statistical maps, depending on the model used—we asked the teams to report the direction of the values in their maps for the relevant hypotheses (5, 6 and 9). Unthresholded maps were corrected to address sign flips for reversed contrasts as reported by the analysis teams. In addition, *t* values were converted to *z* values with Huggett's transform<sup>29</sup>. All subsequent analyses of the unthresholded maps were performed only on voxels that contained non-zero data for all teams (range across hypotheses: 110,062–145,521 voxels).

We assessed the agreement between thresholded statistical maps using per cent agreement, that is, the per cent of voxels that have the same binary value. Because the thresholded maps are very sparse, these values are necessarily high when computed across all voxels. Therefore, we also computed the agreement between pairs of statistical maps only for voxels that were non-zero for at least one member of each pair. To further test the agreement across teams, we performed a coordinate-based meta-analysis with activation likelihood estimation<sup>30,31</sup> (see Supplementary Information).

We further computed the correlation between the unthresholded images of the 64 teams. The correlation matrices were clustered using Ward clustering; the number of clusters was set to three for all hypotheses on the basis of visual examination of the dendograms. A separate mean statistical map was then created for the teams in each cluster (see Fig. 2, Extended Data Fig. 2). Drivers of map similarity were further assessed by modelling the median correlation distance of each team from the average pattern as a function of several analysis decisions (for example, smoothing, whether or not the data preprocessed with fMRIPrep were used, and so on).

To assess the effect of variability in thresholding methods and anatomical definitions across teams, unthresholded *z* maps for each team

were thresholded using a common approach. The *z* maps for each team were translated to *P* values, which were then thresholded using two approaches: a heuristic correction (known to be liberal<sup>32</sup>), and a voxelwise FDR correction. Note that it was not possible to compute the commonly used familywise error correction using Gaussian random field theory because residual smoothness was not available for each team. We then identified whether there were any suprathreshold voxels within the appropriate anatomical ROI for each hypothesis. The ROIs for the ventral striatum and amygdala were defined anatomically on the basis of the Harvard-Oxford anatomical atlas. As there is no anatomical definition for the ventromedial prefrontal cortex, we defined the region using a conjunction of anatomical regions (including all anatomical regions in the Harvard-Oxford atlas that overlap with the ventromedial portion of the prefrontal cortex) and a meta-analytical map obtained from <https://neurosynth.org/> (ref. <sup>33</sup>) for the search term “ventromedial prefrontal”.

An image-based meta-analysis was used to quantify the evidence for each hypothesis across analysis teams (Extended Data Fig. 3b), accounting for the lack of independence due to the use of a common dataset across teams. See Supplementary Information for a description of the image-based meta-analysis method.

### Prediction markets

The second main goal of NARPS was to test the degree to which researchers in the field can predict results, using prediction markets<sup>2–5,34</sup>. We invited team members (researchers that were members of one of the analysis teams) and non-team members (researchers that were neither members of any of the analysis teams nor members of the NARPS research group) to participate in a prediction market<sup>2,35</sup> to measure peer beliefs about the fraction of teams reporting significant whole-brain-corrected results for each of the nine hypotheses. The prediction markets were conducted 1.5 months after all teams had submitted their analysis of the fMRI dataset. Thus, team members had information about the results of their specific team, but not about the results of any other team.

Similar to previous studies<sup>2–5</sup>, participants in the prediction markets were provided with monetary endowments (100 tokens, worth US\$50) and traded on the outcome of the hypotheses through a dedicated online market platform. Each hypothesis constitutes one asset in the market, with asset prices predicting the fraction of teams reporting significant whole-brain-corrected results for the corresponding ex-ante hypothesis examined by the analysis teams using the same dataset. Trading on the prediction markets was incentivized, that is, traders were paid on the basis of their performance in the markets.

**Recruitment.** For the non-team members prediction market, we invited participants via social media (mainly Facebook and Twitter) and emails. The invitation contained a link to an online form on the NARPS website (<https://www.narps.info/>) where participants could sign up using their email address.

Participants for the team members prediction market were invited, after all teams submitted their results, by an email that directed them to an independent registration form (with identical form fields), to separate participants for the two prediction markets already at the time of registration. Note that team members were not aware to start with that they would be invited to participate in a separate prediction market after they had analysed the data. The decision to implement a second market, consisting of traders with partial information about the fundamental values (that is, the team members) was made after the teams obtained access to the fMRI dataset. Thus, team members were only invited to participate in the market after all teams had submitted their analysis results. Once the registration for participating in the prediction markets had been closed, we reconciled the sign-ups with the list of team members to ensure that team members did not mistakenly end up in the non-team members prediction market and vice versa.

# Article

In addition to their email addresses, which were used as the only key to match registrations, accounts in the market platform and the teams' analysis results, registrants were required to provide the following information during sign-up: (i) name, (ii) affiliation, (iii) position (PhD candidate, post-doctoral researcher, assistant professor, senior lecturer, associate professor, full professor, other), (iv) years since PhD, (v) gender, (vi) age, (vii) country of residence, (viii) self-assessed expertise in neuroimaging (Likert scale ranging from 1 to 10), (ix) self-assessed expertise in decision sciences (Likert scale ranging from 1 to 10), (x) preferred mode of payment (Amazon.de voucher, Amazon.com voucher, PayPal payment), and (xi) whether they are a team member of any analysis team (yes or no). The invitations to participate in the prediction markets were first distributed on 9 April 2019; the registration closed on 29 April 2019 at 16:00 UTC. Once registration closed, all participants received a personalized email containing a link to the web-based market software and their login credentials. The prediction markets opened on 2 May 2019 at 16:00 UTC and closed on 12 May 2019 at 16:00 UTC.

**Information available to participants.** All participants had access to detailed information about the data collection, the experimental protocol, the ex-ante hypotheses, the instructions given to the analysis teams, references to related papers and detailed instructions about the prediction markets via the NARPS website (<https://www.narps.info/>).

**Implementation of prediction markets.** To implement the prediction markets, we used a newly developed web-based framework dedicated for conducting continuous-time online market experiments, inspired by the trading platform in the Experimental Economics Replication Project (EERP)<sup>3</sup> and the Social Sciences Replication Project (SSRP)<sup>4</sup>. Similar to these previous implementations, there were two main views on the platform: (i) the market overview and (ii) the trading interface. The market overview showed the nine assets (that is, one corresponding to each hypothesis) in tabular format, including information on the (approximate) current price for buying a share and the number of shares held (separated for long and short positions) for each of the nine hypotheses. Via the trading interface, which was shown after clicking on any of the hypotheses, the participant could make investment decisions and view price developments for the particular asset.

Note that initially, there was an error in the labelling of two assets (that is, hypotheses) in the trading interface and the overview table of the web-based trading platform (the more detailed hypothesis description available via the info symbol on the right-hand side of the overview table contained the correct information): hypotheses 7 and 8 mistakenly referred to negative rather than positive effects of losses in the amygdala. One of the participants informed us about the inconsistency between the information on the trading interface and the information provided on the website on 6 May 2019. The error was corrected immediately on the same day and all participants were informed about the mistake on our part through a personal email notification (on 6 May 2019, 15:28 UTC), pointing out explicitly which information was affected and asking them to double-check their holdings in the two assets to make sure that they were invested in the intended direction.

**Trading and market pricing.** In both prediction markets, traders were endowed with 100 tokens (the experimental currency unit). Once the markets opened, these tokens could be used to trade shares in the assets (that is, hypotheses). Unlike prediction markets on binary outcomes (for example, the outcomes of replications as in previous studies<sup>3,4</sup>), for which market prices were typically interpreted as the predicted probability of the outcome to occur<sup>36</sup> (although see two previous studies for caveats<sup>37,38</sup>), the prediction markets accompanying the team analyses in the current study were implemented in terms of vote-share-markets. Hence, the prediction market prices serve as measures of the aggregate beliefs of traders for the fraction of teams reporting that the hypotheses were supported and can fluctuate

between 0 (no team reported a significant result) and 1 (all teams reported a significant result).

Prices were determined by an automated market maker implementing a logarithmic market scoring rule<sup>39</sup>. At the beginning of the markets, all assets were valued at a price of 0.50 tokens per share. The market maker calculated the price of a share for each infinitesimal transaction and updated the price on the basis of the scoring rule. This ensured both that trades were always possible even when there was no other participant with whom to trade and that participants had incentives to invest according to their beliefs<sup>40</sup>. The logarithmic scoring rule uses the net sales (shares held – shares borrowed) that the market maker has done so far in a market to determine the price for an infinitesimal trade as  $p = e^{s/b}/(e^{s/b} + 1)$ . The parameter  $b$  determines the liquidity provided by the market maker and controls how strongly the market price is affected by a trade. We set the liquidity parameter to  $b = 100$ , implying that by investing 10 tokens, traders could move the price of a single asset from 0.50 to about 0.55.

Investment decisions for a particular hypothesis were made from the market's trading interface. In the trading overview, participants could see the (approximate) price of a new share, the number of shares they currently held (separated for long and short positions) and the number of tokens their current position was worth if they liquidated their shares. The trading page also contained a graph depicting previous price developments. To make an adjustment to their current position, participants could choose either to increase or decrease their position by a number of tokens of their choice. The trading procedures and market pricing are described in more detail in a previous study<sup>3</sup>.

**Incentivization.** Once the markets had been closed, the true 'fundamental value' for each asset (that is, the fraction of teams that reported a significant result for the particular hypothesis) was determined and gains and losses were calculated as follows: if holdings in a particular asset were positive (that is, the trader acted as a net buyer), the payout was calculated as the fraction of analysis teams reporting a significant result for the associated hypothesis multiplied by the number of shares held in the particular asset; if a trader's holdings were negative (that is, the trader acted as a net seller), the (absolute) amount of shares held was valued at the price differential between 1 and the fraction of teams reporting a significant result for the associated hypothesis.

Any tokens that had not been invested into shares when the market closed were voided. Any tokens awarded as a result of holding shares were converted to US dollars at a rate of 1 token = US\$0.5. The final payments were transferred to participants during the months May to September 2019 in form of Amazon.com gift cards, Amazon.de gift cards or PayPal payments, depending on the preferred mode of payment indicated by the participants after registration for the prediction markets.

**Participants.** In total, 96 team members and 91 non-team members signed up to participate in the prediction markets.  $n = 83$  team members and  $n = 65$  non-team members actively participated in the markets. The number of traders active in each of the assets (that is, hypotheses) ranged from 46 to 76 ( $m = 56.4$ , s.d. = 8.9) in the team members set of markets and from 35 to 58 ( $m = 47.1$ , s.d. = 7.9) in the non-team members set of markets. See Extended Data Table 5c for data about trading volume on the prediction markets.

Of the participants, 10.2% did not work in academia (but hold a PhD), 34.2% were PhD students, 43.3% were post-doctoral researchers or assistant professors, 7.5% were lecturers or associate professors and 4.8% were full professors. 27.8% of the participants were female. The average time spent in academia after obtaining the PhD was 4.1 years. Most of the participants lived in Europe (46.3%) and North America (46.3%).

**Preregistration.** All analyses of the prediction markets data reported were preregistered at <https://osf.io/pqeb6/>. The preregistration was

completed after the markets opened, but before the markets closed. Only one member of the NARPS research group, F. Holzmeister, had any information about the prediction market prices before the markets closed (as he monitored the prediction markets). He was not involved in writing the preregistration. Only two members of the NARPS research group, R.B.-N. and T. Schonberg, had any information about the results reported by the 70 analysis teams before the prediction markets closed. Neither of them were involved in writing the preregistration. For additional details on the prediction markets, see Supplementary Information.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

The full fMRI dataset is publicly available on OpenNeuro (<https://doi.org/10.18112/openneuro.ds001734.v1.0.4>) and is described in detail in a Data Descriptor<sup>1</sup>. The results reported by all teams are presented in Extended Data Table 2. A table describing the methods used by the analysis teams is available with the analysis code. NeuroVault collections containing the submitted statistical maps are available via the links provided in Extended Data Table 3a. Source data for Figs. 1, 2 are provided with the paper. Readers may obtain access to the data and run the full analysis stream on the team submissions by following the directions at <https://github.com/poldrack/narps/tree/master/ImageAnalyses>. Access to the raw data requires specifying a URL for the dataset, which is: [https://zenodo.org/record/3528329/files/narps\\_origdata\\_1.0.tgz](https://zenodo.org/record/3528329/files/narps_origdata_1.0.tgz). Results (automatically generated figures, results and output logs) for image analyses are available for anonymous download at <https://doi.org/10.5281/zenodo.3709275>.

## Code availability

Code for all analyses of the reports and statistical maps submitted by the analysis teams is openly shared in GitHub (<https://github.com/poldrack/narps>). Image-analysis code was implemented within a Docker container, with software versions pinned for reproducible execution (<https://hub.docker.com/r/poldrack/narps-analysis/tags>). Python code was automatically tested for quality using the flake8 static analysis tool and the codacy.com code quality assessment tool, and the results of the image-analysis workflow were validated using simulated data. The image-analysis code was independently reviewed by an expert who was not involved in writing the original code. Prediction market analyses were performed using R v.3.6.1; packages were installed using the checkpoint package, which reproducibly installs all package versions as of a specified date (13 August 2019). Analyses reported in this manuscript were performed using code release v.2.0.3 (<https://doi.org/10.5281/zenodo.3709273>). Although not required to, several analysis teams publicly shared their analysis code. Extended Data Table 3d includes these teams along with the link to their code.

23. Gorgolewski, K. J. et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci. Data* **3**, 160044 (2016).
24. Tversky, A. & Kahneman, D. Advances in prospect theory: cumulative representation of uncertainty. *J. Risk Uncertain.* **5**, 297–323 (1992).
25. Nichols, T. E. et al. Best practices in data analysis and sharing in neuroimaging using MRI. *Nat. Neurosci.* **20**, 299–303 (2017).
26. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
27. Lubke, G. H. et al. Assessing model selection uncertainty using a bootstrap approach: an update. *Struct. Equ. Modeling* **24**, 230–245 (2017).
28. Abraham, A. et al. Machine learning for neuroimaging with scikit-learn. *Front. Neuroinform.* **8**, 14 (2014).
29. Hughett, P. Accurate computation of the *F*-to-*z* and *t*-to-*z* transforms for large arguments. *J. Stat. Softw.* **23**, 1–5 (2007).
30. Turkeltaub, P. E., Eden, G. F., Jones, K. M. & Zeffiro, T. A. Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *Neuroimage* **16**, 765–780 (2002).
31. Eickhoff, S. B. et al. Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. *Neuroimage* **137**, 70–85 (2016).
32. Eklund, A., Nichols, T. E. & Knutsson, H. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl. Acad. Sci. USA* **113**, 7900–7905 (2016).
33. Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670 (2011).
34. Open Science Collaboration. Estimating the reproducibility of psychological science. *Science* **349**, aac4716 (2015).
35. Arrow, K. J. et al. Economics. The promise of prediction markets. *Science* **320**, 877–878 (2008).
36. Wolfers, J. & Zitzewitz, E. Interpreting prediction market prices as probabilities. <https://doi.org/10.3386/w12200> (NBER, 2006).
37. Manski, C. F. Interpreting the predictions of prediction markets. *Econ. Lett.* **91**, 425–429 (2006).
38. Fountain, J. & Harrison, G. W. What do prediction markets predict? *Appl. Econ. Lett.* **18**, 267–272 (2011).
39. Hanson, R. Logarithmic market scoring rules for modular combinatorial information aggregation. *J. Prediction Markets* **1**, 3–15 (2007).
40. Chen, Y. *Markets as an Information Aggregation Mechanism for Decision Support*. PhD thesis, Penn State Univ. (2005).

**Acknowledgements** Neuroimaging data collection, performed at Tel Aviv University, was supported by the Austrian Science Fund (P29362-G27), the Israel Science Foundation (ISF 2004/15 to T. Schonberg) and the Swedish Foundation for Humanities and Social Sciences (NHS14-1719:1). Hosting of the data on OpenNeuro was supported by a National Institutes of Health (NIH) grant (R24MH117179). We thank M. C. Frank, Y. Assaf and N. Daw for comments on an earlier draft; the Texas Advanced Computing Center for providing computing resources for preprocessing of the data; the Stanford Research Computing Facility for hosting the data; and D. Roll for assisting with data processing. T. Schonberg thanks The Alfredo Federico Strauss Center for Computational Neuroimaging at Tel Aviv University; A.D. thanks the Knut and Alice Wallenberg Foundation and the Marianne and Marcus Wallenberg Foundation (A.D. is a Wallenberg Scholar), the Austrian Science Fund (FWF, SFB F63) and the Jan Wallander and Tom Hedelius Foundation (Svenska Handelsbankens Forskningsstiftelsen); F. Holzmeister, J. Huber and M. Kirchner thank the Austrian Science Fund (FWF, SFB F63); D.W. was supported by the Research Foundation Flanders (FWO) and the European Union's Horizon 2020 research and innovation programme (<https://ec.europa.eu/programmes/horizon2020/en>) under the Marie Skłodowska-Curie grant agreement no. 665501; L. Tisdall was supported by the University of Basel Research Fund for Junior Researchers; C.B.C. was supported by grant 1207719N from the Research Foundation Flanders; E.L. was supported by grant 12T2517N from the Research Foundation Flanders and Marie Skłodowska-Curie Actions under COFUND grant agreement 665501; A. Eed was supported by a predoctoral fellowship La Caixa-Severo Ochoa from Obra Social La Caixa and also acknowledges Comunidad de Cálculo Científico del CSIC for the high-performance computing (HPC) use; C.L. was supported by the Vienna Science and Technology Fund (WWTF VRG13-007) and Austrian Science Fund (FWF P 32686); A.B.L.V. was supported by the Vienna Science and Technology Fund (WWTF VRG13-007); L.Z. was supported by the Vienna Science and Technology Fund (WWTF VRG13-007), the National Natural Science Foundation of China (no. 7180110), MOE (Ministry of Education in China) Project of Humanities and Social Sciences (no. 18YJC630268) and China Postdoctoral Science Foundation (no. 2018M633270); D.P. is currently supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy 'Science of Intelligence' (EXC 2002/1, project number 390523135); P.H. was supported in part by funding provided by Brain Canada, in partnership with Health Canada, for the Canadian Open Neuroscience Platform initiative; J.-B.P. was partially funded by the NIH (NIH-NIBIB P41 EB019936 (ReproNim), NIH-NIMH R01 MH083320 (CANDIShare) and NIH RF1 MH120021 (NIDM)) and the National Institute Of Mental Health of the NIH under award number R01MH096906 (Neurosynth), as well as the Canada First Research Excellence Fund, awarded to McGill University for the Healthy Brains for Healthy Lives initiative and the Brain Canada Foundation with support from Health Canada; S.B.E. was supported by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 785907 (HBP SGA2); G.M. was supported by the Max Planck Society; S. Heunis has received funding from the Dutch foundation LSH-TKI (grant LSHM16053-SGF); J.F.G.M. was supported by a Graduate Research Fellowship from the NSF and T32 Predoctoral Fellowship from the NIH; B.M. was supported by the Deutsche Forschungsgemeinschaft (grant CRC1193, subproject B01); A.R.L. was supported by NSF 1631325 and NIH R01 DAO41353; M.E.H., T.J. and D.J.W. were supported by the Australian National Imaging Facility, a National Collaborative Research Infrastructure Strategy (NCRIS) capability; P.M.I. was supported by VIDI grant 452.17-013 from the Netherlands Organisation for Scientific Research; B.M.B. was supported by the Max Planck Society; J.P.H. was supported by a grant from the Swedish Research Council; R.W.C. and R.C.R. were supported by NIH IRP project number Z1CMH002888; D.M.N., R.W.C., and R.C.R. used the computational resources of the National Institutes of Health High Performance Computing Biowulf cluster (<http://hpc.nih.gov>); D.M.N. was supported by NIH IRP project number Z1CMH002960; C.F.C. was supported by the Tianqiao and Chrissy Center for Social and Decision Neuroscience Center Leadership Chair; R.G.B. was supported by the Max Planck Society; R.M.W.J.B. was supported by the Max Planck Society; M.B., O.C. and R.G. were supported by the Belgian Excellence of Science program (EOS project 30991544) from the FNRS-Belgium; O.C. is a research associate at the FRS-FNRS of Belgium; A.D.L. was supported by grant R4195 "Repimpact" of EraNET Neuron; Q.S. was funded by grant no. 71971199, 71602175 and 71942004 from the National Natural Science Foundation of China and no. 16YJC630103 of the Ministry of Education of Humanities and Social Science; and T.E.N. was supported by the Wellcome Trust award 100309/Z/12/Z.

**Author contributions** NARPS management team: R.B.-N., F. Holzmeister, C.F.C., A.D., J. Huber, M.J., M. Kirchner, R.A.P. and T. Schonberg. fMRI dataset (experiment design): R.I., J. Durne, R.A.P. and T. Schonberg. fMRI dataset (data collection): R.I. and T. Schonberg. fMRI dataset (preprocessing, quality assurance and data sharing): R.B.-N., K.J.G., R.A.P. and T. Schonberg.

# Article

Analysis teams (recruitment, point of contact and management): R.B.-N., R.A.P. and T. Schonberg. Analysis teams (analysis of the submitted results and statistical maps): R.A.P., T.E.N., J.A.M., J.-B.P., A.P., R.B.-N. and T. Schonberg. Code review: T.G. and K.D. Prediction markets (design and management): F. Holzmeister, C.F.C., A.D., J. Huber, M.J. and M. Kirchler. Prediction markets (analysis): F. Holzmeister, R.B.-N., C.F.C., A.D., J. Huber, M.J., M. Kirchler, S.K., R.A.P. and T. Schonberg. Writing the manuscript: R.B.-N., F. Holzmeister, A.D., J. Huber, M.J., M. Kirchler, T.E.N., R.A.P. and T. Schonberg. Participated as members of analysis teams and reviewed and edited the manuscript: R.A.A., P.A., B.M.B., A. Bajracharya, L.B., S. Ball, M.B., N.B., D.B., J.B., R.G.B., R.M.W.J.B., J.P.B., B.B.B., S.B.-S., T.B., K.L.B., A. Bowring, S. Braem, H.R.B., E.G.B., C.B.C., J.A.C., J.J.C., L.C., E.C.C., Z.J.C., O.C., R.W.C., W.A.C., S.C., K.D., C.P.D., A.D.L., M.R.D., L.D., J.B.D., X.D., E.W.D., E.D., C.L.D., J. Dukart, N.W.D., A. Eed, S.B.E., A. Erhart, L.F., G.M.F., S.F., A.G., R.G., S.G., E.G., J.J.G., S.A.E.G., C.G.-G., K.J.G., C.L.G., M.A.G., J.F.G.M., O.G., S. Hakimi, J.P.H., R.H., G. Handjara, B.B.H., C.H., P.H., G. Herman, S. Heunis, F. Hoffstaedter, J. Hogeveen, S. Holmes, C.-P.H., S.A.H., M.E.H., V.I., A.D.I., P.M.I., A.I.I., A.J., M.R.J., T.J., M.J.E.J., A.C.J., J.W.K., M. Kassinopoulos, C.K., X.-Z.K., T.R.K., N.E.K., B.A.K., A.R.L., C.L., R.L., N.L., H.L., S.L., A. Leemans, A. Leo, E.L., F.L., M.Y.C.L., P.C.L., E.N.L., S.W.L., A.B.L.V., B.C.L., M.L.M., N.M.,

T.M., C.M., K.M., J.T.M., H.M., A.S.M.L., B.M., K.N.M., G.M., G.D.M., J.M., T.E.N., D.M.N., G.N., M.P.N., E.O., A.I.O., P.P., K.R.P., J.E.P., D.P., Y.P., S.R., P.A.R.-L., R.C.R., E.R., J.R.R., A.M.R.-T., A.R., T. Salo, G.R.S.-L., E.S.-M., M.L.S., D.H.S., Q.S., M.A.S., J.A.S., K.S., A.S., D.V.S., P.S.-H., S.R.S., S.M.T., B.T., J.N.T., G.T., L. Tisdall, S.H.T., C.T.-S., J.J.T.T., L. Tozzi, V.T., L. Turella, A.E.v.V., T.V., J.M.V., S.V., K.V., M.B.W., W.D.W., S.W., D.J.W., D.W., A.X.-P., E.A.Y., S.Y., R.Y., K.S.L.Y., L.Z., X.Z. and J.E.Z.

**Competing interests** The authors declare no competing interests.

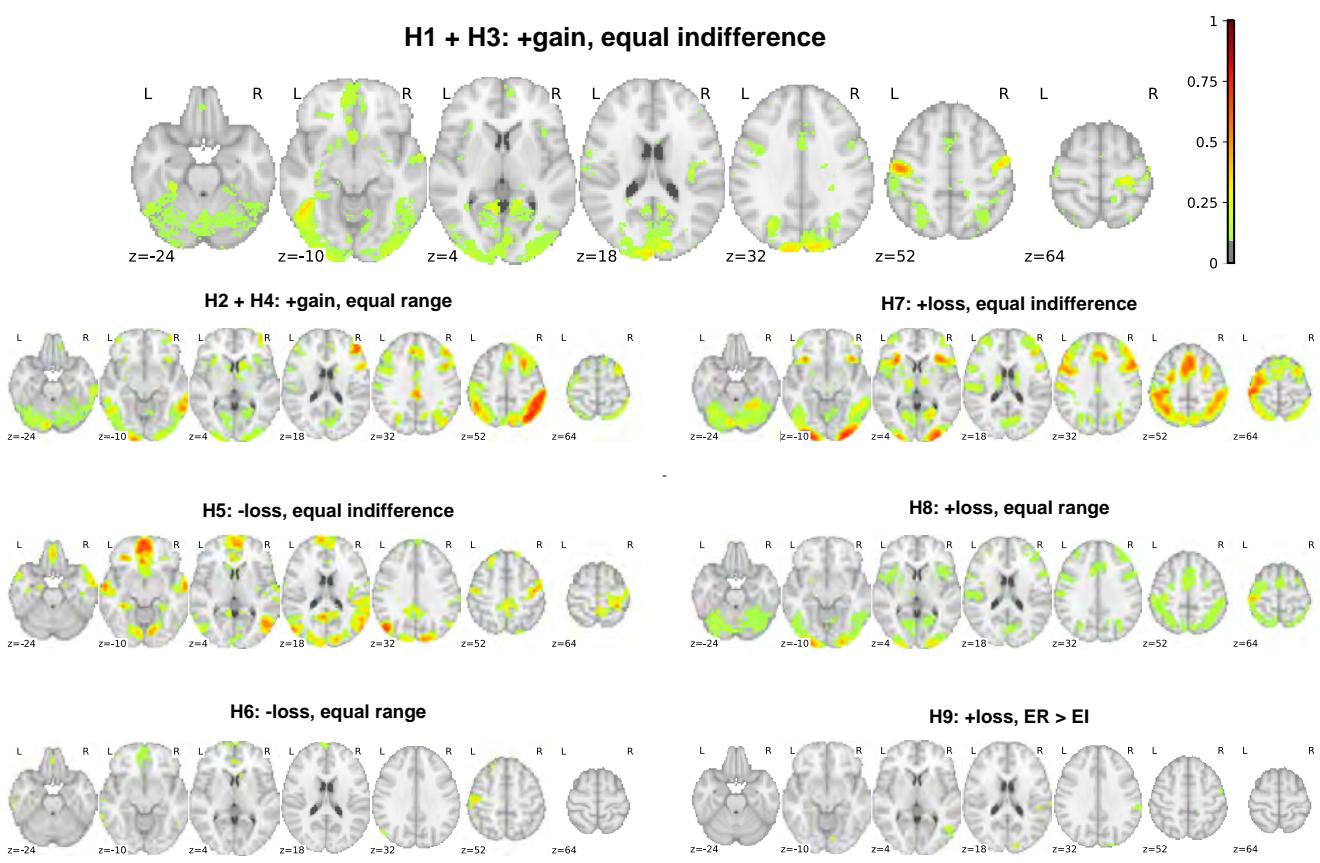
**Additional information**

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2314-9>.

**Correspondence and requests for materials** should be addressed to T.E.N., R.A.P. or T.S.

**Peer review information** *Nature* thanks Martin Lindquist, Marcus Munafò and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

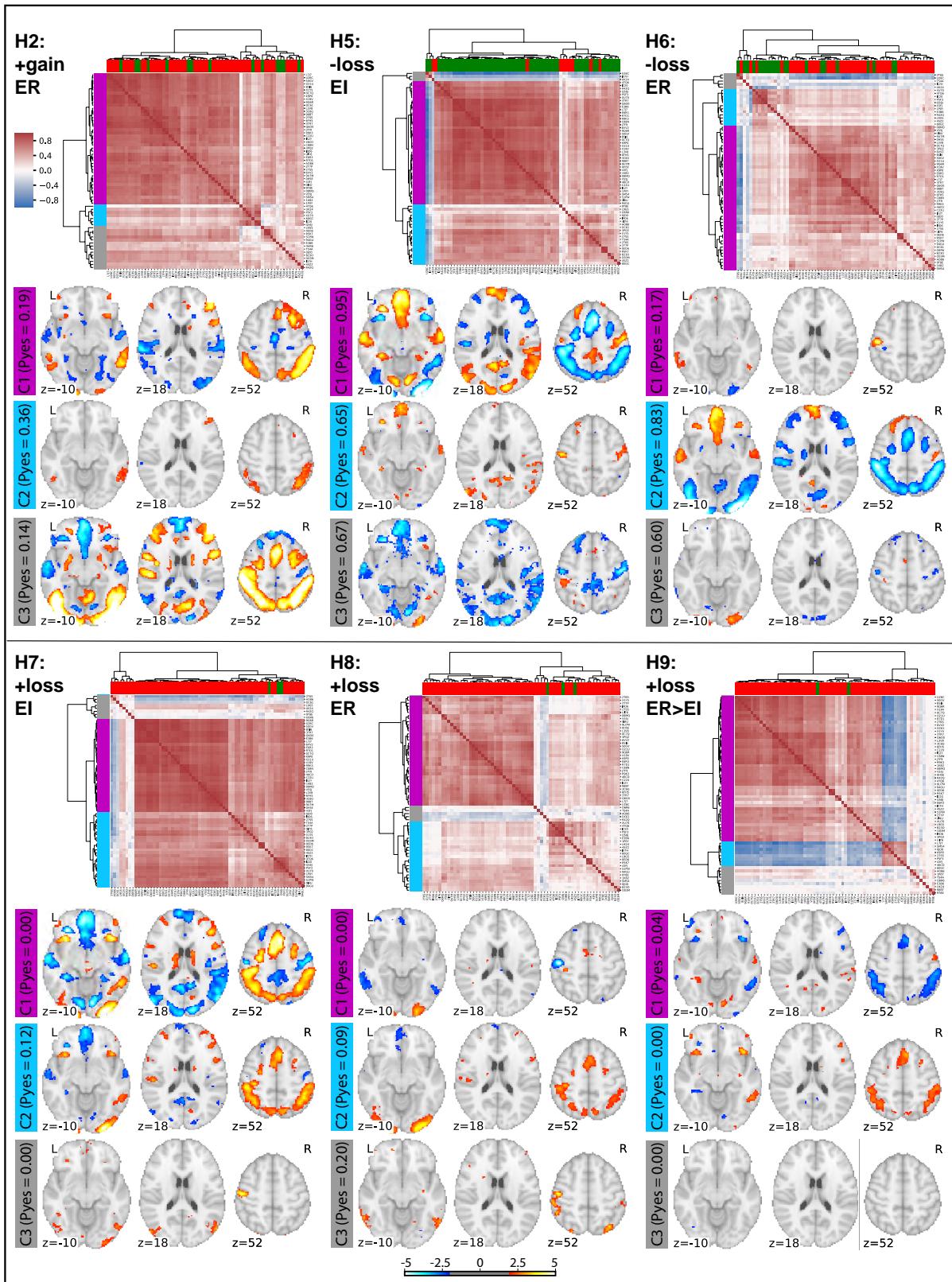
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**Extended Data Fig. 1 | Voxels overlap.** Maps showing at each voxel the proportion of teams (out of  $n=65$  teams) that reported significant activations in their thresholded statistical map, for each hypothesis (labelled H1–H9), thresholded at 10% (that is, voxels with no colour were significant in fewer than 10% of teams). + or – refers to the direction of effect; gain or loss refers to the

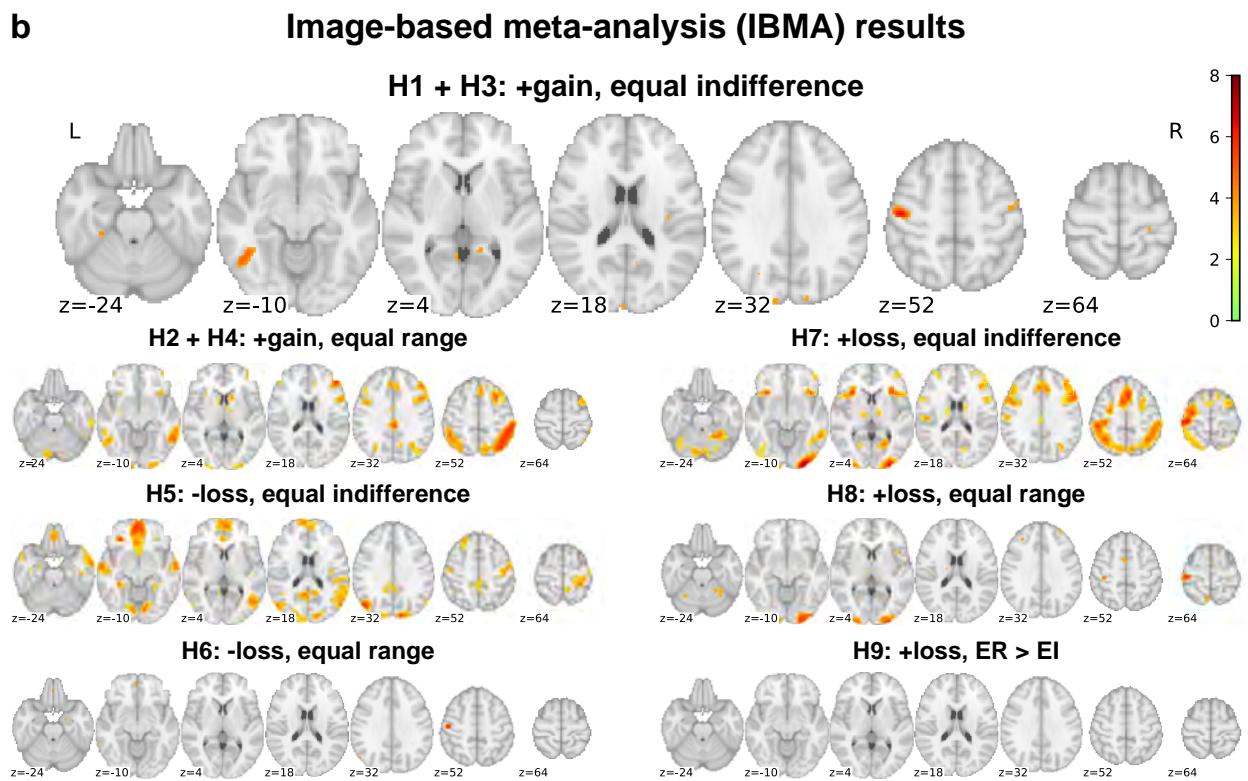
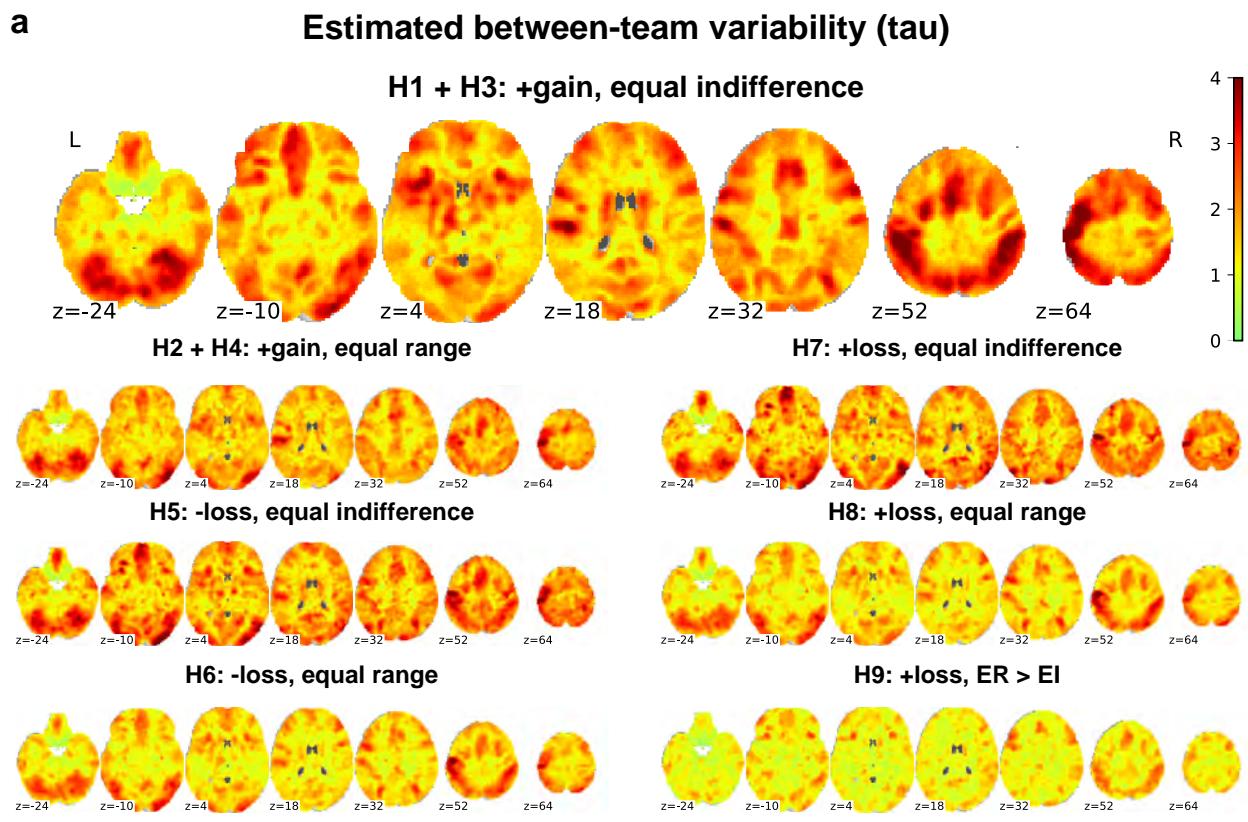
effect being tested; and equal indifference (EI) or equal range (ER) refers to the group being examined or compared. Hypotheses 1 and 3, as well as hypotheses 2 and 4, share the same statistical maps as they relate to the same contrast and experimental group but different regions (see Extended Data Table 1). Images can be viewed at <https://identifiers.org/neurovault.collection:6047>.

# Article



**Extended Data Fig. 2 | Variability of whole-brain unthresholded maps for hypotheses 2 and 4–9.** For each hypothesis, we present a heatmap based on Spearman correlations between unthresholded statistical maps ( $n = 64$ ), clustered according to their similarity, and the average of unthresholded images for each cluster (cluster colours in titles refer to colours in left margin of heatmap). Column colours represent hypothesis decisions (green, yes; red, no)

reported by the analysis teams; row colours denote cluster membership. Maps are thresholded at an uncorrected value of  $z > 2$  for visualization. Unthresholded maps for hypotheses 2 and 4 are identical (as they both relate to the same contrast and group but different regions), and the colours represent reported results for hypothesis 2. For hypotheses 1 and 3, see Fig. 2.

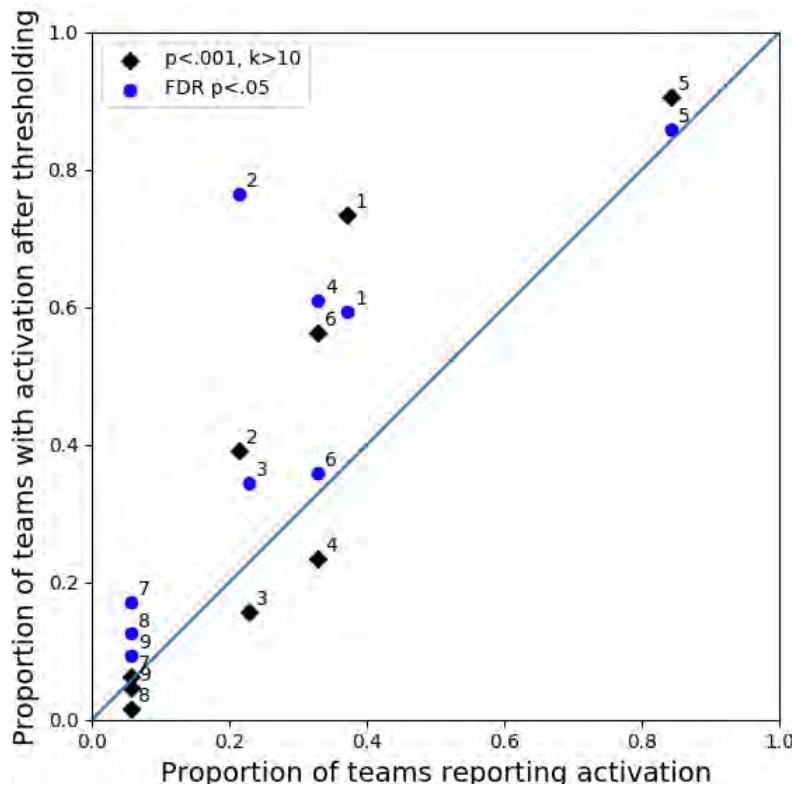


**Extended Data Fig. 3 | Variability and consensus of unthresholded statistical maps.**  $n = 64$ . **a**, Maps of estimated between-team variability ( $\tau_{\alpha}$ ) at each voxel for each hypothesis. **b**, Results of the image-based meta-analysis. A consensus analysis was performed on the unthresholded statistical maps to obtain a group statistical map for each hypothesis, accounting for the correlation between teams owing to the same underlying data (see Methods). Maps are presented for each hypothesis, showing voxels (in colour) in which

the group statistic was significantly greater than zero after voxelwise correction for FDR ( $P < 0.05$ ). Colour bar reflects statistical value ( $z$ ) for the meta-analysis. Hypotheses 1 and 3, as well as hypotheses 2 and 4, share the same unthresholded maps, as they relate to the same contrast and group but different regions (see Extended Data Table 1). Images can be viewed at <https://identifiers.org/neurovault.collection:6051>.

# Article

a



b

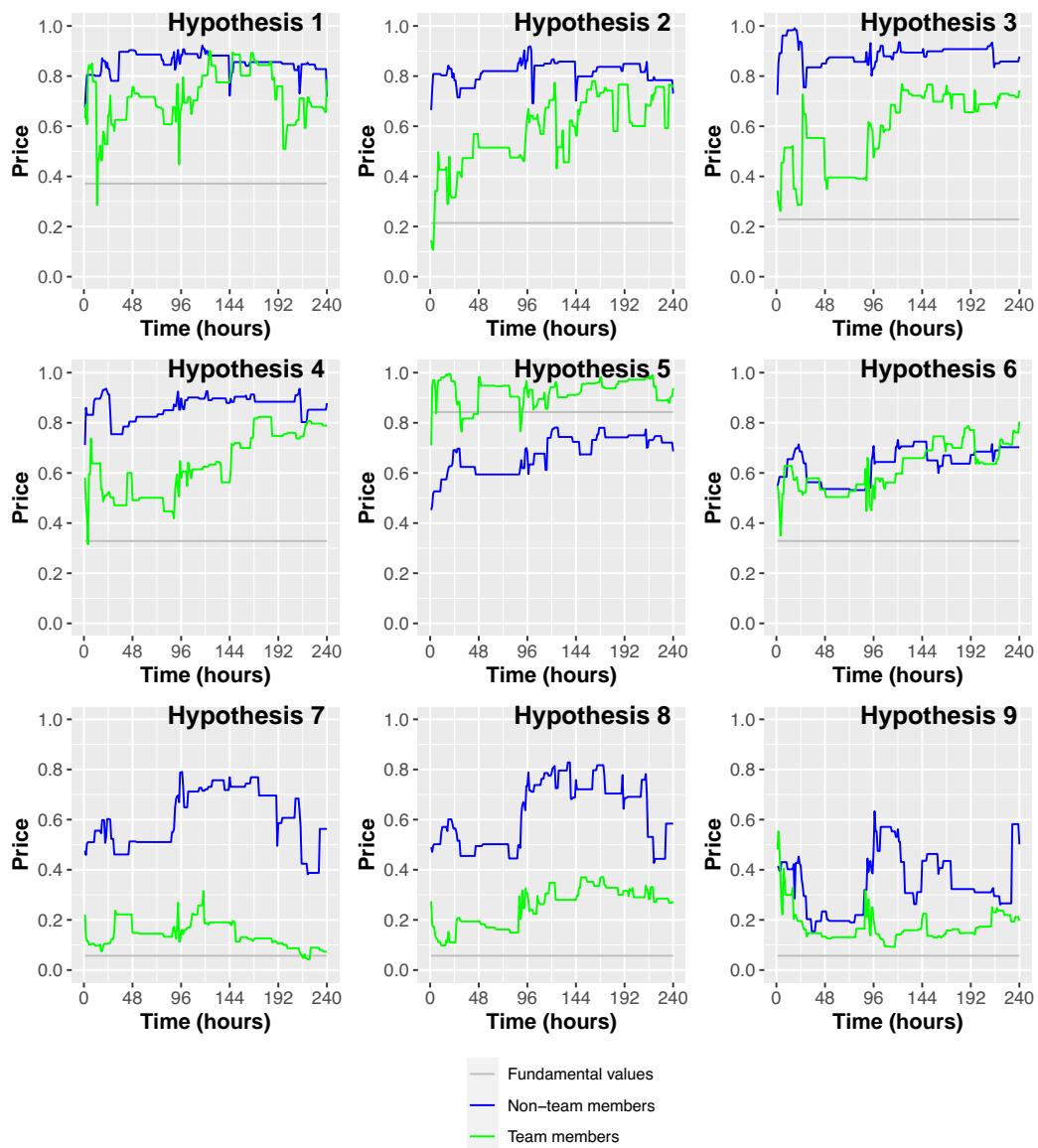
Hypothesis	N voxels in ROI	Proportion of teams reporting activation	Proportion of teams with activation ( $p < 0.001, k > 10$ )	Proportion of teams with activation (FDR)	IBMA (n voxels in ROI)
1	3402	0.371	0.734	0.594	0
2	3402	0.214	0.391	0.766	7
3	173	0.229	0.156	0.344	0
4	173	0.329	0.234	0.609	7
5	3402	0.843	0.906	0.859	2101
6	3402	0.329	0.562	0.359	39
7	672	0.057	0.062	0.172	0
8	672	0.057	0.016	0.125	0
9	672	0.057	0.047	0.094	0

**Extended Data Fig. 4 | Results of the consistent thresholding and ROI selection analysis.**  $n=64$ . a, Activation for each hypothesis as determined using consistent thresholding (black,  $P < 0.001$  and cluster size ( $k$ )  $> 10$  voxels; blue, FDR correction with  $P < 0.05$ ) and ROI selection across teams (y axis), versus the actual proportion of teams reporting activation (x axis). Numbers next to each symbol represent the hypothesis number for each point. b, Results

from re-thresholding of unthresholded maps, using either uncorrected values with the threshold ( $P < 0.001, k > 10$ ) or FDR correction ( $P_{\text{FDR}} < 5\%$ ) and common anatomical ROIs for each hypothesis. A team is recorded as having an activation if one or more significant voxels are found in the ROI. Results for image-based meta-analysis (IBMA) for each hypothesis are presented, also thresholded at  $P_{\text{FDR}} < 5\%$ .

**a**

Effect	Beta (full model)	t (full model)	p (full model)	Beta (no interaction)	t (no interaction)	p (no interaction)
Intercept	0.44	64.12	0.00	0.41	74.61	0.00
Time	0.00	3.38	0.00	0.00	12.48	0.00
Teams	-0.29	-29.50	0.00	-0.22	-45.35	0.00
Time X Teams	0.00	7.78	0.00			
-----						
Adjusted R-squared			0.35			0.34

**b**

**Extended Data Fig. 5 | Prediction markets over time.**  $n = 240$  observations (10 days  $\times$  24 h). **a**, Panel regressions. The table summarizes the results of preregistered fixed-effects panel regressions of the absolute errors of the predictions (that is, the absolute deviation of the market price from the fundamental value) on an hourly basis (average price of all transactions within an hour) on time and prediction market indicators. Standard errors were computed using a robust estimator. **b**, Market prices for each of the nine

hypotheses separated for the team members (green) and non-team members (blue) prediction markets. The figure shows the average prices of the prediction market per hour, separated for the two prediction markets, for the time the markets were open (10 days, that is, 240 h). The grey line indicates the actual share of the analysis teams that reported a significant result for the hypothesis (that is, the fundamental value).

# Article

**Extended Data Table 1 | Hypotheses and results**

Hypothesis description	Fraction of teams reporting a significant result	Median confidence level	Median similarity estimation
#1 Positive parametric effect of gains in the vmPFC (equal indifference group)	0.371	7 (2)	7 (1.5)
#2 Positive parametric effect of gains in the vmPFC (equal range group)	0.214	7 (1.5)	7 (1)
#3 Positive parametric effect of gains in the ventral striatum (equal indifference group)	0.229	6 (1)	7 (1)
#4 Positive parametric effect of gains in the ventral striatum (equal range group)	0.329	6 (1)	7 (1)
#5 Negative parametric effect of losses in the vmPFC (equal indifference group)	0.843	8 (1)	8 (1)
#6 Negative parametric effect of losses in the vmPFC (equal range group)	0.329	7 (1)	7 (1)
#7 Positive parametric effect of losses in the amygdala (equal indifference group)	0.057	7 (1)	8 (1)
#8 Positive parametric effect of losses in the amygdala (equal range group)	0.057	7 (1)	8 (1)
#9 Greater positive response to losses in amygdala (equal range group vs. equal indifference group)	0.057	6 (1)	7 (1)

Each hypothesis is described along with the fraction of teams that reported a whole-brain-corrected significant result (out of  $n = 70$  teams) and two measures reported by the analysis teams for the specific hypothesis: (1) How confident are you about this result? (2) How similar do you think your result is to the other analysis teams? Both of these ordinal measures are rated on a scale of 1–10, and the median values are presented together with the median absolute deviation in brackets. vmPFC, ventromedial prefrontal cortex. See Supplementary Information for analysis of the confidence level and similarity estimation.

**Extended Data Table 2 | Results submitted by analysis teams**

Team ID	H1	H2	H3	H4	H5	H6	H7	H8	H9	Est. smoothing	Package	fMRIPrep	Testing	Movement
08MQ	8	6	8	6	7	7	7	7	6	13.14	FSL	No	Non-parametric	Yes
0C7Q	7	7	8	8	8	7	10	10	9	8.68	Other	Yes	Non-parametric	Yes
0ED6	7	9	8	7	8	8	9	9	6	7.86	SPM	No	Parametric	Yes
0H5E	4	7	7	6	8	5	8	7	1	14.17	SPM	No	Parametric	No
0I4U	4	7	6	8	9	9	9	9	9	8.69	SPM	No	Parametric	Yes
0J0O	7	5	5	5	5	5	5	5	5	8.12	Other	Yes	Parametric	Yes
16IN	8	7	6	6	8	7	8	6	6		Other	Yes	Other	No
1K0E	7	9	6	6	8	7	7	6	9		Other	No	Non-parametric	Yes
1KB2	6	6	8	8	5	5	8	8	7	13.06	FSL	No	Parametric	Yes
1P0Y	8	8	1	1	8	8	5	5	5	9.13	SPM	No	Parametric	No
27SS	4	6	7	7	7	7	6	8	4	11.37	AFNI	No	Parametric	Yes
2T6S	8	9	6	6	10	9	7	8	10	14.93	SPM	Yes	Parametric	Yes
2T7P	8	8	8	8	8	8	8	8	8	7.66	Other	No	Other	Yes
3C6G	6	7	7	5	8	8	8	8	8	14.26	SPM	No	Parametric	Yes
3PQ2	9	8	7	7	7	8	8	8	7	5.79	FSL	No	Parametric	Yes
3TR7	2	2	3	4	8	5	8	6	5	17.4	SPM	Yes	Parametric	Yes
43FJ	3	3	5	5	10	10	10	10	10	10.66	FSL	No	Parametric	Yes
46CD	9	8	5	8	9	8	9	9	5	10.92	Other	No	Parametric	Yes
4SZ2	7	5	6	6	9	9	7	8	7	6.65	FSL	Yes	Parametric	No
4TQ6	7	9	10	9	7	8	10	10	9	14.88	FSL	Yes	Non-parametric	No
50GV	10	10	10	10	10	10	10	10	10	10.26	FSL	Yes	Parametric	No
51PW	8	8	8	8	8	8	6	6	7	11.15	FSL	Yes	Parametric	Yes
5G9K	7	7	7	7	7	7	7	7	7		SPM	Yes	Parametric	Yes
6FH5	9	2	8	8	10	8	8	9	9	12.22	SPM	No	Parametric	Yes
6VV2	8	8	8	6	9	7	8	7	6	7.2	AFNI	No	Parametric	Yes
80GC	9	9	8	4	3	9	6	5	4	4.02	AFNI	Yes	Parametric	Yes
94GU	8	8	8	8	8	8	8	8	8	11.19	SPM	No	Parametric	Yes
98BT	9	7	7	8	9	7	8	8	8	11.48	SPM	No	Parametric	Yes
9Q6R	10	10	10	10	10	10	8	8	8	10.28	FSL	No	Parametric	Yes
9T8E	5	5	5	5	5	5	5	5	4	9.85	SPM	Yes	Non-parametric	Yes
9U7M	7	9	9	9	9	7	9	7	7	14.78	Other	No	Parametric	Yes
AO86	7	7	7	7	7	7	7	7	7	7.49	Other	Yes	Non-parametric	Yes
B23O	6	6	7	7	8	7	6	6	8	3.32	FSL	Yes	Non-parametric	No
B516	10	10	5	5	10	6	8	7	6	9.84	FSL	Yes	Non-parametric	Yes
C22U	8	7	5	8	9	8	8	8	8	11.16	FSL	No	Parametric	No
C88N	7	8	7	4	9	7	8	8	6	11.62	SPM	Yes	Parametric	No
DC61	5	1	5	2	9	5	5	5	5	9.58	SPM	Yes	Parametric	Yes
E3B6	3	7	6	6	8	8	7	7	7	12.8	SPM	Yes	Parametric	Yes
E6R3	5	5	7	3	4	4	7	7	7	9.28	Other	Yes	Other	Yes
I07H	3	3	3	3	9	9	9	9	9	5.59	Other	Yes	Non-parametric	No
I52Y	8	8	8	8	8	8	8	8	8	11.42	FSL	No	Non-parametric	Yes
I9D6	7	7	7	7	1	7	7	6	7	6.21	AFNI	No	Parametric	Yes
IZ20	7	7	7	7	7	7	7	6	6	21.28	Other	No	Parametric	No
J7F9	9	8	9	7	9	7	9	9	9	14.88	SPM	Yes	Parametric	Yes
K9P0	10	10	10	5	10	8	9	9	10	8.05	AFNI	Yes	Parametric	Yes
L1A8	8	5	7	7	8	8	3	8	3		SPM	No	Parametric	Yes
L3V8	9	9	9	9	9	9	9	9	9	14.74	SPM	No	Parametric	No
L7J7	10	9	9	5	8	8	8	9	8	11.76	SPM	Yes	Parametric	Yes
L9G5	5	4	4	6	10	10	9	9	7	7.22	SPM	No	Parametric	No
O03M	3	8	8	2	8	7	7	7	7	3.47	AFNI	Yes	Non-parametric	Yes
O21U	8	8	8	8	8	8	8	8	8	8.26	FSL	Yes	Parametric	Yes
O6R6	8	8	8	8	8	8	8	8	8	3.06	FSL	Yes	Non-parametric	No
P5F3	3	5	7	7	4	4	6	6	7	12.94	FSL	No	Parametric	Yes
Q58J	9	9	9	9	9	9	9	9	9	16.24	FSL	No	Parametric	No
Q600	7	8	8	9	9	8	8	6	7	14.58	SPM	Yes	Parametric	Yes
R42Q	5	5	6	6	6	6	7	8	8	12.73	Other	No	Parametric	Yes
R5K7	6	8	7	9	7	8	8	8	7	12.06	SPM	No	Parametric	Yes
R7D1	4	7	5	5	9	5	8	9	8	8.93	Other	Yes	Non-parametric	Yes
R9K3	5	3	2	5	8	5	3	4	5	11.77	SPM	Yes	Parametric	Yes
SM54	5	9	5	8	8	6	8	8	8	7.05	Other	Yes	Parametric	Yes
T54A	5	9	2	6	9	9	5	5	5	12.28	FSL	Yes	Non-parametric	No
U26C	8	8	8	8	10	8	8	8	9	10.38	SPM	Yes	Parametric	Yes
UI76	10	6	10	10	6	10	10	5		6.6	AFNI	Yes	Parametric	Yes
UK24	4	4	4	4	4	4	4	4	4	10.76	SPM	No	Parametric	No
V55J	4	5	7	7	4	7	5	7	7	12.85	SPM	No	Parametric	No
VG39	6	7	8	8	10	7	9	6	5		SPM	Yes	Parametric	No
X19V	6	7	8	5	9	6	9	9	9	8.48	FSL	Yes	Parametric	Yes
X1Y5	6	6	7	7	8	6	8	8	8	8.69	Other	Yes	Non-parametric	Yes
X1Z4	8	6	4	4	9	5	4	4	4		Other	No	Non-parametric	Yes
XU70	4	5	8	9	9	6	8	8	8	7.17	FSL	No	Parametric	Yes

For each team, the left section of the table represents the reported binary decision (green, yes; red, no) and how confident they were in their result (from 1 (not at all confident) to 10 (extremely confident)) for each hypothesis (H1–H9). The right section displays the information included for each team in the statistical model for hypothesis decisions. Estimated (est.) smoothing values represent full width at half-maximum (FWHM); teams with a blank value were excluded from further analysis. Note that three teams changed their decisions after the end of the project: team L3V8 changed its decision for hypothesis 6 from yes to no; team VG39 changed its decisions for hypotheses 3, 4 and 5 from yes to no; and team U26C changed its decision for hypothesis 5 from yes to no. Results throughout the paper and in this table reflect the final results as they were reported at the end of the project (that is, before this change), as prediction markets were based on those results.

# Article

**Extended Data Table 3 | Data links and analysis-related tables**

a		b	
Team ID	Collection	Team ID	Collection
08MQ	4953	C88N	4812
0C7Q	5652	DC61	4963
0ED6	4994	E3B6	4782
0H5E	4936	E6R3	4959
0I4U	4938	I07H	5001
0JO0	4807	I52Y	4933
16IN	4927	I9D6	4978
1K0E	4974	IZ20	4979
1KB2	4945	J7F9	4949
1P0Y	5649	K9P0	4961
27SS	4975	L1A8	5680
2T6S	4881	L3V8	4888
2T7P	4917	L7J7	4866
3C6G	4772	L9G5	5173
3PQ2	4904	O03M	4972
3TR7	4966	O21U	4779
43FJ	4824	O6R6	4907
46CD	5637	P5F3	4967
4SZ2	5665	Q58J	5164
4TQ6	4869	Q6O0	4968
50GV	4735	R42Q	5619
51PW	5167	R5K7	4950
5G9K	4920	R7D1	4954
6FH5	5663	R9K3	4802
6VV2	4883	SM54	5675
80GC	4891	T54A	4876
94GU	5626	U26C	4820
98BT	4988	UI76	4821
9Q6R	4765	UK24	4908
9T8E	4870	V55J	4919
9U7M	4965	VG39	5496
AO86	4932	X19V	4947
B23O	4984	X1Y5	4898
B5I6	4941	X1Z4	4951
C22U	5653	XU70	4990

c			
Effects	Chi-squared	P value	Delta R <sup>2</sup>
Hypothesis	185.390	0.000	0.350
Estimated smoothness	13.210	0.000	0.040
Used fMRIprep data	2.270	0.132	0.010
Software package	13.450	0.004	0.040
Multiple correction method	7.500	0.024	0.020
Movement modeling	1.160	0.281	0.000

d	
Team ID	Link to shared analysis codes
16IN	<a href="https://github.com/jennyrieck/NARPS">https://github.com/jennyrieck/NARPS</a>
2T7P	<a href="https://osf.io/3b57r">https://osf.io/3b57r</a>
E3B6	<a href="doi.org/10.5281/zenodo.3518407">doi.org/10.5281/zenodo.3518407</a>
Q58J	<a href="https://github.com/amrka/NARPS_Q58J">https://github.com/amrka/NARPS_Q58J</a>

**a.** Numbers of public NeuroVault collections of all analysis teams <https://neurovault.org/collections/>. **b.** Descriptions of teams that were excluded from the analyses of statistical maps. **c.** Summary of mixed-effects logistic regression modelling of decision outcomes ( $n = 64$  per hypothesis) as a function of different factors including the hypothesis (1–9) and various aspects of statistical modelling (for modelling details see <https://github.com/poldrack/narps/blob/master/ImageAnalyses/DecisionAnalysis.Rmd>). **d.** Links to shared analysis code of some of the analysis teams.

**Extended Data Table 4 | Variability of statistical maps across teams**

**a**

Hypothesis	Minimum sig. voxels	Maximum sig. voxels	Median sig. voxels	N empty images
1	0	118181	1940	8
2	0	135583	8120	2
3	0	118181	1940	8
4	0	135583	8120	3
5	0	76569	6527	11
6	0	72732	167	25
7	0	147087	9383	8
8	0	129979	475	16
9	0	49062	266	29

**b**

Hypothesis	Correlation (mean)	Cluster1		Cluster2		Cluster3	
		Correlation	Cluster size	Correlation	Cluster size	Correlation	Cluster size
1+3	0.394	0.670	50	0.680	7	0.095	7
2+4	0.521	0.736	43	0.253	14	0.659	7
5	0.485	0.777	41	0.329	20	0.342	3
6	0.259	0.442	47	0.442	12	0.156	5
7	0.487	0.851	31	0.466	25	0.049	8
8	0.302	0.593	36	0.256	23	-0.044	5
9	0.205	0.561	47	0.568	8	0.106	9

**a**, Variability in the number of significantly (sig.) activated voxels reported across teams ( $n = 65$  teams). **b**, Mean Spearman correlation between the unthresholded statistical maps for all pairs of teams and separately for pairs of teams within each cluster, for each hypothesis ( $n = 64$  teams).

# Article

**Extended Data Table 5 | Results of prediction markets and additional data**

a

Hypothesis	FV	CI	Non-teams market prediction	Teams market prediction
1	0.37	[0.26-0.48]	0.727 *	0.814 *
2	0.21	[0.12-0.31]	0.73 *	0.753 *
3	0.23	[0.13-0.33]	0.881 *	0.743 *
4	0.33	[0.22-0.44]	0.882 *	0.789 *
5	0.84	[0.76-0.93]	0.686 *	0.952 *
6	0.33	[0.22-0.44]	0.685 *	0.805 *
7	0.06	[0.00-0.11]	0.563 *	0.073
8	0.06	[0.00-0.11]	0.584 *	0.274 *
9	0.06	[0.00-0.11]	0.476 *	0.188 *

b

Hypothesis	1	2	3	4	5	6	7	8	9
Spearman rho	0.58	0.56	0.58	0.64	0.47	0.74	0.23	0.37	0.31
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.01	0.02
Share of consistent holdings	0.71	0.68	0.70	0.80	0.89	0.74	0.80	0.80	0.75
Z (signed rank test)	3.40	2.78	2.82	4.24	6.81	3.24	4.34	4.34	3.64
p-value (signed rank test)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average holdings if consistent	5.61	21.14	25.80	13.11	-115.50	7.31	34.61	24.23	23.54
Average holdings if inconsistent	1.04	-6.90	-8.03	0.03	18.26	1.58	-14.63	-8.29	-11.61

c

Hypothesis	Tokens invested (Non-teams)	Volume (Non-teams)	# Traders (Non-teams)	# Transactions (Non-teams)	Tokens invested (Teams)	Volume (Teams)	# Traders (Teams)	# Transactions (Teams)
1	8.568	20.175	55	139	12.643	25.671	64	213
2	10.51	22.544	53	98	11.632	22.908	58	171
3	12.818	24.709	58	132	7.773	15.837	52	141
4	11.134	20.397	49	112	8.126	15.479	52	127
5	6.873	14.636	38	71	14.48	30.76	76	244
6	6.806	12.663	35	72	8.097	16.676	46	134
7	7.99	15.209	41	98	7.131	15.864	52	160
8	8.791	19.072	45	91	7.085	14.598	52	141
9	10.427	21.118	50	131	9.506	18.812	56	178

**a.** Summary of the prediction market results. FV refers to the fundamental value, that is, the actual fraction of teams (out of  $n = 70$  teams) that reported significant results for the hypothesis. CI refers to the 95% confidence interval corresponding to the fundamental value (estimated with a normal approximation to the binomial distribution). Values marked with an asterisk are not within the corresponding 95% CI. **b.** Consistency of traders' holdings and team results. The top two rows show two-sided Spearman rank correlations between traders' final holdings and the binary result reported by their team, and the corresponding P value for each hypothesis. The bottom five rows show the share of traders' holdings that are consistent with the results reported by their team. Consistent refers to positive (negative) holdings if the team reported a significant (non-significant) result; z and P values refer to Wilcoxon signed-rank tests for the share of consistent holdings being equal to 0.5; and average holdings if (in)consistent refer to the mean final holdings, separated for consistent and inconsistent traders. **c.** Additional data for each of the nine hypotheses. Tokens invested indicates the average number of tokens invested per transaction; volume refers to the mean number of shares bought or sold per transaction; # traders refers to the number of traders who bought or sold shares of the particular asset at least once; and # transactions describes the overall number of transactions recorded.

Corresponding author(s): Tom Schonberg

Last updated by author(s): Mar 25, 2020

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
  - Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted
  - Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

A full description of the experimental procedures, validations and the fMRI dataset is available in a Data Descriptor (<https://doi.org/10.1038/s41597-019-0113-7>). Code used for fMRI data collection are available at [https://github.com/rotemb9/NARPS\\_scientific\\_data](https://github.com/rotemb9/NARPS_scientific_data).

Data analysis

Fully reproducible code for the analyses of the analysis teams' submitted results and statistical maps, as well as the prediction markets, are available at DOI: 10.5281/zenodo.3709273. The full list of software and versions used within the code are available in the dockerfile: <https://github.com/poldrack/narps/blob/master/Dockerfile>

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The fMRI dataset is openly available via OpenNeuro at DOI:10.18112/openneuro.ds001734.v1.0.4. Additional data are included with the analyses code at DOI:10.5281/zenodo.3709273

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences     Behavioural & social sciences     Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Quantitative
Research sample	The fMRI dataset included neuroimaging and behavioral data of 108 participants. Demographic information of the participants can be found at DOI:10.18112/openneuro.ds001734.v1.0.4. 70 analysis teams analyzed the dataset. 96 "team members" and 91 "non-team members" signed up to participate in the prediction markets. N = 83 "team members" and N = 65 "non-team members" actively participated in the markets. Members of the analysis teams and traders in the predictions market were researchers in the field from around the world.
Sampling strategy	Relevant information for the fMRI dataset is available at the Data Descriptor ( <a href="https://doi.org/10.1038/s41597-019-0113-7">https://doi.org/10.1038/s41597-019-0113-7</a> ). With regard to the number of analysis teams and traders in the prediction markets, we aimed to recruit as many as possible within the time frame.
Data collection	Relevant information for the fMRI dataset is available at the Data Descriptor ( <a href="https://doi.org/10.1038/s41597-019-0113-7">https://doi.org/10.1038/s41597-019-0113-7</a> ). Shortly, data was collected using MRI scanner and computers.
Timing	The fMRI dataset was collected between November 2017 and May 2018. Analysis teams were recruited and analyzed the data between November 2018 and March 2019. The prediction markets were open between May 2nd to May 12th 2019.
Data exclusions	One team was excluded from all analyses since their reported results were not based on a whole-brain analysis as instructed. Of the remaining 69 teams, thresholded maps of 65 teams and unthresholded (z / t) maps of 64 teams were included in the analyses (see Extended Data Table 3b for detailed reasons for exclusion of the other teams).
Non-participation	12 out of the 82 analysis teams that signed the non-disclosure form and were provided with access to the data did not submit their results by the deadline. 13 traders in the "team members" and 26 traders in the "non-team members" prediction markets registered but did not actively participate in the prediction markets.
Randomization	fMRI dataset- participants were pseudo-randomly (alternately) assigned to one of two experimental conditions (Equal Indifference or Equal Range). Analysis teams were not allocated into experimental groups.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines	<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input type="checkbox"/>	<input checked="" type="checkbox"/> MRI-based neuroimaging
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms		
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	fMRI dataset- demographic information of the participants can be found at DOI:10.18112/openneuro.ds001734.v1.0.4. 108 participants where included in the dataset: 54 in the Equal Indifference group (30 females, mean age = 26.06 years, SD age = 3.02 years) and 54 in the Equal Range group (30 females, mean age = 25.04 years, SD age = 3.99 years). All participants were right-handed, had normal or corrected-to-normal vision and reported no history of psychiatric or neurologic diagnoses, or use any medications that would interfere with the experiment.
----------------------------	---

## Recruitment

Analysis teams were recruited via social media, mainly Twitter and Facebook, as well as during the 2018 annual meeting of The Society for Neuroeconomics. Prediction market traders were recruited via social media (mainly Facebook and Twitter) and e-mails. This recruitment method may increase the chances of specific researchers to participate in an analysis team or in the prediction markets, for example researchers that are more active in social media or attended the 2018 meeting of The Society for Neuroeconomics. Researchers who advocate for replication attempts and "open science" practices may also be more inclined to join such study. However, our results strongly suggest that they were not biased. For example, the fact that several hypotheses were only affirmed by roughly 5% of teams, while Hypothesis #5 was affirmed by 84% of teams, suggests that there was no overall bias towards either affirmation or rejection of hypotheses. In addition, each of the 70 analysis teams chose to use a different analysis pipeline, which suggests evidence against a potential bias in methods used by the specific analysis teams that joined the study. With regard to the prediction markets, traders that were exposed to the recruitment ads on social media may be biased with regard to their predictions, but as there is a debate in the published literature regarding most of the hypotheses included in our study, we do not have a specific reason to assume such bias.

## Ethics oversight

MRI data collection was approved by the Helsinki committee at Sheba Tel Hashomer Medical Center and the ethics committee at Tel Aviv University, and all participants gave written informed consent (as described in the Scientific Data Descriptor of this dataset). The Board for Ethical Questions in Science at the University of Innsbruck approved the data collection in regards of the prediction markets, and certified that the project is in correspondence with all requirements of the ethical principles and the guidelines of good scientific practice. The Stanford University IRB determined that the analysis of the submitted team results did not meet the definition of human subject research, and thus no further IRB review was required.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Magnetic resonance imaging

## Experimental design

## Design type

Task

## Design specifications

The fMRI dataset was published in a Data Descriptor (<https://doi.org/10.1038/s41597-019-0113-7>)

## Behavioral performance measures

The fMRI dataset was published in a Data Descriptor (<https://doi.org/10.1038/s41597-019-0113-7>)

## Acquisition

## Imaging type(s)

functional and structural

## Field strength

3T

## Sequence &amp; imaging parameters

Imaging data were acquired using a 3T Siemens Prisma MRI scanner with a 64-channel head coil, at the Strauss Imaging Center on the campus of Tel Aviv University. Functional data during the mixed gambles task were acquired using T2\*-weighted echo-planar imaging sequence with multi-band acceleration factor of 4 and parallel imaging factor (iPAT) of 2, TR=1000ms, TE=30ms, flip angle=68 degrees, field of view (FOV)=212×212mm, in plane resolution of 2×2mm 30 degrees off the anterior commissure-posterior commissure line to reduce the frontal signal dropout<sup>27</sup>, slice thickness of 2mm, 64 slices and a gap of 0.4mm between slices to cover the entire brain. For each functional run, we acquired 453 volumes.

## Area of acquisition

Whole brain

## Diffusion MRI

 Used Not used

## Preprocessing

## Preprocessing software

Each team performed their own preprocessing. Raw data and data preprocessed with fMRIprep v. 1.1.4 were shared with the teams.

## Normalization

Each team performed their own preprocessing. Raw data and data preprocessed with fMRIprep v. 1.1.4 were shared with the teams.

## Normalization template

Each team performed their own preprocessing. Raw data and data preprocessed with fMRIprep v. 1.1.4 were shared with the teams.

## Noise and artifact removal

Each team performed their own preprocessing. Raw data and data preprocessed with fMRIprep v. 1.1.4 were shared with the teams.

## Volume censoring

Each team performed their own preprocessing. Raw data and data preprocessed with fMRIprep v. 1.1.4 were shared with the teams.

## Statistical modeling &amp; inference

## Model type and settings

Each team performed their own analysis.

## Effect(s) tested

Each team performed their own analysis.

Specify type of analysis:  Whole brain  ROI-based  Both

Statistic type for inference  
(See [Eklund et al. 2016](#))

Each team performed their own analysis.

Correction

Each team performed their own analysis.

## Models & analysis

n/a | Involved in the study

- Functional and/or effective connectivity
- Graph analysis
- Multivariate modeling or predictive analysis

## Article

# C9orf72 suppresses systemic and neural inflammation induced by gut bacteria

<https://doi.org/10.1038/s41586-020-2288-7>

Received: 17 May 2019

Accepted: 9 April 2020

Published online: 13 May 2020

 Check for updates

Aaron Burberry<sup>1,2</sup>, Michael F. Wells<sup>1,2</sup>, Francesco Limone<sup>1,2,3</sup>, Alexander Couto<sup>1,2</sup>, Kevin S. Smith<sup>1,2</sup>, James Keaney<sup>4</sup>, Gaëlle Gillet<sup>4</sup>, Nick van Gastel<sup>1,5</sup>, Jin-Yuan Wang<sup>1,2</sup>, Olli Pietilainen<sup>1,2</sup>, Menglu Qian<sup>1,2,6</sup>, Pierce Eggan<sup>1,2</sup>, Christopher Cantrell<sup>1,2</sup>, Joanie Mok<sup>1,2</sup>, Irena Kadiu<sup>4</sup>, David T. Scadden<sup>1,5</sup> & Kevin Eggan<sup>1,2,6</sup>✉

A hexanucleotide-repeat expansion in *C9ORF72* is the most common genetic variant that contributes to amyotrophic lateral sclerosis and frontotemporal dementia<sup>1,2</sup>. The *C9ORF72* mutation acts through gain- and loss-of-function mechanisms to induce pathways that are implicated in neural degeneration<sup>3–9</sup>. The expansion is transcribed into a long repetitive RNA, which negatively sequesters RNA-binding proteins<sup>5</sup> before its non-canonical translation into neural-toxic dipeptide proteins<sup>3,4</sup>. The failure of RNA polymerase to read through the mutation also reduces the abundance of the endogenous *C9ORF72* gene product, which functions in endolysosomal pathways and suppresses systemic and neural inflammation<sup>6–9</sup>. Notably, the effects of the repeat expansion act with incomplete penetrance in families with a high prevalence of amyotrophic lateral sclerosis or frontotemporal dementia, indicating that either genetic or environmental factors modify the risk of disease for each individual. Identifying disease modifiers is of considerable translational interest, as it could suggest strategies to diminish the risk of developing amyotrophic lateral sclerosis or frontotemporal dementia, or to slow progression. Here we report that an environment with reduced abundance of immune-stimulating bacteria<sup>10,11</sup> protects *C9orf72*-mutant mice from premature mortality and significantly ameliorates their underlying systemic inflammation and autoimmunity. Consistent with *C9orf72* functioning to prevent microbiota from inducing a pathological inflammatory response, we found that reducing the microbial burden in mutant mice with broad spectrum antibiotics—as well as transplanting gut microflora from a protective environment—attenuated inflammatory phenotypes, even after their onset. Our studies provide further evidence that the microbial composition of our gut has an important role in brain health and can interact in surprising ways with well-known genetic risk factors for disorders of the nervous system.

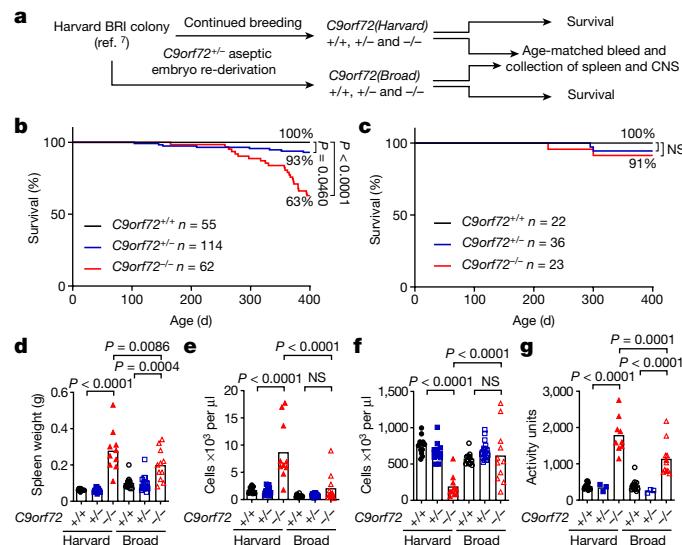
To understand the consequences of the long-term reduction in *C9ORF72* activity found in patients with amyotrophic lateral sclerosis (ALS) or frontotemporal dementia (FTD) who carry the repeat expansion, mice that contain loss-of-function (LOF) mutations in the orthologous gene (*C9orf72*) have previously been studied<sup>6,7,12,13</sup>. It was previously reported<sup>7</sup>, and later corroborated, that reduced *C9orf72* function led to age-dependent inflammation, characterized by cytokine storm<sup>7,14</sup>, neutrophilia<sup>6,7,14</sup>, pseudothrombocytopenia<sup>7</sup>, autoimmunity<sup>7,14</sup>, splenomegaly<sup>6,7,13,14</sup> and neuroinflammation<sup>6,7</sup>. Informed by these observations and validating their importance, it was subsequently found that patients with ALS or FTD who had mutations in *C9ORF72* were significantly more likely to have been diagnosed with autoimmune disease before their neurological diagnosis<sup>15,16</sup>.

However, the long-term survival of *C9orf72*LOF mutant mice varied markedly between reports, despite many groups studying the same allele on a similar genetic background. In some studies, it was found<sup>7,12</sup> that the loss of one (+/−) or both (−/−) alleles of *C9orf72* increased the risk of premature mortality, whereas others<sup>13</sup> noted a reduced survival of *C9orf72*<sup>+/−</sup> but not *C9orf72*<sup>−/−</sup> mice—and another group<sup>6</sup> reported no survival differences between control and mutant mice (Extended Data Fig. 1). These findings suggested that the environment in which the mice were reared might be an important modifier of survival when *C9orf72* levels are reduced. To test this hypothesis, we aseptically re-derived *C9orf72*-mutant mice into a new facility at the Broad Institute (hereafter referred to as *C9orf72*(Broad) mice) while continuing our colony at the Harvard Biological Research Infrastructure (BRI)

<sup>1</sup>Harvard Stem Cell Institute, Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA. <sup>2</sup>Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. <sup>3</sup>Hubrecht Institute for Developmental Biology and Stem Cell Research, Royal Netherlands Academy of Arts and Sciences, Utrecht, The Netherlands.

<sup>4</sup>Neuroscience Therapeutic Area, New Medicines, UCB Biopharma SPRL, Braine-l'Alleud, Belgium. <sup>5</sup>Center for Regenerative Medicine, Massachusetts General Hospital, Boston, MA, USA.

<sup>6</sup>Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA, USA. ✉e-mail: Eggan@mcb.harvard.edu



**Fig. 1 | Environment governs survival, inflammation and autoimmunity in C9orf72LOF mice.** **a**, Aseptic embryo transfer of *C9orf72* neo-deleted allele from Harvard BRI to the Broad Institute. Male and female mice were aged for survival or tissue collection. CNS, central nervous system. **b, c**, Survival of mice at Harvard BRI (*C9orf72*<sup>+/+</sup>, *n* = 55; *C9orf72*<sup>+/−</sup>, *n* = 114; *C9orf72*<sup>−/−</sup>, *n* = 62) (**b**) or the Broad Institute (*C9orf72*<sup>+/+</sup>, *n* = 22; *C9orf72*<sup>+/−</sup>, *n* = 36; *C9orf72*<sup>−/−</sup>, *n* = 23) (**c**) (Gehan-Breslow-Wilcoxon). NS, not significant. **d–g**, Age-matched (48-week-old) mice reared at Harvard BRI (*C9orf72*<sup>+/+</sup>, *n* = 12; *C9orf72*<sup>+/−</sup>, *n* = 13; *C9orf72*<sup>−/−</sup>, *n* = 10) or the Broad Institute (*C9orf72*<sup>+/+</sup>, *n* = 12; *C9orf72*<sup>+/−</sup>, *n* = 18; *C9orf72*<sup>−/−</sup>, *n* = 11) were assessed for spleen weight (**d**), blood neutrophil count (**e**), blood platelet count measured at 0 °C (**f**) and plasma anti-double-stranded (ds)DNA antibody activity (**g**). In **d–g**, one-way ANOVA with Sidak's multiple comparisons. Each dot represents one mouse.

facility (hereafter referred to as *C9orf72(Harvard)* mice) (Fig. 1a). To assess the reproducibility of the original findings at Harvard, we aged an independent cohort of *C9orf72(Harvard)* mice (*C9orf72(Harvard)*<sup>+/+</sup>, *n* = 55; *C9orf72(Harvard)*<sup>+/−</sup>, *n* = 114; *C9orf72(Harvard)*<sup>−/−</sup>, *n* = 62) and again found that *C9orf72(Harvard)*<sup>+/−</sup> mice (Gehan-Breslow-Wilcoxon *P* = 0.0460) and *C9orf72(Harvard)*<sup>−/−</sup> mice (Gehan-Breslow-Wilcoxon *P* < 0.0001) were at an increased risk of premature mortality (Fig. 1b). The causes of death in these mice—which included cervical lymphadenopathy, wasting and severe ataxia—were indistinguishable from those that were observed previously<sup>7</sup>, and were closely tied to their underlying autoimmune condition (Extended Data Fig. 2a, b). By contrast, we observed no early mortality or motor behaviour deficit in either heterozygous or homozygous mutant mice at the Broad Institute (*C9orf72(Broad)*<sup>+/+</sup>, *n* = 22, *C9orf72(Broad)*<sup>+/−</sup>, *n* = 36, *C9orf72(Broad)*<sup>−/−</sup>, *n* = 23) (Fig. 1c, Extended Data Fig. 2c). As a result, *C9orf72*<sup>−/−</sup> mice were significantly more likely to die prematurely when reared at Harvard than at the Broad Institute (Gehan-Breslow-Wilcoxon test, *P* = 0.0179). We therefore conclude that signals from the environment can be significant modifiers of lifespan when *C9orf72* function declines.

To determine whether the improved survival that we observed in *C9orf72(Broad)* mice was associated with a diminution of inflammatory and autoimmune endophenotypes<sup>7</sup>, we jointly analysed age-matched mice reared at each facility (Fig. 1d–g, Extended Data Fig. 2d, e). As previously reported<sup>7</sup>, *C9orf72(Harvard)* mice exhibited autoimmune and inflammatory phenotypes, including significantly elevated levels of IL-23, IL-10, IL-22, G-CSF, IL-17a, TNF, IFN $\gamma$ , IL-1 $\beta$  and IL-12p70 (*P* < 0.05) (Extended Data Fig. 2e) as well as splenomegaly (*P* < 0.0001) (Fig. 1d), neutrophilia (*P* < 0.0001) (Fig. 1e), pseudo-thrombocytopenia (*P* < 0.0001) (Fig. 1f, Extended Data Fig. 2f, g) and development of auto-antibodies (*P* < 0.0001) (Fig. 1g). Notably and in every case, these inflammatory phenotypes were significantly reduced

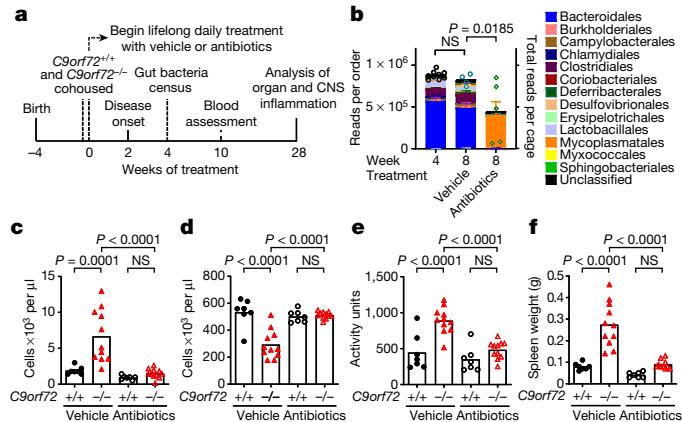
in *C9orf72(Broad)*<sup>−/−</sup> mice relative to their *C9orf72(Harvard)* mutant counterparts (Fig. 1d–g). In fact, the reduction of inflammation in the pro-survival Broad Institute environment was sufficiently reduced that many inflammatory phenotypes that we routinely observed in mutant mice at Harvard were no longer significantly different between *C9orf72*<sup>−/−</sup> and *C9orf72*<sup>+/−</sup> mice at the Broad Institute. It is notable that the few phenotypes that remained significantly different between *C9orf72(Broad)*<sup>−/−</sup> and *C9orf72(Broad)*<sup>+/−</sup> mice, such as modest splenomegaly (one-way analysis of variance (ANOVA) with Sidak's multiple comparisons, *P* = 0.0004), were those that have been most widely reported<sup>6,13,14</sup>. Thus, an environment that improved survival also ameliorated the underlying inflammatory and autoimmune disease found in *C9orf72* mutant mice.

## Antibiotics prevent inflammation

We next considered variables between the two environments that might have contributed to such marked differences in the severity of mutant phenotypes. We found that diet, light cycle and many other features of the two environments were similar. However, a review of microbial screening reports from the two facilities indicated that murine norovirus (Fisher's exact test, *P* = 0.0140), *Helicobacter* spp. (Fisher's exact test, *P* < 0.0001), *Pasteurella pneumotropica* (Fisher's exact test, *P* = 0.0070) and *Tritrichomonas muris* (Fisher's exact test, *P* < 0.0001) were significantly more common in *C9orf72(Harvard)* mice than in *C9orf72(Broad)* mice (Supplementary Table 1). It is important to note that the differences between the two colonies were well within norms for Assessment and Accreditation of Laboratory Animal Care processes. The differential components of the microflora that we found at Harvard are not generally considered pathogenic, consistent with the normal health and lifespan of control mice in that environment<sup>7</sup> (Fig. 1b). However, *Helicobacter* spp. have previously been suggested to have immune-stimulating properties<sup>10</sup>, which raises the possibility that changes in gut microflora between the two environments might underlie the increased rate of mortality and inflammatory phenotypes that we found in *C9orf72(Harvard)* mutant mice.

To learn whether the resident microflora contributed to the inflammation and autoimmunity seen in *C9orf72(Harvard)* mutant mice, we weaned new *C9orf72(Harvard)* mice (*C9orf72(Harvard)*<sup>+/+</sup>, *n* = 14; *C9orf72(Harvard)*<sup>−/−</sup>, *n* = 22) and administered either vehicle or broad-spectrum antibiotics before the onset of inflammatory disease (day 30), then monitored related phenotypes for 200 days (Fig. 2a). As expected, antibiotics significantly reduced the abundance and diversity of bacterial species (including *Helicobacter* spp.), without affecting levels of murine norovirus. The guts of vehicle-treated control mice were largely unaltered (Fig. 2b, Extended Data Fig. 3a). We found that vehicle had no effect on the development of either inflammatory or autoimmune phenotypes in *C9orf72(Harvard)*<sup>−/−</sup> mice, including cytokine storm (Extended Data Fig. 3b), neutrophilia (Fig. 2c), pseudo-thrombocytopenia (Fig. 2d), autoimmunity (Fig. 2e) and splenomegaly (Fig. 2f, Extended Data Fig. 3c). By contrast, providing lifelong antibiotics treatment to *C9orf72(Harvard)*<sup>−/−</sup> mice completely suppressed the emergence of all of these phenotypes (Fig. 2c–f, Extended Data Fig. 3b–i). Thus, our experiments suggest that signals derived from gut bacteria promote inflammation and autoimmunity when *C9orf72* function is diminished. However, we found that chronic administration of antibiotics resulted in previously reported health consequences (including hepatotoxicity)<sup>17</sup>, which prevented us from assessing behavioural and survival outcomes.

We next asked whether the acute suppression of gut microbiota could ameliorate inflammatory and autoimmune phenotypes after their establishment in *C9orf72(Harvard)* mutant mice. To this end, we obtained another independent cohort of *C9orf72(Harvard)* mice (*C9orf72(Harvard)*<sup>+/+</sup>, *n* = 25; *C9orf72(Harvard)*<sup>−/−</sup>, *n* = 24; day 250), demonstrated these mice displayed the expected inflammatory phenotypes



**Fig. 2 | Lifelong suppression of gut microflora prevents inflammation and autoimmunity in C9orf72LOF mice.** **a**, Male and female C9orf72(Harvard)<sup>+/+</sup> and C9orf72(Harvard)<sup>-/-</sup> neo-deleted mice of weaning age were cohoused by treatment group, then administered vehicle (C9orf72(Harvard)<sup>+/+</sup>,  $n=7$ ; C9orf72(Harvard)<sup>-/-</sup>,  $n=11$ ) or antibiotics (C9orf72(Harvard)<sup>+/+</sup>,  $n=7$ ; C9orf72(Harvard)<sup>-/-</sup>,  $n=11$ ) daily for life. **b–f**, Mice were assessed for gut microbial composition at 4 weeks (**b**) and blood measurements at 8 weeks (**c–e**), and killed for organ and central nervous system assessment at 28 weeks (**f**). **b**, 16S rDNA sequencing of bacteria diversity in faeces. Each dot represents total sequencing reads per cage. One-way ANOVA with Dunnett's multiple comparisons. **c**, Blood neutrophil count. **d**, Blood platelet count measured at 0 °C. **e**, Plasma anti-dsDNA antibody activity. **f**, Spleen weight. In **c–f**, one-way ANOVA with Sidak's multiple comparisons. Each dot represents one mouse.

relative to control mice and showed that they exhibited poor performance on the accelerating rotarod (Fig. 3a–d, Extended Data Fig. 4a). Then, we began acute administration of broad-spectrum antibiotics and monitored associated phenotypes over the course of 60 days. We found that this treatment significantly reduced each of the inflammatory and autoimmune phenotypes in mutant mice (Fig. 3b–d), including splenomegaly ( $P=0.0002$ ) (Extended Data Fig. 4c, d) and improved rotarod performance ( $P=0.0398$ ) (Extended Data Fig. 4a). By contrast, treatment with vehicle had no effect on these measures (Fig. 3b–d, Extended Data Fig. 4a–d).

### Faecal transplants mitigate inflammation

To more directly investigate whether the phenotypic improvements were due to the microbial communities of the gut (rather than unrelated consequences of antibiotics treatment), we performed faecal transplant experiments. We produced another cohort of C9orf72(Harvard) mice (C9orf72(Harvard)<sup>+/+</sup>,  $n=27$ ; C9orf72(Harvard)<sup>-/-</sup>,  $n=32$ ; day 100) and demonstrated that these mice displayed the expected inflammatory phenotypes relative to control mice. We then suppressed the gut microflora of these mice with transient antibiotic treatment, and transplanted with faeces from either the pro-inflammatory (Harvard BRI) or pro-survival (Broad Institute) environment (Fig. 3e). Transplantation of pro-survival gut microflora significantly improved each of the inflammatory and autoimmune phenotypes (Fig. 3f–h, Extended Data Fig. 4e). By contrast, transplant with microflora from the pro-inflammatory facility did not improve these measures, which suggests that the benefits we observed when transplanting faeces from the protective environment were not merely due to the brief antibiotic treatment that enabled microbial engraftment. Therefore, our studies establish that the inflammatory and autoimmune disease that underlies premature mortality in C9orf72(Harvard) mutant mice can be therapeutically prevented, and that signals from particular gut microbiota help to maintain this disease.

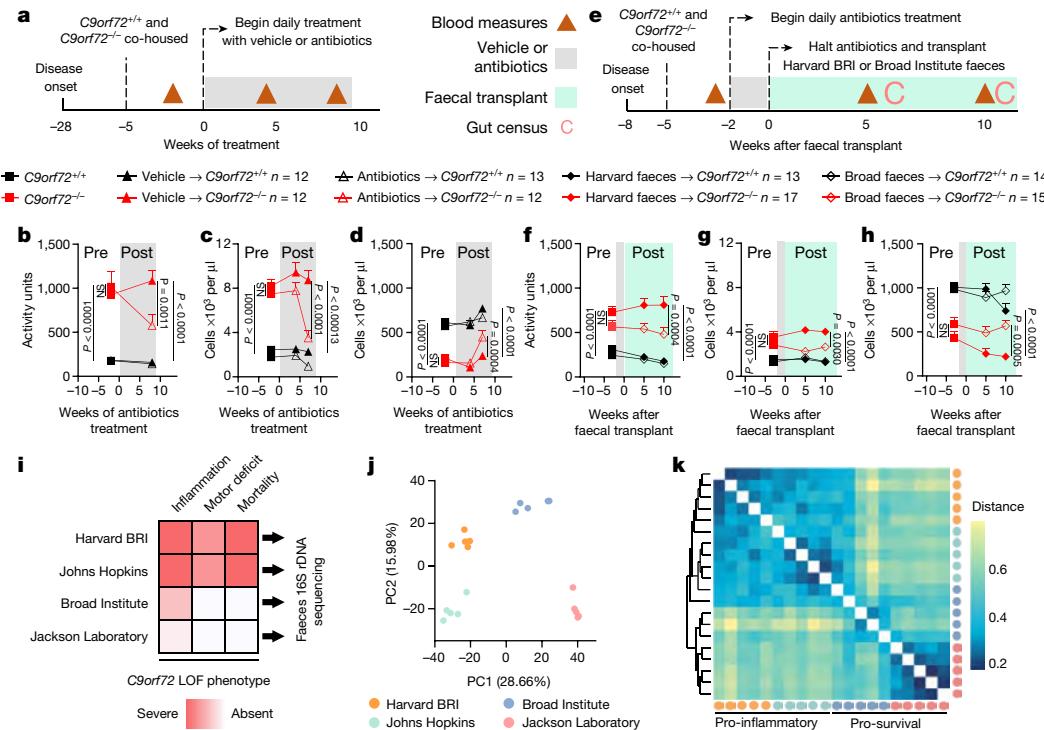
### Profiling gut bacteria

To identify the bacterial species of the gut that are associated with severe phenotypes in C9orf72(Harvard) mutant mice, we surveyed the composition of faeces from two pro-inflammatory environments in which the mutant mice perished<sup>7,12</sup>, as well as from one previously published<sup>6</sup> and one additional pro-survival environment (Fig. 3i). Principal component analysis readily separated samples from the four environments, with the largest principal component (principal component 1, 28.7% of variance) separating the two pro-inflammatory environments from the two pro-survival environments (Fig. 3j). Deeper investigation of this axis of variance revealed a shared significant decrease in  $\alpha$ -diversity in the two pro-inflammatory environments, and unsupervised hierarchical clustering demonstrated that samples from the pro-inflammatory environments showed  $\beta$ -diversity disparate from that in pro-survival environments (Fig. 3k, Extended Data Fig. 5f). Exemplifying these considerable differences in community structure, 62 of 301 bacterial species we identified (20.6%) were significantly altered in their abundance when jointly comparing the two pro-survival environments to the two pro-inflammatory environments ( $P<0.0002$ ) (Extended Data Fig. 5a–e). Consistent with initial observations (Supplementary Table 1), *Helicobacter* spp. were found in both of the pro-inflammatory environments (Extended Data Fig. 5g, h, Supplementary Fig. 1) but were absent in pro-survival environments.

We next characterized the extent of microbial reconstitution in our faecal transplant recipients (Extended Data Fig. 6a–f). Hierarchical clustering of  $\beta$ -diversity revealed that the microbial composition in mice that received faecal transplants from the Broad Institute environment were more similar to faeces from mice housed at the Broad Institute than to faeces from mice housed at Harvard BRI or faeces from mice that received faecal transplants from the Harvard environment (Extended Data Fig. 6d). Analysis of individual bacteria similarly supported the success of our transplants: 85% (199 out of 234) of bacterial species identified in faeces from mice housed at Harvard BRI being detected in mice that received faecal transplants from the Harvard environment, and 75% (178 out of 236) of bacterial species identified in faeces from mice housed at the Broad Institute detected in mice that received faecal transplants from the Broad Institute environment (Extended Data Fig. 6f). Semi-quantitative PCR for *Helicobacter* spp. ribosomal (r)DNA further confirmed the reconstitution of Harvard-specific microorganisms in the recipients of faecal transplants from the Harvard environment, and their elimination from the recipients of faecal transplants from the Broad Institute environment (Extended Data Fig. 5i, Supplementary Fig. 1).

### Gut components regulate myeloid cytokines

To mechanistically explore how varied faecal components in separate environments alter cytokine burden and autoimmunity in C9orf72<sup>-/-</sup> mice, we stimulated bone-marrow-derived macrophages from C9orf72(Harvard) mice with chemical analogues of microbial components and found that both C9orf72(Harvard)<sup>+/+</sup> and C9orf72(Harvard)<sup>-/-</sup> bone-marrow-derived macrophages released higher levels of several pro-inflammatory cytokines than C9orf72(Harvard)<sup>+/+</sup> control cells in response to bacterial lipopeptide, single-stranded (ss)RNA and ssDNA (Extended Data Fig. 7a–c). Given these findings, we next asked whether faecal material from mice housed at Harvard BRI contained higher levels of innate-immune stimulating factors than faeces from mice at the Broad Institute. To this end, we individually administered normalized concentrations of faecal Eubacteria from both institutions to C9orf72(Harvard)<sup>-/-</sup> and C9orf72(Harvard)<sup>+/+</sup> bone-marrow-derived macrophages. We found that C9orf72<sup>-/-</sup> bone-marrow-derived macrophages produced significantly higher levels of TNF when exposed to faecal material from Harvard BRI than when exposed to faeces from the Broad Institute (Extended Data Fig. 7d). In addition, serial dilutions



**Fig. 3 | Gut bacteria propagates inflammation and autoimmunity in C9orf72LOF mice.** **a**, Age-matched (36-week-old) female *C9orf72(Harvard)*<sup>+/+</sup> and *C9orf72(Harvard)*<sup>-/-</sup> neo-deleted mice were cohoused by treatment group, then administered vehicle (*C9orf72(Harvard)*<sup>+/+</sup>,  $n=12$ ; *C9orf72(Harvard)*<sup>-/-</sup>,  $n=12$ ) or antibiotics (*C9orf72(Harvard)*<sup>+/+</sup>,  $n=13$ ; *C9orf72(Harvard)*<sup>-/-</sup>,  $n=12$ ) daily. **b-d**, These mice were then assessed for plasma anti-dsDNA antibody activity (b), blood neutrophil count (c) and blood platelet count measured at 0 °C (d). **e**, Age-matched (13-week-old) female *C9orf72(Harvard)*<sup>+/+</sup> and *C9orf72(Harvard)*<sup>-/-</sup> neo-deleted mice were cohoused by treatment group, administered antibiotics for two weeks, and then gavaged with faeces from the Harvard BRI environment (*C9orf72(Harvard)*<sup>+/+</sup>,  $n=13$ ; *C9orf72(Harvard)*<sup>-/-</sup>,

$n=17$ ) or Broad Institute environment (*C9orf72(Harvard)*<sup>+/+</sup>,  $n=14$ ; *C9orf72(Harvard)*<sup>-/-</sup>,  $n=15$ ). **f-h**, These mice were then assessed for plasma anti-dsDNA antibody activity (f), blood neutrophil count (g) and blood platelet count measured at 0 °C (h). **i-k**, Faecal pellets ( $n=5$  each) from two pro-inflammatory environments (Harvard BRI and Johns Hopkins) and two pro-survival environments (Broad Institute and Jackson Laboratory) were subjected to 16S rRNA sequencing (i), and then assessed by principal component (PC) analysis (j) and Bray–Curtis dissimilarity matrix of β-diversity (k). In b-d, f-h, one-way ANOVA with Sidak's multiple comparisons. Each dot represents one mouse.

revealed that a combination of faeces from Harvard BRI and a *C9orf72*<sup>-/-</sup> genotype leads to TNF release at the lowest faecal concentrations (Extended Data Fig. 7d).

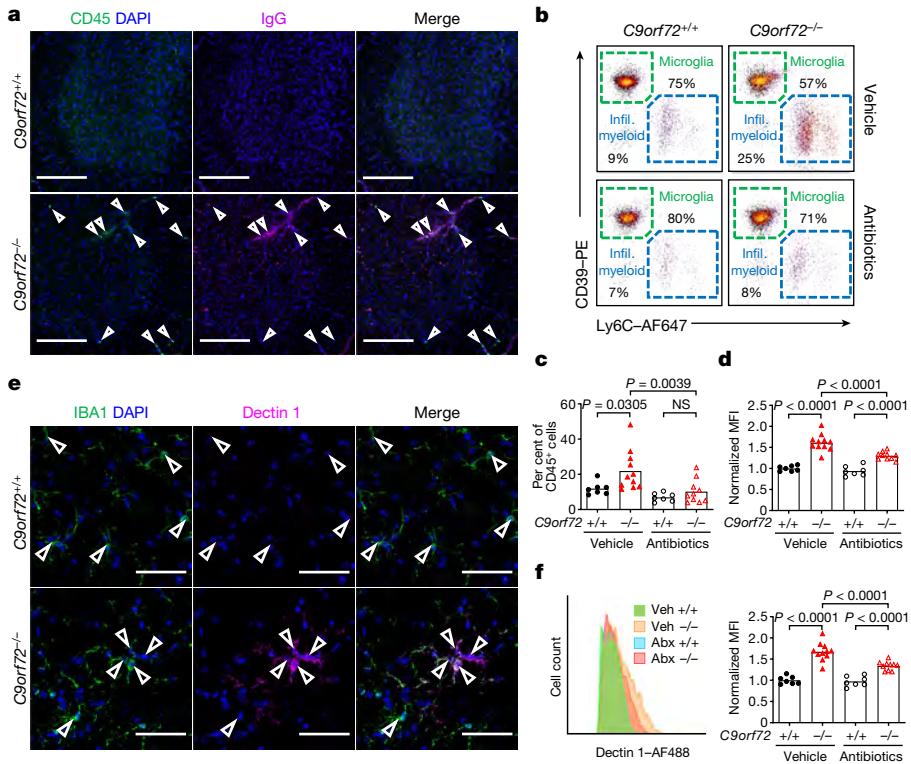
## Environment governs neuro-inflammation

Neuro-inflammation is a pathological hallmark of ALS and FTD associated with mutations in *C9ORF2*<sup>18,19</sup>, with substantial infiltration of peripheral immune cells noted in the spinal cord of patients with ALS<sup>20,21</sup>. We used the pan-haematopoietic marker CD45 to distinguish CD45<sup>mid</sup> resident microglia from peripherally derived CD45<sup>high</sup> cells<sup>22</sup> and found that infiltrating cells were present at sites of focal inflammation within the spinal cord parenchyma of *C9orf72(Harvard)*<sup>-/-</sup> mice (Fig. 4a, Extended Data Fig. 8a–h, Supplementary Videos 1, 2). Mass cytometry analysis revealed that the CD45<sup>high</sup> cells that infiltrated the spinal cord were mostly CD11b<sup>+</sup>Ly6C<sup>+</sup>Ly6G<sup>+</sup>CD39<sup>-</sup> neutrophils and CD3e<sup>+</sup> T cells (Extended Data Fig. 8a–f; gating strategy is in Supplementary Information). Notably, lifelong suppression of gut microflora with antibiotics prevented the accumulation of infiltrating myeloid cells within the spinal cord of *C9orf72(Harvard)*<sup>-/-</sup> mice (Fig. 4b, c).

In addition to infiltrating peripheral immune cells, there are also substantial changes to resident microglia in the nervous systems of patients with ALS or FTD<sup>23</sup>. Previous studies<sup>6,8,9,24</sup> have implicated *C9orf72* and its interactor SMCR8 in regulation of endolysosomal trafficking and autophagy, particularly in myeloid derivatives. We found that

microglia from the spinal cord of *C9orf72(Harvard)*<sup>-/-</sup> mice expressed higher levels of the lysosome-associated proteins LAMP1<sup>6</sup> (Extended Data Fig. 9a, Supplementary Videos 3, 4) and cathepsin B (Extended Data Fig. 9b, Supplementary Videos 5, 6). Lifelong suppression of gut microflora did not significantly decrease LAMP1 or cathepsin B levels in *C9orf72(Harvard)*<sup>-/-</sup> microglia (Extended Data Fig. 9c, d), which suggests that *C9orf72* regulates lysosomal constituents independently from microbial signals.

To examine the activation status of resident microglia in *C9orf72(Harvard)* mutant mice and to ask whether microglial activation might be altered by signals from the microbiota, we measured levels of the pattern recognition receptor dectin 1, the chemokine receptor CCR9 and the lipoprotein lipase LPL, which have previously been associated with pro-inflammatory microglial states<sup>25–27</sup>. Consistent with the notion that microglia become activated when *C9orf72* levels decline, we found that dectin 1 and CCR9 were enriched in microglia from *C9orf72(Harvard)*<sup>-/-</sup> mice (Fig. 4d–f, Extended Data Fig. 9e–g, Supplementary Videos 7, 8). Importantly, dectin 1 and CCR9 expression were significantly reduced in microglia from *C9orf72(Harvard)* mutant mice with gut microflora that was chronically suppressed with antibiotics (Fig. 4d–f). These results demonstrate that when *C9orf72* function is reduced, peripheral immune cells can infiltrate the spinal cord, where they associate with sites of neuro-inflammation, and that treatment with antibiotics, which suppresses the microbiota, modulates both infiltration and microglial activation.



**Fig. 4 | Gut microflora promotes myeloid cell infiltration and microgliosis in *C9orf72*LOF spinal cord.** **a**, Orthogonal projection of CD45 and mouse immunoglobulin G (IgG) in lumbar spinal cord of 55-week-old *C9orf72*(*Harvard*) neo-deleted mice (*C9orf72*(*Harvard*)<sup>+/+</sup>,  $n=3$ ; *C9orf72*(*Harvard*)<sup>-/-</sup>,  $n=3$ ), showing cells infiltrating the lumbar spinal cord. Scale bars, 200  $\mu$ m. **b**, Representative gating of CD45<sup>+</sup>CD11b<sup>+</sup> cells from spinal cord of *C9orf72*(*Harvard*) mice in Fig. 2. Infil., infiltrating. **c**, Quantification of CD45<sup>high</sup>CD11b<sup>+</sup>Ly6C<sup>+</sup> myeloid cells infiltrating the spinal cord, shown in **b**. **d**, Quantification of CCR9 expression on CD45<sup>mid</sup>CD11b<sup>+</sup>CD39<sup>+</sup> microglia from

the spinal cord of *C9orf72*(*Harvard*) mice, shown in Fig. 2. MFI, mean fluorescence intensity. **e**, Orthogonal projection of dectin 1 in IBA1<sup>+</sup> microglia in lumbar spinal cord from 55-week-old *C9orf72*(*Harvard*) neo-deleted mice (*C9orf72*(*Harvard*)<sup>+/+</sup>,  $n=3$ ; *C9orf72*(*Harvard*)<sup>-/-</sup>,  $n=3$ ). Scale bars, 50  $\mu$ m. **f**, Dectin 1 in CD45<sup>mid</sup>CD11b<sup>+</sup>CD39<sup>+</sup> microglia from spinal cord of *C9orf72*(*Harvard*) mice in Fig. 2. <sup>+/+</sup>, *C9orf72*(*Harvard*)<sup>+/+</sup>; <sup>-/-</sup>, *C9orf72*(*Harvard*)<sup>-/-</sup>; abx, antibiotics-treated; veh, vehicle-treated. In **c**, **d**, **f**, one-way ANOVA with Sidak's multiple comparisons. Each dot represents one mouse.

## Discussion

Our results indicate that when *C9orf72* function declines, the environment generally—and the gut microbiota specifically—become potent modifiers of whether autoimmunity, neural inflammation, motor deficits and premature mortality occur. The effect of environment and accompanying changes in microbial microflora are so strong in this mouse model that in one environment, inflammatory disease and death were highly penetrant phenotypes, whereas in another they were essentially absent. We therefore provide the probable explanation for the considerable phenotypic variation that has been observed across groups studying this *C9orf72*LOF allele in mice<sup>6,7,12,13</sup>. These conclusions are important because they re-emphasize that the 50% reduction in the levels of *C9orf72* found in patients with ALS or FTD who have mutations in *C9orf72* are a credible cause for the neural inflammation that are characteristic in their condition. Most provocatively, our findings also suggest that variance in microbiota could explain why some carriers of the *C9orf72* mutation develop ALS or FTD, or overt inflammatory conditions such as lupus<sup>15,16</sup>, while others do not.

It should be re-emphasized that the microorganisms present in the environments we studied here are not considered mouse pathogens per se, and that their abundances were within the scope found in comparable institutions<sup>28</sup>. Importantly, the environmental conditions that triggered severe phenotypes in our *C9orf72*(*Harvard*) mice were reproducible elsewhere. Previous reports<sup>7,12</sup> have noted a relationship between the reduction in *C9orf72* function and an increased rate of premature mortality comparable to that described here. It is notable

that these two pro-inflammatory environments were most similar in their microbial constituents and also shared many microorganisms that were not present in the two pro-survival locations we surveyed. Given the large number of species that we found significantly differ in their abundance between pro-inflammatory and pro-survival environments (Fig. 3i, Extended Data Fig. 5c), future studies will be needed to elucidate the relative contribution of individual bacterial species to variation in the inflammatory and autoimmune phenotypes we report here. However, microorganism-by-microorganism analysis of varying environments and our transplant mice would seem to rule out the previously reported protective effects of *Akkermansia muciniphila*<sup>29,30</sup> (Extended Data Fig. 6f) and potential inflammatory influences of *T. muris* (Extended Data Fig. 5e).

It is increasingly appreciated that gut microorganisms alter the maturation and function of microglia<sup>31</sup>, can influence the activity of neurons in the central nervous system<sup>32</sup> and contribute to neuro-inflammation and neuropathology in models of Alzheimer's<sup>33</sup> and Parkinson's disease<sup>34</sup>. However, only initial surveys of the gut microbiota have been reported in patients with neurological conditions<sup>35</sup> and, thus far, results from initial studies in patients with ALS have been mixed<sup>36,37</sup>. One study has reported significant differences between the microbial constituents of patients with ALS and controls<sup>36</sup>, whereas another found no clear distinctions<sup>37</sup>. Given the genetic heterogeneity exhibited within patients with ALS, it is perhaps not surprising that early studies have not reached consensus.

Consistent with the idea there are complex interactions between the germline genotype of a patient and their gut microflora in ALS, it

was recently reported that *SOD1*-transgenic mice displayed a faster decline when bacterial load was reduced, which was linked to reduced bacterial production of nicotinamide<sup>29</sup>. Although we cannot rule out the presence of protective microorganisms in some environments, our studies suggest that lowering the bacterial load in *C9orf72*-mutant mice was in aggregate protective, probably by reducing the exposure of their genetically sensitized innate immune response to inflammatory factors derived from microorganisms. In sum, our studies suggests that the microbiome may be an important governor of the onset and progression of neurological disease in patients with *C9orf72* mutations, including those experiencing autoimmune and inflammatory conditions before a diagnosis of ALS or FTD<sup>15,16</sup>. To test this idea, a key future experiment will be to identify *C9orf72* repeat expansion carriers within families known to be predisposed to developing ALS or FTD, and to determine whether the gut microbiota differs between individuals that remain healthy and those that acquire the conditions.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2288-7>.

1. DeJesus-Hernandez, M. et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of *C9orf72* causes chromosome 9p-linked FTD and ALS. *Neuron* **72**, 245–256 (2011).
2. Majounie, E. et al. Frequency of the *C9orf72* hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol.* **11**, 323–330 (2012).
3. Mori, K. et al. The *C9orf72* GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science* **339**, 1335–1338 (2013).
4. Ash, P. E. A. et al. Unconventional translation of *C9orf72* GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* **77**, 639–646 (2013).
5. Donnelly, C. J. et al. RNA toxicity from the ALS/FTD *C9orf72* expansion is mitigated by antisense intervention. *Neuron* **80**, 415–428 (2013).
6. O'Rourke, J. G. et al. *C9orf72* is required for proper macrophage and microglial function in mice. *Science* **351**, 1324–1329 (2016).
7. Burberry, A. et al. Loss-of-function mutations in the *C9orf72* mouse ortholog cause fatal autoimmune disease. *Sci. Transl. Med.* **8**, 347ra93 (2016).
8. Nassif, M., Woehlbier, U. & Manque, P. A. The enigmatic role of *C9orf72* in autophagy. *Front. Neurosci.* **11**, 442 (2017).
9. Shi, Y. et al. Haploinsufficiency leads to neurodegeneration in *C9orf72* ALS/FTD human induced motor neurons. *Nat. Med.* **24**, 313–325 (2018).
10. Whary, M. T. & Fox, J. G. Natural and experimental *Helicobacter* infections. *Comp. Med.* **54**, 128–158 (2004).
11. Flannigan, K. L. & Denning, T. L. Segmented filamentous bacteria-induced immune responses: a balancing act between host protection and autoimmunity. *Immunology* **154**, 537–546 (2018).
12. Ugolini, J. et al. Loss of *C9orf72* enhances autophagic activity via deregulated mTOR and TFEB signaling. *PLoS Genet.* **12**, e1006443 (2016).
13. Jiang, J. et al. Gain of toxicity from ALS/FTD-linked repeat expansions in *C9orf72* is alleviated by antisense oligonucleotides targeting GGGGCC-containing RNAs. *Neuron* **90**, 535–550 (2016).
14. Atanasio, A. et al. *C9orf72* ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production, and glomerulonephropathy in mice. *Sci. Rep.* **6**, 23204 (2016).
15. Miller, Z. A. et al. Increased prevalence of autoimmune disease within C9 and FTD/MND cohorts. *Neurology. Neuroimmunol. Neuroinflamm.* **3**, e301 (2016).
16. Fredi, M. et al. *C9orf72* intermediate alleles in patients with amyotrophic lateral sclerosis, systemic lupus erythematosus, and rheumatoid arthritis. *Neuromolecular Med.* **21**, 150–159 (2019).
17. Stine, J. G. & Lewis, J. H. Hepatotoxicity of antibiotics: a review and update for the clinician. *Clin. Liver Dis.* **17**, 609–642, ix (2013).
18. Ransohoff, R. M. How neuroinflammation contributes to neurodegeneration. *Science* **353**, 777–783 (2016).
19. McCauley, M. E. & Baloh, R. H. Inflammation in ALS/FTD pathogenesis. *Acta Neuropathol.* **137**, 715–730 (2019).
20. Zhao, W., Beers, D. R. & Appel, S. H. Immune-mediated mechanisms in the pathoprophgression of amyotrophic lateral sclerosis. *J. Neuroimmune Pharmacol.* **8**, 888–899 (2013).
21. Zondler, L. et al. Peripheral monocytes are functionally altered and invade the CNS in ALS patients. *Acta Neuropathol.* **132**, 391–411 (2016).
22. Zhang, G. X., Li, J., Ventura, E. & Rostami, A. Parenchymal microglia of naïve adult C57BL/6J mice express high levels of B7.1, B7.2, and MHC class II. *Exp. Mol. Pathol.* **73**, 35–45 (2002).
23. Lall, D. & Baloh, R. H. Microglia and *C9orf72* in neuroinflammation and ALS and frontotemporal dementia. *J. Clin. Invest.* **127**, 3250–3258 (2017).
24. Zhang, Y. et al. The *C9orf72*-interacting protein Smcr8 is a negative regulator of autoimmunity and lysosomal exocytosis. *Genes Dev.* **32**, 929–943 (2018).
25. Li, H. et al. Different neurotropic pathogens elicit neurotoxic CCR9- or neurosupportive CXCR3-expressing microglia. *J. Immunol.* **177**, 3644–3656 (2006).
26. Krasemann, S. et al. The TREM2-APOE pathway drives the transcriptional phenotype of dysfunctional microglia in neurodegenerative diseases. *Immunity* **47**, 566–581.e9 (2017).
27. Keren-Shaul, H. et al. A unique microglia type associated with restricting development of Alzheimer's disease. *Cell* **169**, 1276–1290.e17 (2017).
28. Nilsson, H.-O. et al. High prevalence of *Helicobacter* species detected in laboratory mouse strains by multiplex PCR-denaturing gradient gel electrophoresis and pyrosequencing. *J. Clin. Microbiol.* **42**, 3781–3788 (2004).
29. Blacher, E. et al. Potential roles of gut microbiome and metabolites in modulating ALS in mice. *Nature* **572**, 474–480 (2019).
30. Zhai, R. et al. Strain-specific anti-inflammatory properties of two *Akkermansia muciniphila* strains on chronic colitis in mice. *Front. Cell. Infect. Microbiol.* **9**, 239 (2019).
31. Erny, D. et al. Host microbiota constantly control maturation and function of microglia in the CNS. *Nat. Neurosci.* **18**, 965–977 (2015).
32. Olson, C. A. et al. The gut microbiota mediates the anti-seizure effects of the ketogenic diet. *Cell* **173**, 1728–1741.e13 (2018).
33. Harach, T. et al. Reduction of Abeta amyloid pathology in APPPS1 transgenic mice in the absence of gut microbiota. *Sci. Rep.* **7**, 41802 (2017).
34. Sampson, T. R. et al. Gut microbiota regulate motor deficits and neuroinflammation in a model of Parkinson's disease. *Cell* **167**, 1469–1480.e12 (2016).
35. Tremlett, H., Bauer, K. C., Appel-Cresswell, S., Finlay, B. B. & Waubant, E. The gut microbiome in human neurological disease: a review. *Ann. Neurol.* **81**, 369–382 (2017).
36. Fang, X. et al. Evaluation of the microbial diversity in amyotrophic lateral sclerosis using high-throughput sequencing. *Front. Microbiol.* **7**, 1479 (2016).
37. Brenner, D. et al. The fecal microbiome of ALS patients. *Neurobiol. Aging* **61**, 132–137 (2018).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Methods

### Mice

All experimental procedures were approved by the Institutional Animal Care and Use Committee of Harvard University and the Broad Institute, and were in compliance with all relevant ethical regulations. The KOMP and neo-deleted *C9orf72* loss of function strains were generated as previously described<sup>7</sup>. Mice were housed with nestlet bedding, red hut for enrichment, provided water ad libitum and fed ad libitum either with Prolab Isopro RMH 3000 (Harvard BRI) or with PicoLab Rodent Diet 20 (Broad Institute) and kept on a 12-h light–dark cycle. Embryo re-derivation was performed by collecting embryos from super-ovulated *C9orf72*<sup>+/−</sup> females, washing embryos, then surgical transfer using an aseptic technique into the reproductive tract of pseudopregnant recipient females. For experiments involving antibiotics, mice were cohoused for at least a week before initiation of dosing. Mice were administered either vehicle (water) or a freshly prepared cocktail of four antibiotics including ampicillin sodium salt (200 mg/kg/d), neomycin trisulfate salt hydrate (200 mg/kg/d), metronidazole (200 mg/kg/d), and vancomycin hydrochloride from *Streptomyces orientalis* (100 mg/kg/d) (all from Sigma) administered by twice daily gavage. The number (*n*), sex and ages of the mice used in each study are described in figure legends or text. Power calculations (G\*Power 3.1.9.2) using the mean and standard error of endophenotype data was used to estimate necessary cohort sizes for antibiotics and faecal transplant studies. Before administration of antibiotics, mice were assessed for systemic inflammatory measures and mice were allocated into groups so that no significant differences were present before treatment initiation.

### Motor behaviour

Naive mice were trained on the rotarod at constant speed of 4 rpm for 300 s at least 1 d before competitive assessment. For performance trials, the rotarod accelerated from 4 to 40 rpm over 300 s using Ugo Basile mouse RotaRod NG (Harvard FAS BRI) or Panlab Rota Rod (Broad Institute). Each trial day consisted of 3 tests per mouse, with each test separated by at least 20 min. The operator was blinded to mouse genotype during trials.

### Faecal transplantation

Using sterilized forceps, donor faecal pellets were collected directly from the anus, or donor upper and lower intestinal contents were isolated from euthanized mice and immediately frozen on dry ice. Recipient mice received antibiotics twice daily by gavage for two weeks, then a two day secession of antibiotics, then faecal transplantation once per day for two days. Faecal pellets and intestinal contents from donor mice were weighed, pooled, diluted to 200 mg/ml in degassed PBS and administered by oral gavage to recipient mice at 2 mg faeces per g of body weight. All cage changes were performed in HEPA filtered hoods with freshly autoclaved cages, bedding and enrichment.

### Blood and cytokine measures

Peripheral blood was collected via mandible puncture into EDTA-coated tubes. Blood counts were assessed using a Hemavet (Abaxis). Samples were then centrifuged to pellet cells and plasma was collected from the supernatant. Plasma was diluted 1:2 for luminex-based multiplexed fluorescence assay to assess 36 cytokines and chemokines. Plasma was diluted 1:200 to assess mouse anti-dsDNA total IgG autoantibodies (Alpha Diagnostic International).

### Tissue preparation

Mice were anaesthetized with isoflurane followed by transcardial perfusion with HBSS supplemented with 10 U/ml heparin. Spleens were dissociated by repeated trituration with a glass pipetman in HBSS, subjected to 10-min RBC lysis (eBioscience), washed in autoMACS

(Miltenyi), filtered (40 µm) and counted using a Countess (Invitrogen) for antibody staining. For flow cytometry and mass cytometry of the central nervous system, spinal cords were digested by papain and DNase diluted in EBSS (Worthington) for 10 min at 37 °C, triturated with a glass pipetman to generate large tissue chunks, and then allowed to digest for 20 min at 37 °C. DMEM supplemented with glutamax was added, samples triturated to single cells, ovomucoid (Worthington) and DNase diluted in EBSS added to inhibit protease activity, cells filtered, washed in autoMACS buffer, and pelleted at 500g for 15 min at 4 °C. Cell pellets were brought up in isotonic Percoll Plus (Sigma) diluted to 30% in autoMACS and spun for 15 min at room temperature with no brake. Floating myelin was gently removed using plastic transfer pipette. Cell pellets were resuspended, filtered, washed in autoMACS and re-pelleted at 4 °C. Cells were fixed in 4% paraformaldehyde (PFA) (Electron Microscopy Sciences) either before or after antibody staining depending on need. Samples were collected on a BD LSRII or Helios mass cytometer. Data were analysed using FlowJo and/or Cytobank. For immunofluorescence experiments, following HBSS perfusion, mice were perfused with 4% PFA and central nervous system tissue was post-fixed in 4% PFA overnight at 4 °C. The next day, samples were washed with PBS overnight at 4 °C. Tissue was submerged in 30% sucrose for 2 d. After cryoprotection, lumbar regions were mounted in OCT and cryostat-sectioned at 30 µm.

### Immunofluorescence

Spinal cord sections were washed three times in PBS to remove residual OCT. Sections were incubated in a blocking solution (10% donkey serum, 0.1M glycine, 0.1% Tween20 or 0.3% Triton X100, PBS, Image-iT FX Signal Enhancer (Thermo) for 1 h at room temperature. Following blocking, sections were incubated with primary antibodies for 2 d on a rocker at 4 °C. Primary antibodies include: rat CD11b–FITC 1:200 (M1/70, BioLegend), rabbit cathepsin B 1:400 (D1C7Y CST), rat CD45–488 1:200 (30-F11 BioLegend), guinea pig IBA1:500 (234004 Synaptic Systems), rat LAMP1:200 (1D4B SCB), rat CCR9–FITC (9B1 BioLegend), rat Dectin 1/CLEC7A (mabg-mdect Invivogen) and mouse LPL (ab21356 Abcam). Sections were then washed with 0.1% Tween20 in PBS (for stains with CD11b, CD45, CCR9, and cathepsin B) or 0.3% TritonX100 in PBS (for stains with IBA1, LAMP1, CLEC7A (dectin 1) and LPL) at least 5 times. Secondary antibodies include: donkey-anti-rat-AlexaFluor-488, -mouse IgG-555, -rabbit-555, -rabbit-647, -rat-647, -guinea pig-647, all 1:500 dilution (Invitrogen), for 2 h at room temperature. Sections were washed again, and mounted on microscope slides in Fluoromount for curing overnight. Spinal cords were imaged on a ZEISS LSM700 with either a 10× and 40× objective, or an Axio scan Z.1 at 20× objective. Images were stitched and processed on ZEISS ZEN 2.6 image processing software and Bitplane Imaris 9.2. All comparative stains between control and mutant mice were acquired using identical laser and microscope settings, and images were processed with viewer blinded to genotype.

### Flow cytometry

Dissociated single cells were stained in autoMACS on ice using the following antibodies (BioLegend): CD45–BV421 or APC-Cy7 1:200 (30-F11), rabbit cathepsin B 1:100 (D1C7Y CST) and goat-anti-rabbit-AlexaFluor-488 1:500 (Invitrogen), CCR9–FITC 1:200 (9B1), F4/80–PE–Cy5 1:400 (BM8), CD11b–AlexaFluor-700 1:400 (M1/70 Invitrogen), LAMP1–APC–Cy7 1:400 (1D4B), TruStain FCX 1:250 (93), CD39–PE 1:400 (Duha59), Ly6G–PE–Cy7 1:600 (1A8), Ly6C–AlexaFluor-647 (HK1.4). To retrieve the cathepsin B epitope, fixed cells were slowly permeabilized in 90% methanol before staining for cathepsin B.

### 16S sequencing and PCR assays for rodent infectious agents

DNA was isolated by Powersoil (Qiagen) per the manufacturer's protocol, and recovery yield and DNA quality were determined by fluorometric analysis. DNA concentration was standardized and amplified using

# Article

16S rRNA primers spanning the V3 and V4 regions (Illumina). Resulting amplified PCR products were analysed on a Bioanalyzer (Agilent Technologies), and then purified and amplified with primers containing unique sample nucleotide barcodes (Illumina). PCR products were analysed with the Bioanalyzer for product quality control and also by SYBR green PCR to determine the quantity. All samples were pooled and standardized to a final concentration of 4.0 nM representation for each sample. The 16S PCR product pool was denatured with sodium hydroxide then adjusted to 4.0 pM and combined with 5% PhiX control DNA before loading onto a sequencing flow cell (Illumina) with 300-bp paired ends and a unique molecular tag for each sample. Following the sequencing run, the sequence data were separated on the basis of the nucleotide bar code, and then compared to the Greengenes database<sup>38</sup>. Relative abundance,  $\alpha$ -diversity,  $\beta$ -diversity and principal coordinate analysis were performed using QIIME analysis software<sup>39</sup>. PCR assays for rodent infectious agents were performed as described<sup>40</sup>.

## PCR

Faecal DNA was isolated from faecal pellets using QIAamp Fast DNA Stool Mini Kit (Qiagen). *Helicobacter* spp. 16S rDNA was amplified using primers 5'-CTATGACGGGTATCCGCC-3' and 5'-ATTCCACCTACCTCTCCA-3'. Total Eubacteria 16S rDNA was amplified using primers 5'-TCCTACGGGAGGCAGCAG-3' and 5'-GGACTACCAGGGTATCTAATCCTGTT-3'. *T. muris* 28S rDNA was amplified using primers 5'-GCTTTGCAAGCTAGGTCCC-3' and 5'-TTTCTGATGGGGCGTACAC-3'. RNA was isolated from tissue by dissociating cortex in Trizol LS (Thermo) using a pellet pestle and reverse transcriptase with iScript (Biorad). Quantitative PCR with reverse transcription as performed using SYBR (Biorad). *Ly6c* was amplified using primers 5'-TACTGTGTGCAGAAAGAGCTCAG-3' and 5'-TTCCTTCTTGAGAGTCCTCAATC-3'. *Gapdh* was amplified using primers 5'-TGCGACTTCAACAGCAACTC-3' and 5'-GCCTCTC TTGCTCAGTGTCC-3'.

## Bone-marrow-derived macrophages

Two femurs and tibias were stripped of musculature, flushed and cultured in IMDM supplemented with 10% FCS, NEAA, glutamax, penicillin-streptomycin and 20 ng/ml mouse M-CSF (PeproTech). The medium was changed on day 3 and cells plated for experiments after 6 days. Cells were plated at  $4 \times 10^4$  cells per well of a 96-well plate, and allowed to attach overnight, followed by stimulation with microbial moieties (Invivogen), including Pam3csk4 (10–1,000 ng/ml; tlrl-pms), zymosan (1  $\mu$ g/ml; tlrl-zyn), HMW poly(I:C) (10  $\mu$ g/ml; tlrl-pic), LPS (10 ng/ml; tlrl-peklps), R848 (20 ng/ml; tlrl-r848), CpG ODN (25  $\mu$ g/ml; tlrl-1826) or PGN (20  $\mu$ g/ml; tlrl-pgnb3). For faecal stimulations, previously frozen faeces were thawed, diluted to 200 mg/ml in PBS, passed through a 40- $\mu$ m filter, quick spun, and supernatant was collected and kept on ice. Bacterial DNA was isolated from each sample using QIAamp Fast DNA Stool Mini Kit (Qiagen) and total Eubacteria 16S rDNA abundance determined by qPCR. The more concentrated sample was diluted in PBS to normalize the relative Eubacteria abundance, which was confirmed

again by bacterial DNA isolation and qPCR. Dilution curves were prepared for each normalized faecal sample and added to macrophage cultures. Penicillin-streptomycin was added to cultures after 2 h, then the medium was collected after 18 h for testing by TNF enzyme-linked immunosorbent assay at 1:2 and 1:10 dilution (BioLegend).

## Statistics

Statistical calculations were performed using GraphPad prism 8.0. Tests between two groups used a two-tailed Student's *t*-test. A Bonferroni-corrected *t*-test was used to assess differentially abundant bacterial species between pro-inflammatory and pro-survival environments. Tests between multiple groups used one-way ANOVA with either Tukey's or Sidak's multiple comparisons. Tests between multiple groups over time used two-way ANOVA with Dunnett's multiple comparisons. Survival curves were evaluated by generalized Wilcoxon test.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

The 16S rDNA sequencing dataset are available through the Gene Expression Omnibus repository at GSE147325. All other data generated or analysed are included in the published Article and its Supplementary Information.

38. DeSantis, T. Z. et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**, 5069–5072 (2006).
39. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
40. Henderson, K. S. et al. Efficacy of direct detection of pathogens in naturally infected mice by using a high-density PCR array. *J. Am. Assoc. Lab. Anim. Sci.* **52**, 763–772 (2013).

**Acknowledgements** Support to K.E. was provided by The Merkin Fund at the Broad Institute, Target ALS, NIH 5R01NS089742, Harvard Stem Cell Institute and UCB. A.B. was supported by NIH 5K99AG057808-02, M.F.W. was supported by NIH 1K99MH119327-01. We thank J. Wang for providing faeces from mice housed Johns Hopkins University.

**Author contributions** A.B., M.F.W., J.M. and K.E. conceived the study. Experiments were performed by A.B. (Figs. 1–4, Extended Data Figs. 1–9), F.L. (Fig 1, Extended Data Fig. 2), M.F.W. (Fig. 1, Extended Data Fig. 2), K.S.S. (Fig. 4, Extended Data Figs. 8, 9), A.C. (Figs. 2–4, Extended Data Figs. 3–9), J.M. (Figs. 1–3, Extended Data Figs. 2–6, 8), N.v.G. (Extended Data Fig. 2), J.-Y.W. (Fig. 4, Extended Data Fig. 8), J.K. and G.G. (Extended Data Fig. 8), and M.Q. and P.E. and C.C. (Figs. 2, 3, Extended Data Figs. 3, 4, 8), A.B., F.L., K.S.S., J.K., G.G., N.v.G., J.-Y.W., O.P., I.K., D.T.S. and K.E. interpreted results. A.B. and K.E. wrote the manuscript.

**Competing interests** K.E. is a co-founder of Q-State Biosciences, Quralis and Enclear Therapies.

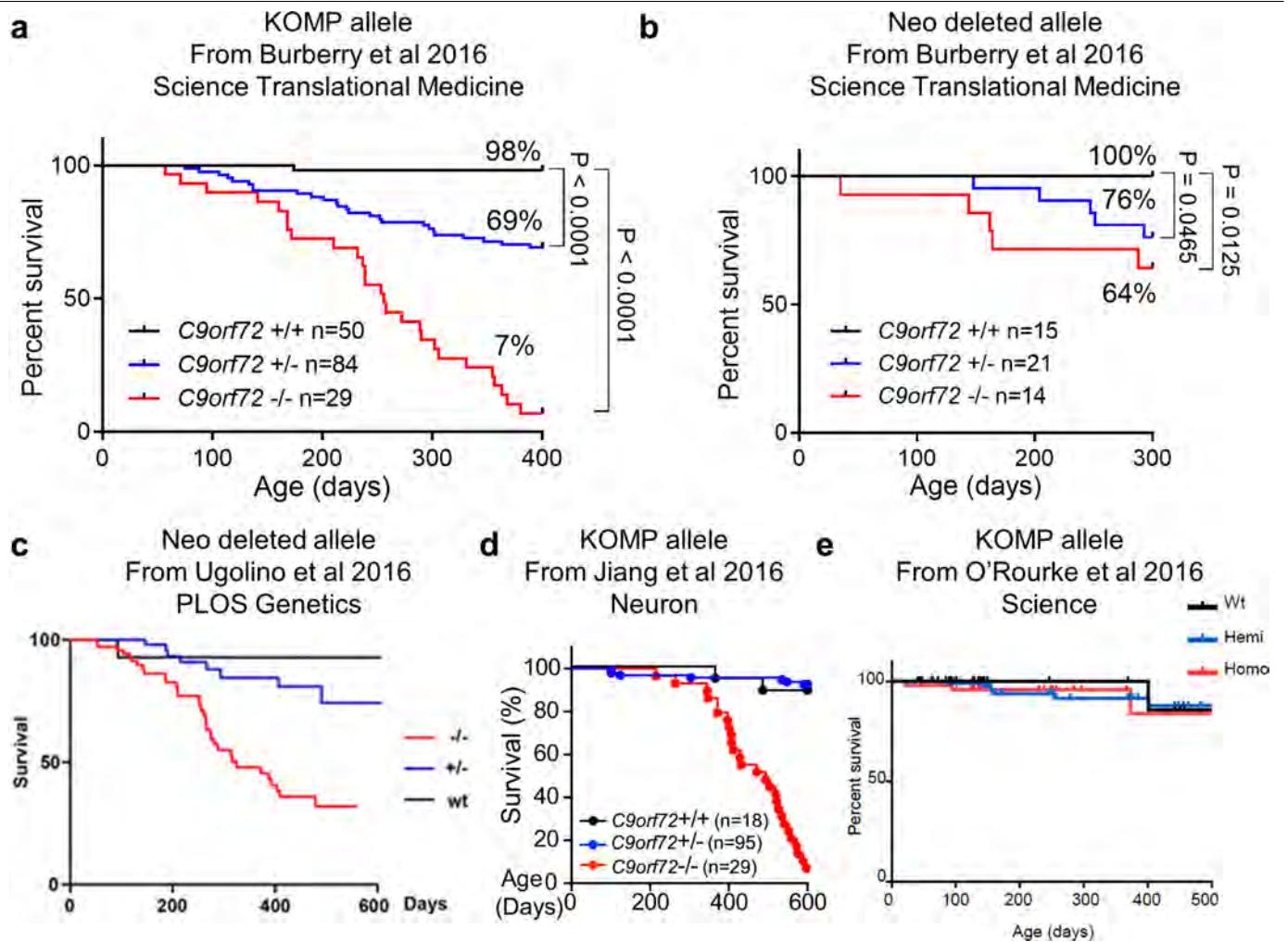
## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2288-7>.

**Correspondence and requests for materials** should be addressed to K.E.

**Peer review information** *Nature* thanks Michael Fischbach, Aaron D. Gitler and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

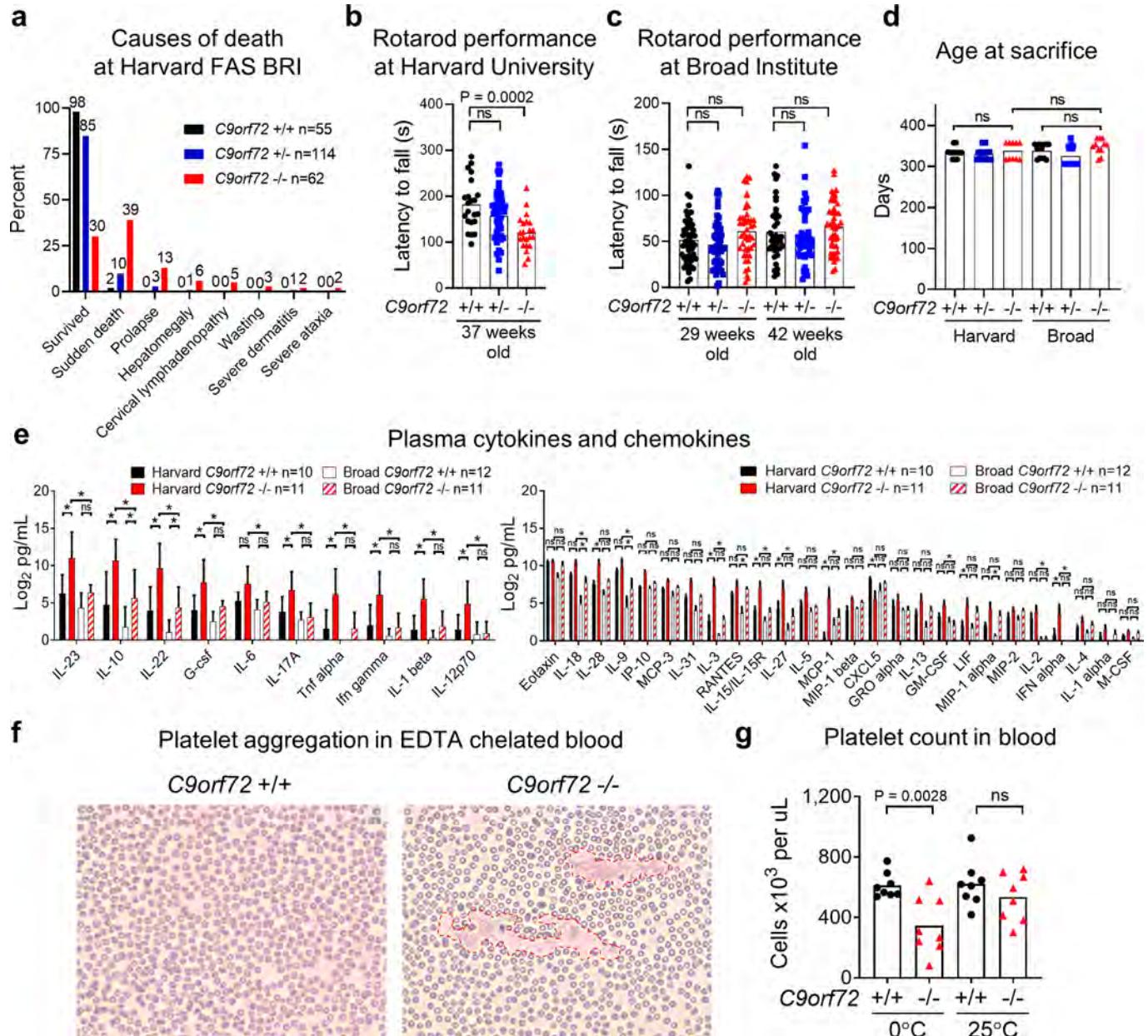


**Extended Data Fig. 1 | Previously published *C9orf72* LOF survival studies.**

**a**, From ref.<sup>7</sup>. Reproduced with permission from AAAS. **b**, From ref.<sup>7</sup>. Reproduced with permission from AAAS. **c**, Reproduced from ref.<sup>12</sup>, licensed

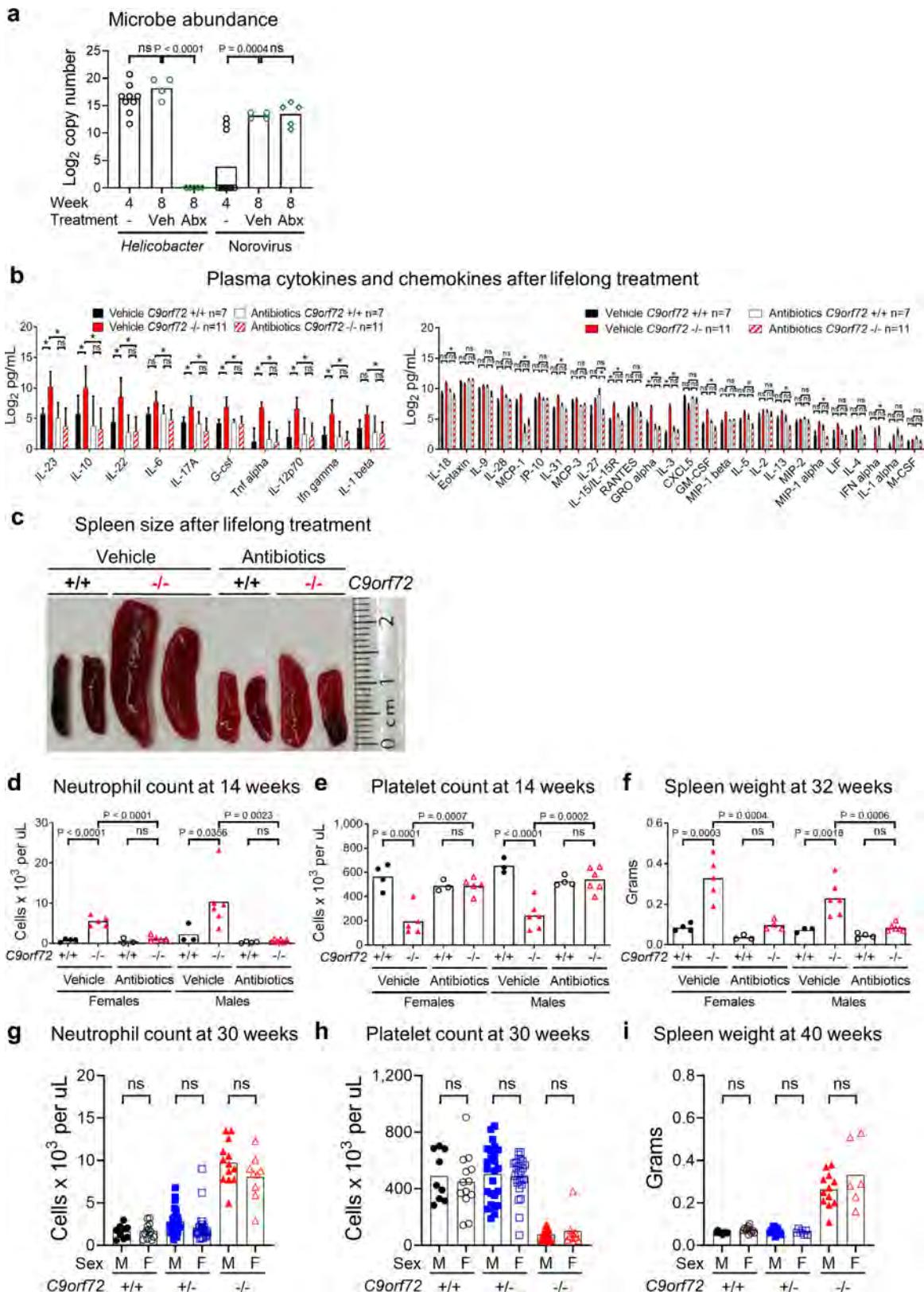
under a CC BY 4.0 licence. **d**, Reprinted from ref.<sup>13</sup>, with permission from Elsevier. **e**, From ref.<sup>6</sup>. Reproduced with permission from AAAS.

# Article



**Extended Data Fig. 2 | Causes of death, motor performance, levels of plasma cytokines and identification of pseudothrombocytopenia in *C9orf72*LOF mice.** **a**, Causes of death or premature mortality of *C9orf72*(Harvard) mice in Fig. 1b. **b**, Accelerating rotarod performance of 37-week-old *C9orf72*(Harvard) neo-deleted mice (*C9orf72*(Harvard)<sup>+/+</sup>, *n* = 22; *C9orf72*(Harvard)<sup>+/−</sup>, *n* = 50; *C9orf72*(Harvard)<sup>−/−</sup>, *n* = 22). **c**, Accelerating rotarod performance of *C9orf72*(Broad) neo-deleted mice at 29 weeks of age (*C9orf72*(Broad)<sup>+/+</sup>, *n* = 53; *C9orf72*(Broad)<sup>+/−</sup>, *n* = 52; *C9orf72*(Broad)<sup>−/−</sup>, *n* = 48) or 42 weeks of age (*C9orf72*(Broad)<sup>+/+</sup>, *n* = 38; *C9orf72*(Broad)<sup>+/−</sup>, *n* = 48; *C9orf72*(Broad)<sup>−/−</sup>, *n* = 48). In **b**, **c**, one-way ANOVA with Dunnett's multiple comparisons. Each point represents the average of three trials per mouse.

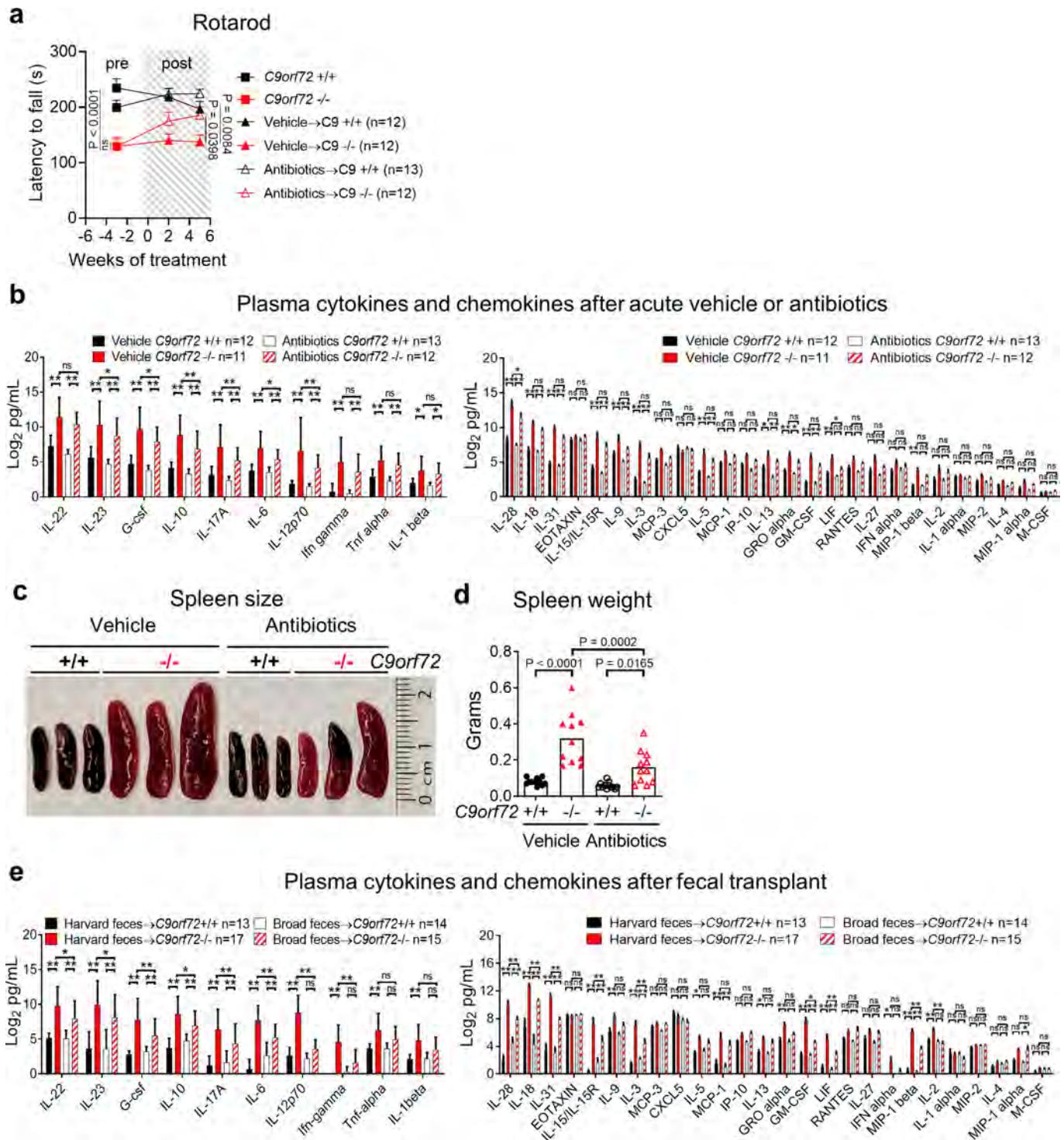
**d**, Age at which mice in Fig. 1d–g were killed. One-way ANOVA with Sidak's multiple comparisons. **e**, Plasma cytokines and chemokines at death from mice in Fig. 1d–g. Mean ± s.d. Two-way ANOVA with Tukey's multiple comparisons. **f**, Peripheral blood smear of 18-week-old *C9orf72*(Harvard) neo-deleted mice. Platelets from *C9orf72*(Harvard)<sup>−/−</sup> mice were prone to aggregate (outlined by red dashed lines) in the presence of EDTA at 0 °C. **g**, Pseudothrombocytopenia could be reversed by warming the blood to room temperature. The reduced platelet count in the *C9orf72*(Harvard)<sup>−/−</sup> model therefore represents an indirect measure of anti-platelet auto-antibodies, rather than a reduction in platelet abundance. Two-way ANOVA with Tukey's multiple comparisons. Each dot represents one mouse.



**Extended Data Fig. 3 | Cytokines and chemokines in lifelong-antibiotic-treated *C9orf72*LOF mice and sex stratification of inflammatory phenotypes.** **a**, PCR analysis of *Helicobacter* spp. and norovirus DNA in faecal pellets. Each dot represents faeces from one cage. One-way ANOVA with Dunnett's multiple comparisons. **b**, Plasma cytokines and chemokines of mice in Fig. 2. Mean  $\pm$  s.d. Two-way ANOVA with Tukey's multiple comparisons. **c**, Representative spleen size of mice in Fig. 2. **d-f**, Total blood neutrophil count (**d**), platelet count (**e**) and spleen weight (**f**) from mice in Fig. 2.

stratified by sex. **g, h**, Total blood neutrophil count (**g**) and platelet count (**h**) in 30-week-old *C9orf72(Harvard)* neo-deleted mice stratified by sex (*C9orf72(Harvard)*<sup>+/+</sup>,  $n=9$  male and 13 female; *C9orf72(Harvard)*<sup>-/-</sup>,  $n=25$  male and 27 female; *C9orf72(Harvard)*<sup>-/+</sup>,  $n=13$  male and 9 female). **i**, Spleen weight in 40-week-old *C9orf72(Harvard)* neo-deleted mice stratified by sex (*C9orf72(Harvard)*<sup>+/+</sup>,  $n=8$  male and 11 female; *C9orf72(Harvard)*<sup>-/-</sup>,  $n=13$  male and 7 female; *C9orf72(Harvard)*<sup>-/+</sup>,  $n=12$  male and 6 female). In **d–i**, each dot represents one mouse. One-way ANOVA with Sidak's multiple comparisons.

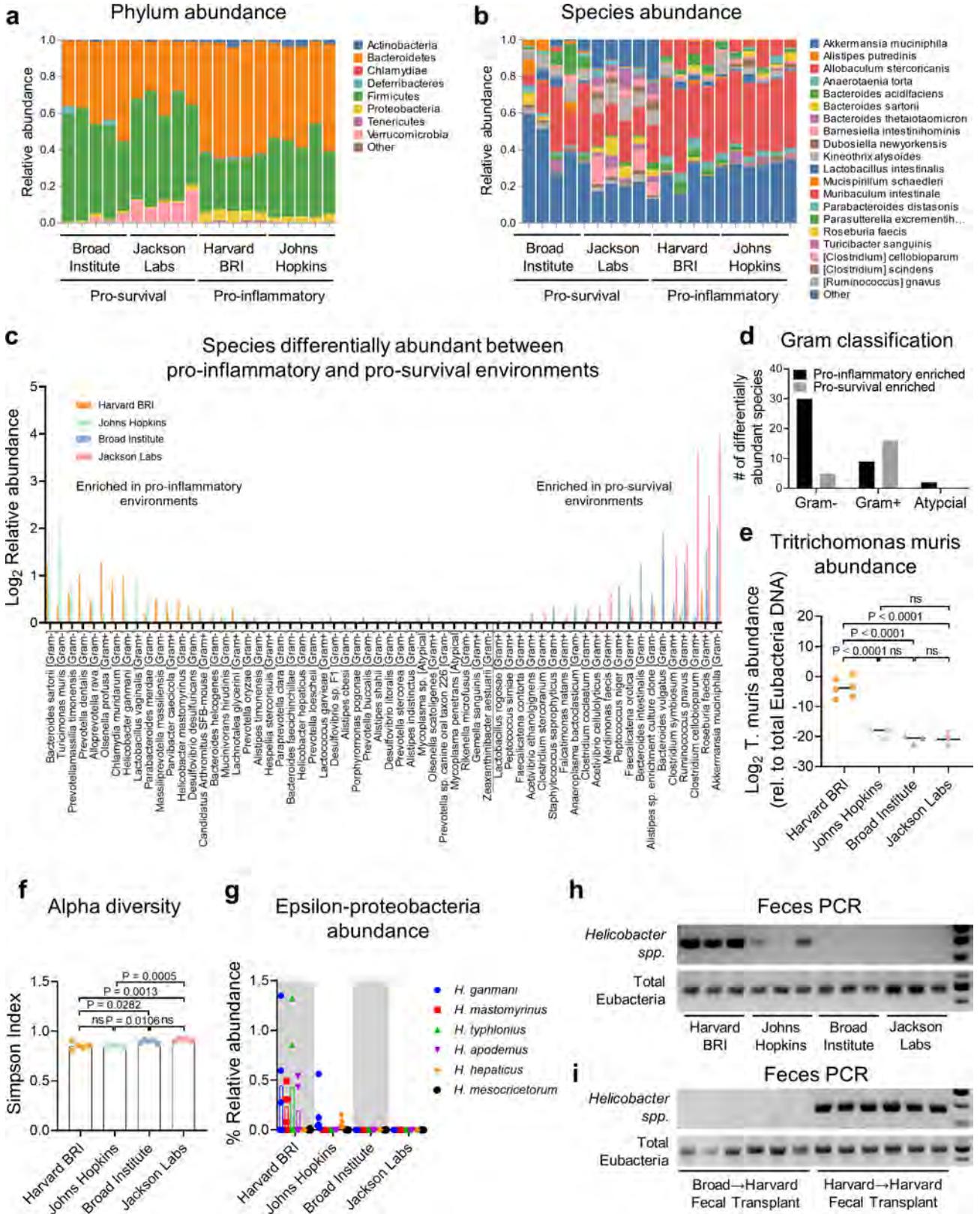
# Article



**Extended Data Fig. 4 | Acute antibiotic treatment improves motor function, and mitigates splenomegaly and cytokine burden, in C9orf72 LOF mice.** **a**, Accelerating rotarod performance of mice in Fig. 3a. Each point represents the average of three trials per mouse. Two-way ANOVA with Dunnett's multiple comparisons. **b**, Plasma cytokines and chemokines of mice in Fig. 3a–d after seven weeks of treatment. **c, d**, Representative spleen size (**c**)

and spleen weight (**d**) of mice in Fig. 3a after eight weeks of treatment. Each dot represents one mouse. One-way ANOVA with Sidak's multiple comparisons.

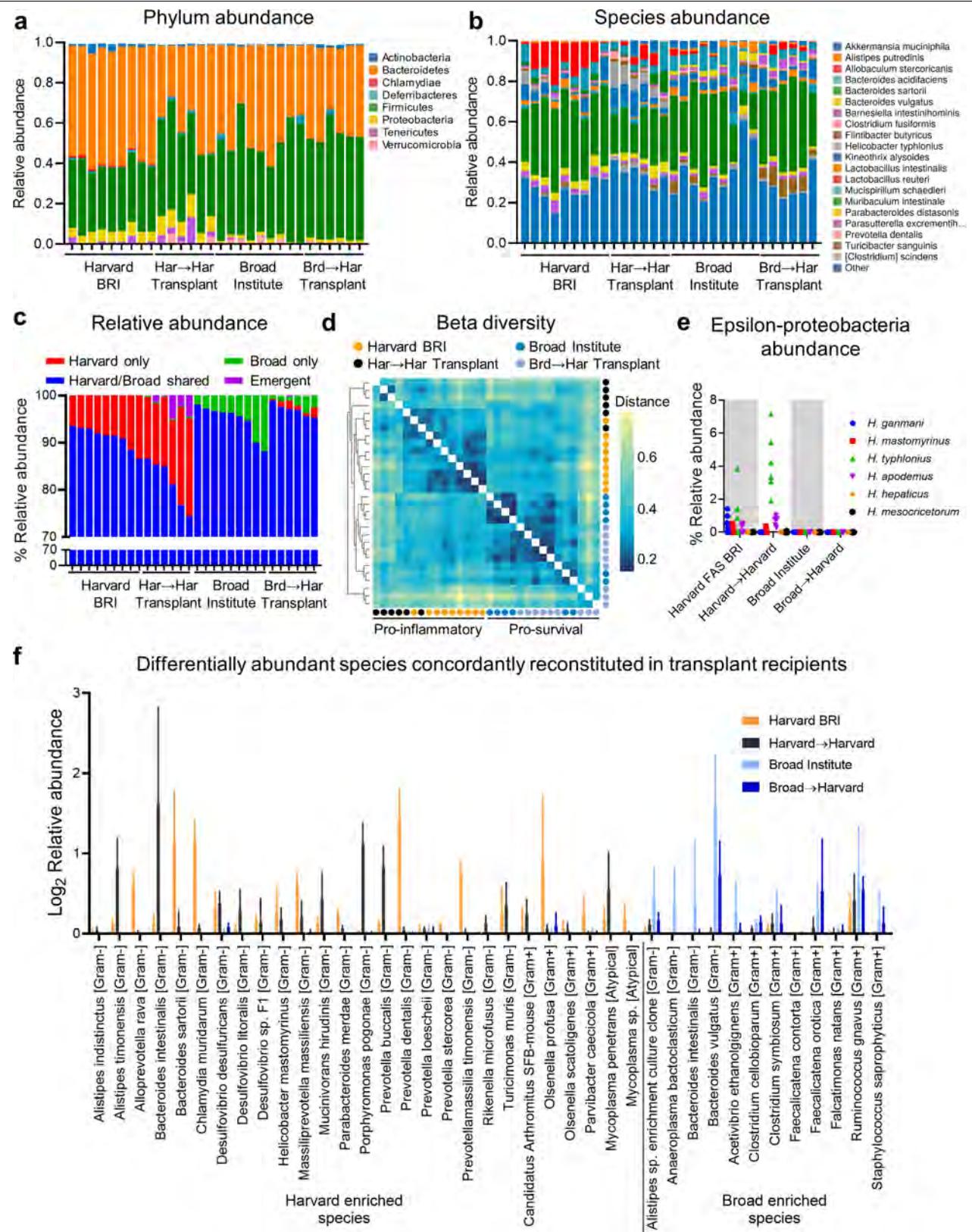
**e**, Plasma cytokines and chemokines of mice in Fig. 3e–h 10 weeks after faecal transplant. In **b, e**, mean  $\pm$  s.d. Two-way ANOVA with Tukey's multiple comparisons.



Extended Data Fig. 5 | See next page for caption.

# Article

**Extended Data Fig. 5 | Bacteria and protozoa diversity across environments.** **a, b**, Phylum-level (**a**) and species-level (**b**) relative abundance of bacteria from 16S rDNA sequencing in Fig. 3i. Each bar represents sequencing from one pellet per cage. **c, d**, Relative abundance (**c**) and Gram-stain classification (**d**) of bacterial species, the abundance of which was significantly different between pro-inflammatory environments (Harvard BRI and Johns Hopkins University) and pro-survival environments (Broad Institute and Jackson Laboratory). *t*-test with Bonferroni multiple comparisons; 62/301 detected species had significance  $P < 0.0002$ .  $n = 5$  faecal pellets per environment. Mean  $\pm$  s.d. **e**, Quantitative PCR with reverse transcription analysis of *T. muris* 28S rDNA relative to total Eubacteria 16S rDNA in faeces. **f**, Simpson index of faecal  $\alpha$ -diversity. **g**, Relative abundance of  $\epsilon$  proteobacteria (*Helicobacter*). In **e–g**, each dot represents a faecal pellet from one cage. One-way ANOVA with Tukey's multiple comparisons. **h**, PCR analysis of *Helicobacter* spp. 16S rDNA and total Eubacteria 16S rDNA in faeces (six weeks after transplant) from Fig. 3e. **i**, PCR analysis of *Helicobacter* spp. 16S rDNA and total Eubacteria 16S rDNA in faeces (six weeks after transplant) from Fig. 3e.

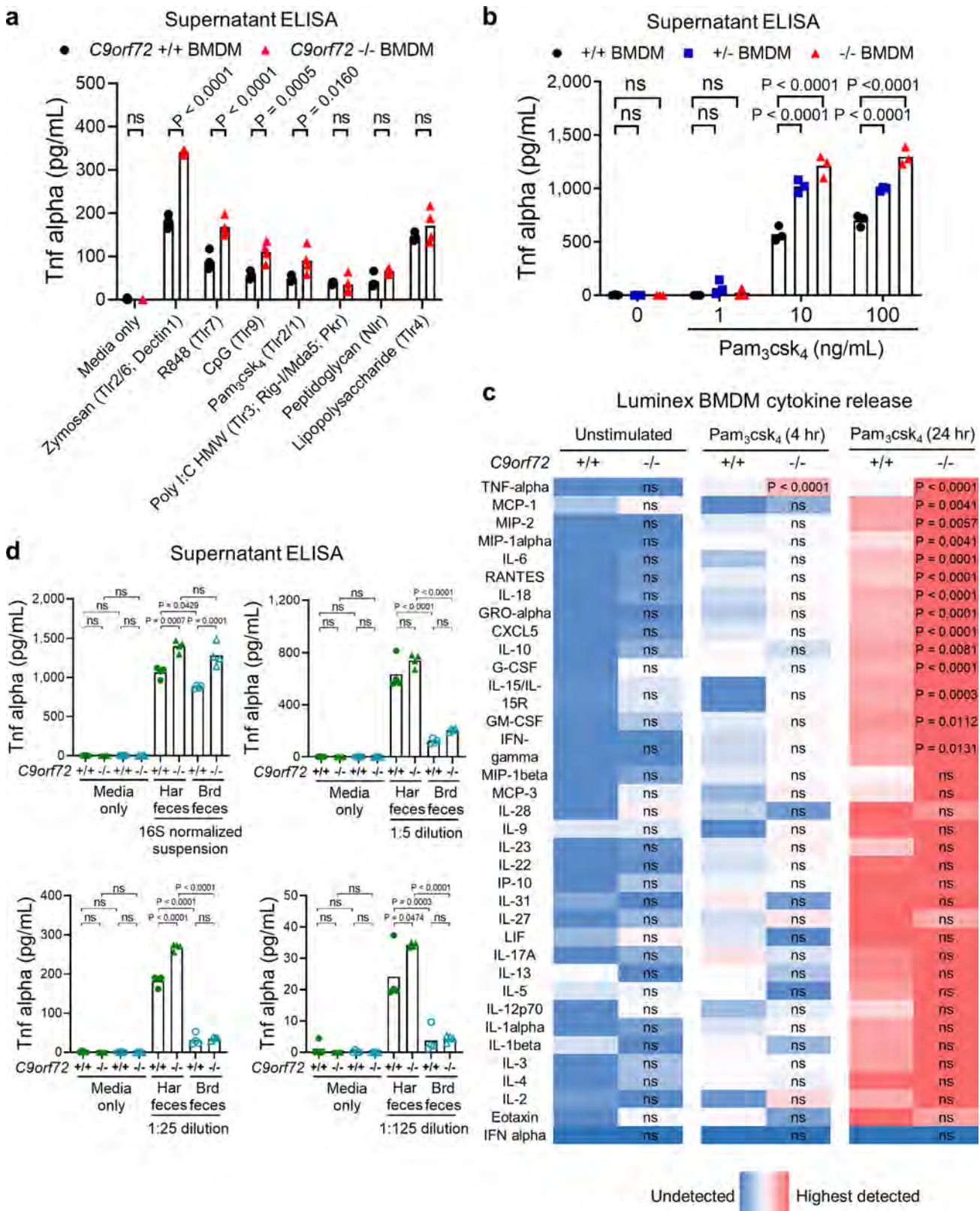


**Extended Data Fig. 6** | See next page for caption.

## Article

**Extended Data Fig. 6 | Environment-enriched bacteria engraft faecal-transplant recipients.** **a–f**, Analysis of bacteria in faeces at 10 weeks after transplant (from mice in Fig. 3e) by 16S rDNA sequencing. Each bar represents a faecal sample from an individual cage. **a, b**, Phylum-level (**a**) and species-level (**b**) relative abundance. **c**, Relative abundance of bacterial species grouped as those only observed in cages from Harvard BRI (Harvard-only), those only observed in cages from the Broad Institute (Broad-only), those observed in cages from Harvard BRI and the Broad Institute (Harvard/Broad-shared) or those not observed in Harvard BRI or Broad Institute cages

but detectable in transplant recipient cages (emergent). **d**, Bray–Curtis dissimilarity matrix of faeces  $\beta$ -diversity. **e**, Relative abundance of  $\varepsilon$  proteobacteria (*Helicobacter*). **f**, Putative pro-inflammatory species ( $n=27$ ) enriched in pro-inflammatory environments (Harvard BRI and Johns Hopkins University) that were also enriched in Harvard-to-Harvard recipients, and putative pro-survival species ( $n=12$ ) enriched in pro-survival environments (Broad Institute and Jackson Laboratory) and enriched in Broad-to-Harvard recipients.

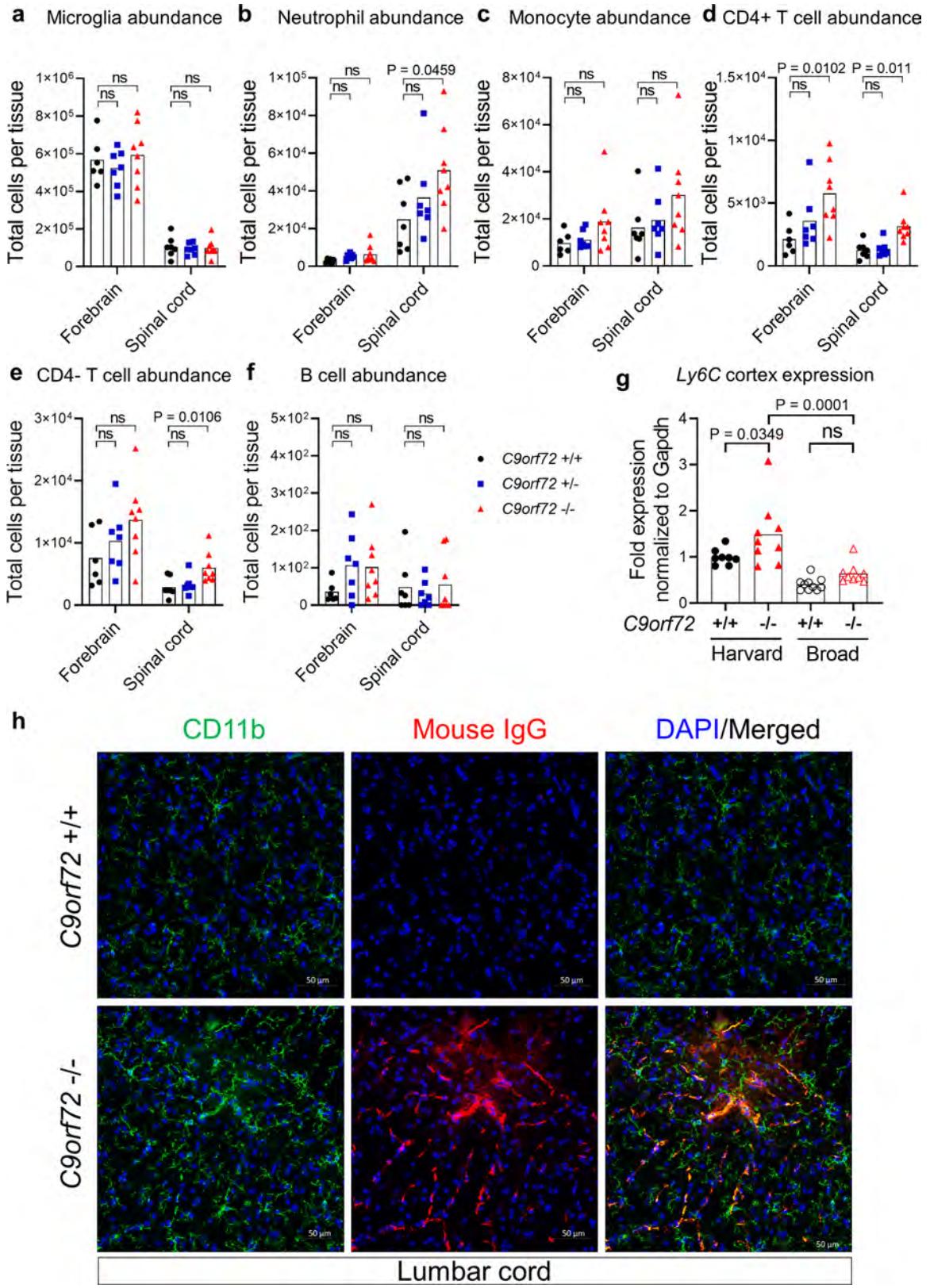


**Extended Data Fig. 7** | See next page for caption.

## Article

**Extended Data Fig. 7 | *C9orf72* restricts myeloid cytokine release in response to foreign stimuli.** **a–d,** Analysis of cytokines and chemokines in supernatant 24 h after stimulation of bone-marrow-derived macrophages (BMDM) with activators of Toll-like receptor (Tlr) or NOD-like receptor (Nlr) agonists (**a–c**) or filtered Eubacteria-normalized faecal preparations (**d**). **c,** The abundance of cytokine and chemokine in the supernatant was normalized and colour-coded (blue, low; red, high) relative to the average level of each molecule in unstimulated *C9orf72*<sup>+/+</sup> bone-marrow-derived macrophage wells. Levels of each analyte were measured by Luminex in multiplex. **d,** The abundance of total Eubacteria in each faecal sample was measured by qPCR for

16S rDNA and this value was used to normalize faecal Eubacteria bacteria concentration before generation of the dilution curve. Each dot represents one well. Panels are representative of  $n=2$  replicate experiments (**a**);  $n=5$  replicate experiments (**b**); one representative experiment with average of  $n=3$  technical replicates per condition (**c**);  $n=2$  replicate experiments (**d**). In **a**, two-way ANOVA with Sidak's multiple comparison. In **b**, two-way ANOVA with Dunnett's multiple comparisons. In **c**, two-way ANOVA with Sidak's multiple comparison for each analyte tested. In **d**, one-way ANOVA with Sidak's multiple comparisons.



**Extended Data Fig. 8** | See next page for caption.

## Article

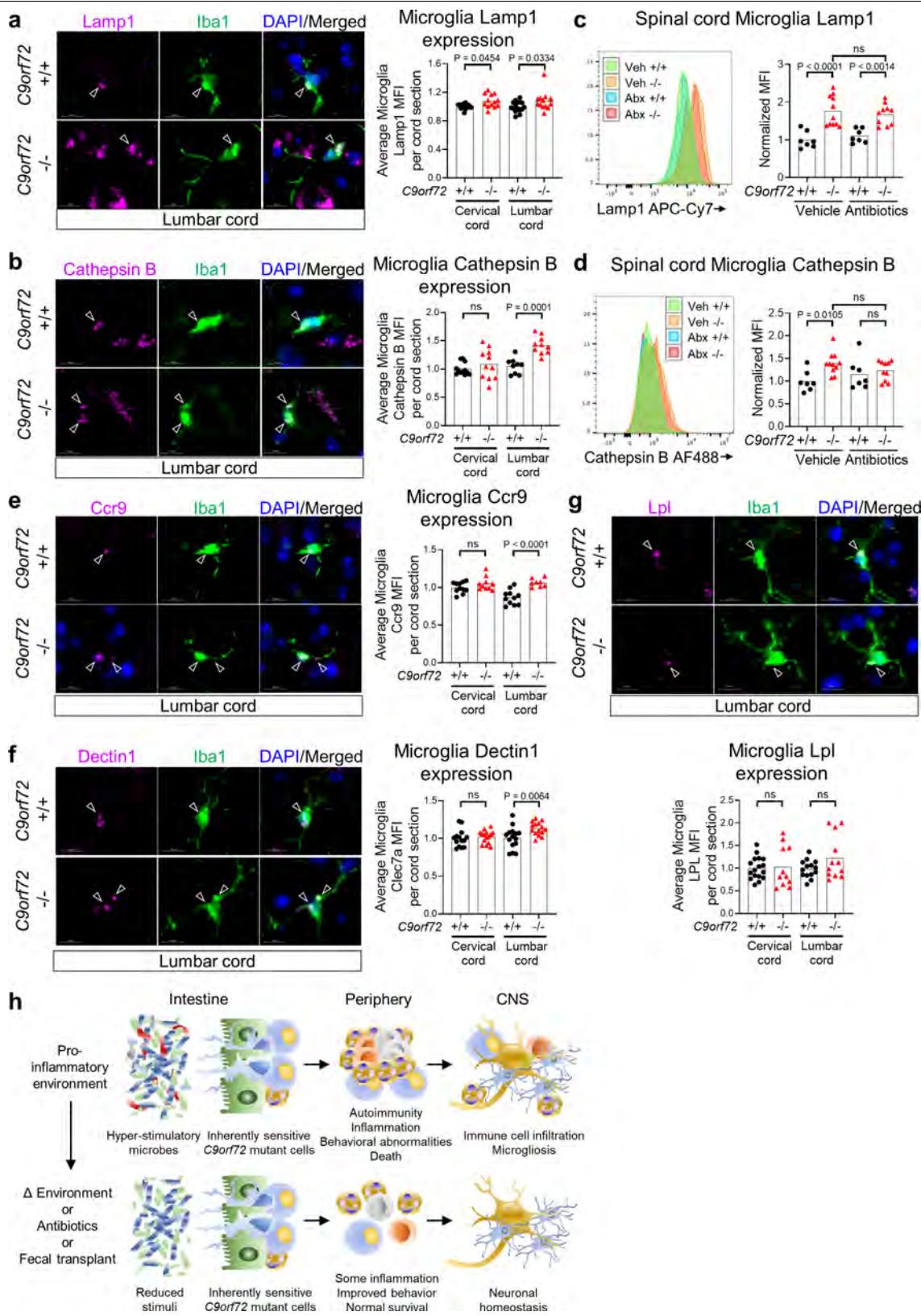
### Extended Data Fig. 8 | Neutrophils and T cells infiltrate spinal cord of

**C9orf72 LOF mice.** **a–f,** Mass cytometry investigation of single-cell-dissociated forebrain or spinal cord from 36-week-old *C9orf72(Harvard)* neo-deleted male and female mice (*C9orf72(Harvard)*<sup>+/+</sup>,  $n=7$ ; *C9orf72(Harvard)*<sup>-/-</sup>,  $n=7$ ; *C9orf72(Harvard)*<sup>-/-</sup>,  $n=8$ ). One *C9orf72*<sup>+/+</sup> forebrain sample failed, and was excluded from analysis. Representative gating scheme can be found in Supplementary Information. Populations were defined as CD45<sup>mid</sup> CX3CR1<sup>+</sup>CD39<sup>+</sup> microglia (**a**), CD45<sup>high</sup>Ly6C<sup>+</sup>Ly6G<sup>high</sup> neutrophils (**b**), CD45<sup>high</sup>Ly6C<sup>+</sup>Ly6G<sup>low</sup> monocytes (**c**), CD45<sup>high</sup>CD3e<sup>+</sup>CD4<sup>+</sup> T cells (**d**), CD45<sup>high</sup>CD3e<sup>+</sup>CD4<sup>-</sup> T cells (**e**) and CD45<sup>high</sup>CD19<sup>+</sup> B cells (**f**). Quantification of total cells per tissue was obtained by multiplying the percentage of each gated

population by the total cells recovered from the tissue of that mouse. Each dot represents one mouse. Two-way ANOVA with Dunnett's multiple comparisons.

**g,** Quantitative PCR with reverse transcription of *Ly6c* expression in total cortex tissue of 47-week-old *C9orf72(Harvard)* neo-deleted mice (*C9orf72(Harvard)*<sup>+/+</sup>,  $n=8$ ; *C9orf72(Harvard)*<sup>-/-</sup>,  $n=9$ ) or *C9orf72(Broad)* neo-deleted mice (*C9orf72(Broad)*<sup>+/+</sup>,  $n=10$ ; *C9orf72(Broad)*<sup>-/-</sup>,  $n=9$ ). Each dot represents one mouse. One-way ANOVA with Sidak's multiple comparisons.

**h,** Orthogonal projection of confocal imaging of CD11b and mouse immunoglobulin IgG in lumbar spinal cord of a 43-week-old *C9orf72(Harvard)* mouse.



**Extended Data Fig. 9** | See next page for caption.

## Article

**Extended Data Fig. 9 | Elevated lysosomal proteins and microgliosis in spinal cords of *C9orf72*LOF mice.** **a, b, e–g,** Orthogonal projection and quantification of confocal imaging of LAMP1 (**a**), cathepsin B (**b**), CCR9 (**e**), dectin 1 (CLEC7A) (**f**) and LPL (**g**) in IBA1<sup>+</sup> microglia in spinal cord of a 55-week-old *C9orf72(Harvard)* mouse. One-way ANOVA with Sidak's multiple comparisons. Each dot represents the average mean fluorescent intensity (MFI) of the antigen within microglia on a given spinal cord section. Over 100 microglia were surveyed per section. Sections from  $n=3$  *C9orf72*<sup>+/+</sup> and  $n=3$  *C9orf72*<sup>-/-</sup> mice were surveyed. **c, d,** Flow cytometry quantification of

LAMP1 (**c**) or cathepsin B (**d**) in CD45<sup>mid</sup>CD11b<sup>+</sup>CD39<sup>+</sup> microglia from spinal cord of *C9orf72(Harvard)* neo-deleted mice in Fig. 2. One-way ANOVA with Sidak's multiple comparisons. **h,** Graphical illustration of *C9orf72* functioning within the haematopoietic system to restrict the development of inflammation, autoimmunity, peripheral immune infiltration into the central nervous system (CNS) and microgliosis in response to hyper-stimulatory communities of gut microflora. The microglia image was modified from Servier Medical Art ([https://smart.servier.com/smart\\_image/microglia-2/](https://smart.servier.com/smart_image/microglia-2/)) under a CC BY 3.0 licence.

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

Flow cytometry and mass cytometry data was analyzed by Flowjo v10.6 or Cytobank Premium.

Data analysis

Image analysis was performed using ZEISS ZEN 2.6 image processing software. Videos were rendered using Bitplane Imaris 9.2. For 16S rDNA sequencing following the sequencing run, the sequence data was separated based on the nucleotide bar code and then compared to the Greengenes database. Relative abundance, alpha diversity, beta diversity and principal coordinate analysis was performed using QIIME analysis software (v2).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The 16S rDNA sequencing datasets generated and analyzed during the current study are available through the GEO repository at GSE147325. All other data generated or analyzed during this study are included in this published article (and its supplementary information files).

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](http://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Effect sizes of all phenotypic measures were estimated using GPower 3.1 software based on the average and standard deviation of each measure either from small trial experiments or previously published work on the model. These effect sizes were used to estimate necessary sample sizes for larger studies involving multiple treatment groups.
Data exclusions	Data were not excluded and values from each animal or replicate are indicated on the graphs as stated in figure legends
Replication	All attempts at replication were successful. To ensure reproducibility, independent cohorts of animals were repeatedly generated and tested for the presence of phenotypes described. In the case where survival differed when mice were reared in different environments, mechanistic experiments were performed to understand what environmental factors contributed to differences observed.
Randomization	In Figures 1, 2 and 4, age matched animals born from heterozygous inter-crosses were assigned to each cohort for study. In Figure 3, mutant animals were pre-screened for behavioral performance, neutrophilia and psuedothrombocytopenia and animals demonstrating these phenotypes were licensed into the trial as described in the manuscript text.
Blinding	Researchers performing behavioral tasks, bleeding animals, and administering antibiotics or vehicle treatment or feces were blinded to genotype.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

### Antibodies used

Immunofluorescence: Rat-anti-mouse/human CD11b-FITC 1:200 (M1/70, BioLegend 101205), rabbit-anti-mouse/human Cathepsin B 1:400 (D1C7Y CST, 31718S), rat-anti-mouse Ly6C-647 1:1000 (HK1.4 BioLegend, 128009), rat-anti-mouse CD45-488 1:200 (30-F11 BioLegend, 103121), guinea pig-anti-mouse Iba1 1:500 (234004 Synaptic Systems), rat-anti-mouse/human Lamp1 1:200 (1D4B SCB, sc-19992), rat-anti-mouse Dectin1/Clec7a 1:30 (Invivogen, mabg-mdect), mouse-anti-mouse-Lpl LPL.A4 1:200 (Abcam, ab21356). Donkey-anti-rat-AlexaFluor-488 (Invitrogen, A-21208), Donkey-anti-mouse IgG-555 (Invitrogen, A-31570), Goat-anti-rabbit-555 (Invitrogen, A-21428), Goat-anti-rabbit-647 (Invitrogen, A27040), Goat-anti-rat-647 (Invitrogen, A-21247), Goat-anti-guinea pig-647 (Invitrogen, A-21450), all 1:500 dilution.  
Flow cytometry: (BioLegend): rat-anti-mouse CD45-BV421 (103103) or APC-Cy7 (103115) 1:200 (30-F11), rabbit-anti-mouse/human Cathepsin B 1:100 FACS (D1C7Y CST, 31718S) and donkey anti-rabbit-AlexaFluor-488 1:500 (Invitrogen, A-21206), rat-anti-mouse-Ccr9-FITC 1:200 (9B1, Biolegend, 129705), rat-anti-mouse F4/80-PE-Cy5 1:400 (BM8, Biolegend, 123111), rat-anti-mouse/human CD11b-AlexaFluor-700 1:400 (M1/70, Invitrogen, 56-0112-82), rat-anti-mouse Lamp1-APC-Cy7 1:400 (1D4B, Biolegend, 121615), TruStain FCX rat-anti-mouse CD16/32 1:250 (93, Biolegend, 101319), rat-anti-mouse CD39-PE 1:400 (Duoh59, Biolegend, 143803), rat-anti-mouse Ly6G-PE-Cy7 1:600 (1A8, Biolegend, 127617), rat-anti-mouse Ly6C-AlexaFluor-647 (HK1.4, Biolegend, 128009).  
Mass cytometry: (All Fluidigm 1:150) Rat-anti-mouse Ly6G-141Pr 1A8 (3141008B), Rat-anti-mouse CD39-142Nd 24DMS1 (3142005B), rat-anti-mouse CD4-145Nd RM4-5 (3145002B), Rat-anti-mouse CD45-147Sm 30-F11 (3147003B), Rat-anti-mouse/human CD11b-148Nd M1/70 (3148003B), Rat-anti-mouse CD3e-152Sm 145-2C11 (3152004B), Rat-anti-mouse CD19-166Er 6D5 (3166015B), rat-anti-mouse CX3CR1-164Dy SA011F11 (3164023B). Rat-anti-mouse Ly6C-143Nd HK1.4 1:150 (Biolegend, 128039).

## Validation

Description of antibody validations are provided by each commercial manufacturer on their associated website. Each lot of Biolegend antibodies were quality control tested by immunofluorescent staining with flow cytometric analysis. The Cell Signaling Technology Cathepsin B (D1C7Y) XP antibody met all of the quality control standards and was approved for Flow Cytometry, Immunofluorescence, Immunohistochemistry and Western. The SCB Lamp1 1D4B antibody was recommended for detection of LAMP1 of mouse, rat and human origin by WB, IP, IF and FCM. The Synaptic Systems guinea pig anti-mouse Iba1 was approved for WB, IP, ICC, and IHC and reacts with mouse, human, rat, sheep and ape with no signal in zebrafish. The Invivogen rat anti-mouse Dectin1 mabg-mdect antibody was screened for flow cytometry and neutralization activity and recognizes both mDectin-1 isoforms A and B. The Abcam mouse anti-mouse-Lpl LPLA4 antibody was suitable for ICC/IF, IHC-P, Flow Cyt, WB and ELISA with reactivity against mouse, human and cow. The Invitrogen CD11b-AF700 M1/70 antibody was published to react with mouse, human and fish. For Invitrogen secondary antibodies, they show minimum cross-reactivity to species other than the intended target and are approved for Flow, ICC, IF and WB. Fluidigm 600 Maxpar metal conjugated antibodies were predesigned for mass cytometry. Each antibody was validated empirically by performing dilution series on tissue or cells expressing the relevant antigens in relation to isotype or buffer controls

## Animals and other organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research

## Laboratory animals

C9orf72 loss of function neo deleted animals were previously generated using the Knockout Mouse Project targeting vector on C57BL6/J background and crossed with mice expressing Sox2-cre to remove the neomycin resistance cassette. Male and Female mice were studied in Figure 1 (Harvard BRI C9orf72 +/+ n=12; +/- n=13; -/- n=10 with mean age 47.7 weeks s.d. 2.3 weeks; Broad Institute C9orf72 +/+ n=12; +/- n=18; -/- n=11 with mean age 47.9 weeks s.d. 3.2 weeks) and Figure 2 (vehicle +/+ n=7; -/- n=11 mean age 3.3 weeks s.d. 0.8 weeks or antibiotics +/+ n=7; -/- n=11 mean age 3.4 weeks s.d. 1.2 weeks) while only female mice were studied in Figure 3a (vehicle +/+ n=12; -/- n=12 mean age 35.3 s.d. 1.9 weeks or antibiotics +/+ n=13; -/- n=12 mean age 35.8 weeks s.d. 1.4 weeks) and 3e (Harvard BRI feces +/+ n=13; -/- n=17 mean age 13.3 weeks s.d. 3.8 weeks or Broad Institute feces +/+ n=14; -/- n=15 mean age 13.3 weeks s.d. 3.8 weeks). For confocal imaging in Figure 4 n=3 +/+ and n=3 -/- male 55 week old mice used.

## Wild animals

The study did not involve wild animals.

## Field-collected samples

The study did not involve samples collected from the field.

## Ethics oversight

All experimental procedures were approved by the Institutional Animal Care and Use Committee of Harvard University and the Broad Institute.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Flow Cytometry

## Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

## Methodology

## Sample preparation

Animals were anesthetized with isoflurane followed by transcardial perfusion with HBSS supplemented with 10 U/mL heparin. Spleens were dissociated by repeated trituration with glass pipettman in HBSS, subjected to 10-minute RBC lysis (eBioscience), washed in autoMACS (Miltenyi), filtered (40 µm) and counted using Countess (Invitrogen) for antibody staining. For flow cytometry of the CNS, spinal cords were digested by papain and DNase diluted in EBSS (Worthington) for 10 minutes at 37°C, triturated with glass pipettman to generate large tissue chunks then allowed to digest for 20 minutes at 37°C. DMEM supplemented with glutamax was added, samples triturated to single cells, ovomucoid (Worthington) and DNase diluted in EBSS added to inhibit protease activity, cells filtered, washed in autoMACS buffer, and pelleted at 500xg for 15 minutes at 4°C. Cell pellets brought up in isotonic Percoll Plus (Sigma) diluted to 30% in autoMACS and spun for 15 minutes at room temperature with no brake. Floating myelin was gently removed using plastic transfer pipette. Cell pellets were resuspended, filtered, washed in autoMACS and re-pelleted at 4°C. Cells were fixed in 4% paraformaldehyde (PFA; Electron Microscopy Sciences) either before or after antibody staining depending on need.

## Instrument

Samples collected on BD LSRII

## Software

Flow cytometry and mass cytometry data was analyzed by Flowjo v10.6 or Cytobank Premium.

## Cell population abundance

No experiments involving cell sorting are described

## Gating strategy

Cells were defined based on Fsc-A and Ssc-A profile, then single cells were gated based on horizontal ratio of Fsc-A/Fsc-W and then Ssc-A/Ssc-W. CD45+ cells were sub-gated based on CD45/Ssc-A using samples without anti-CD45 antibody to define the

negative population. From the CD45+ gate, CD45/CD11b was used to define CD45mid CD11b+ microglia and CD45hi CD11b+ infiltrating myeloid cells. From the CD45+ gate, CD45mid CD11b+ microglia were confirmed to be CD39+, while CD45hi CD11b+ infiltrating myeloid cells were confirmed to be Ly6C+.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

## Article

# Landscape and function of multiple mutations within individual oncogenes

<https://doi.org/10.1038/s41586-020-2175-2>

Received: 23 July 2019

Accepted: 13 February 2020

Published online: 08 April 2020

 Check for updates

Yuki Saito<sup>1,2,10</sup>, Junji Koya<sup>1,10</sup>, Mitsugu Araki<sup>3</sup>, Yasunori Kogure<sup>1</sup>, Sumito Shingaki<sup>1</sup>, Mariko Tabata<sup>1,4</sup>, Marni B. McClure<sup>1</sup>, Kota Yoshifiji<sup>1,5</sup>, Shigeyuki Matsumoto<sup>6</sup>, Yuta Isaka<sup>7</sup>, Hiroko Tanaka<sup>8</sup>, Takanori Kanai<sup>2</sup>, Satoru Miyano<sup>8</sup>, Yuichi Shiraishi<sup>9</sup>, Yasushi Okuno<sup>3</sup> & Keisuke Kataoka<sup>1,✉</sup>

Sporadic reports have described cancer cases in which multiple driver mutations (MMs) occur in the same oncogene<sup>1,2</sup>. However, the overall landscape and relevance of MMs remain elusive. Here we carried out a pan-cancer analysis of 60,954 cancer samples, and identified 14 pan-cancer and 6 cancer-type-specific oncogenes in which MMs occur more frequently than expected: 9% of samples with at least one mutation in these genes harboured MMs. In various oncogenes, MMs are preferentially present in cis and show markedly different mutational patterns compared with single mutations in terms of type (missense mutations versus in-frame indels), position and amino-acid substitution, suggesting a cis-acting effect on mutational selection. MMs show an overrepresentation of functionally weak, infrequent mutations, which confer enhanced oncogenicity in combination. Cells with MMs in the *PIK3CA* and *NOTCH1* genes exhibit stronger dependencies on the mutated genes themselves, enhanced downstream signalling activation and/or greater sensitivity to inhibitory drugs than those with single mutations. Together oncogenic MMs are a relatively common driver event, providing the underlying mechanism for clonal selection of suboptimal mutations that are individually rare but collectively account for a substantial proportion of oncogenic mutations.

The advent of next-generation sequencing has enabled the processing of tens of thousands of tumours of many types for the systematic discovery of driver alterations<sup>3,4</sup>. These efforts have identified thousands of recurrent somatic mutations across cancers, with few highly frequent (major) mutations and a much larger number of rare (minor) mutations. Although the latter account for a substantial portion of the accumulated mutations, even for oncogenes, the vast majority of them are regarded as functionally weak mutations or of uncertain significance; however, how these minor mutations are clonally selected despite limited functionality and what their genetic differences are from major mutations are poorly understood.

Tumour suppressor genes (TSGs) are frequently affected by multiple (typically biallelic) loss-of-function mutations<sup>3,4</sup>. There have also been many cancer cases in which specific secondary mutations—including the *EGFR* T790M and *KIT* V654A mutations—have been acquired in already mutated oncogenes following tyrosine kinase inhibitor (TKI) therapy<sup>5–8</sup>. However, only sporadic reports have investigated MMs arising in the same oncogene during cancer initiation and development, before any therapy<sup>1,2</sup>. Therefore, the frequency, spectrum and genetic features of MMs across oncogenes remain unclear. In addition, whether oncogenic MMs occur as a combination of driver–driver or driver–passenger mutations, what their biological and clinical implications are, and how they

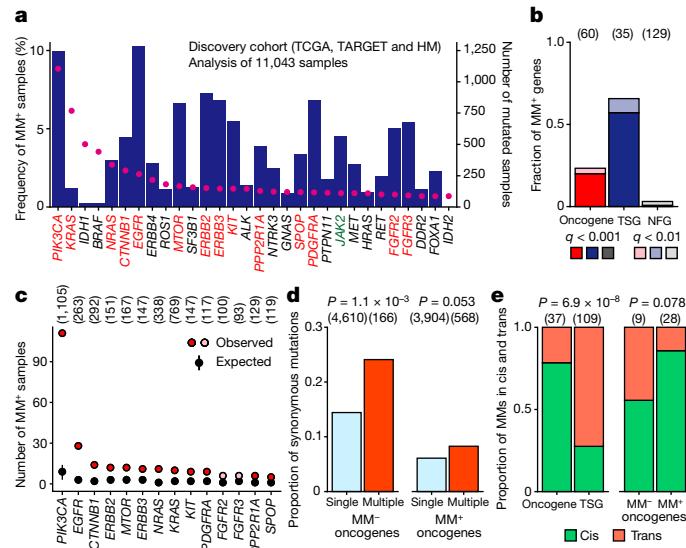
differ from single mutations have not been well studied. Here we present a systematic pan-cancer analysis of more than 60,000 samples and delineate the overall landscape and genetic properties of oncogenic MMs. We explore their functionality and influences on molecular machinery and drug sensitivity, as well as potential synergistic interactions (the epistatic effect) between mutations, focusing on *PIK3CA* and *NOTCH1*.

## Frequent MMs in a variety of oncogenes

We assembled, annotated and systematically curated a large repository of cancer genome data from 60,954 samples of primary and metastatic cancer across more than 150 cancer types from five cohorts: The Cancer Genome Atlas (TCGA), Therapeutically Applicable Research to Generate Effective Treatments (TARGET), Foundation Medicine (FM), the American Association for Cancer Research (AACR) Project Genomics Evidence Neoplasia Information Exchange (GENIE), and the haematological malignancy (HM) cohort, comprising acute myeloid leukaemia (AML) and non-Hodgkin's lymphoma (NHL) (Extended Data Figs. 1, 2a, b and Supplementary Table 1; see Methods). In particular, to eliminate probable false-positive MMs, we rigorously reviewed certain variant types (Extended Data Fig. 2c–e), namely dinucleotide and trinucleotide

<sup>1</sup>Division of Molecular Oncology, National Cancer Center Research Institute, Tokyo, Japan. <sup>2</sup>Department of Gastroenterology, Keio University School of Medicine, Tokyo, Japan. <sup>3</sup>Department of Clinical System Onco-Informatics, Graduate School of Medicine, Kyoto University, Kyoto, Japan. <sup>4</sup>Department of Urology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan.

<sup>5</sup>Department of Hematology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan. <sup>6</sup>Medical Sciences Innovation Hub Program, RIKEN Cluster for Science, Technology and Innovation Hub, Yokohama, Japan. <sup>7</sup>Research and Development Group for In Silico Drug Discovery, Center for Cluster Development and Coordination, Foundation for Biomedical Research and Innovation, Kobe, Japan. <sup>8</sup>Laboratory of Sequence Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan. <sup>9</sup>Center for Cancer Genomics and Advanced Therapeutics, National Cancer Center, Tokyo, Japan. <sup>10</sup>These authors contributed equally: Yuki Saito, Junji Koya. <sup>✉</sup>e-mail: kkataoka-tky@umin.ac.jp



**Fig. 1 | Entire landscape of MMs in various oncogenes.** **a**, Number of mutated samples (dots) and frequency of MM<sup>+</sup> samples (bars) for 30 oncogenes (with 80 or more mutated samples) in the discovery cohort ( $n=11,043$ ). Oncogenes significantly affected by MMs ( $q<0.01$  and three or more MMs) in pan-cancer and cancer-type-specific analyses are indicated in red and green, respectively. **b**, Fraction of MM<sup>+</sup> genes in oncogenes, TSGs and NFGs. Numbers of genes examined are shown in parentheses. **c**, Number of MM<sup>+</sup> samples for 14 MM<sup>+</sup> oncogenes (with  $q<0.01$  and three or more MMs). Red ( $q<0.001$ ) and pink ( $q<0.01$ ) circles show observed values; black circles indicate expected values (median with 95% confidence intervals). Values in parentheses indicate numbers of samples examined. **d**, **e**, One-sided permutation test ( $n=10,000$ ) with Benjamini–Hochberg correction. **d**, Proportion of synonymous to total mutations according to MM status (single or multiple) in MM<sup>-</sup> and MM<sup>+</sup> oncogenes. Values in parentheses indicate numbers of mutations examined. **e**, Proportion of MMs in cis and trans (with distances between mutations of 25 base pairs (bp) or more) by phasing from RNA-seq or WES/WGS in oncogenes and TSGs (left), and in MM<sup>-</sup> and MM<sup>+</sup> oncogenes (right). Numbers in parentheses indicate numbers of MMs examined. **d**, **e**, Two-sided Fisher's exact test.

variants (DNVs/TNVs)—single genetic events affecting multiple bases, which are usually misannotated as multiple single-nucleotide variants (SNVs) with different amino-acid substitutions, and which frequently affect certain mutational hotspots, such as KRAS amino-acid residue G12 and BRAFV600; and SNVs around insertions and deletions (indels), which are also often miscalled owing to ambiguous mapping.

First, we focused on whole-exome/whole-genome sequencing (WES/WGS) data of 11,043 primary untreated samples from TCGA, TARGET and HM (together the discovery cohort), and evaluated the frequency of MMs in 60 well-described oncogenes (Extended Data Figs. 1a, 2f and Supplementary Table 2). Although several highly mutated oncogenes, such as IDH1 and BRAF, contained no or few MMs, MMs were frequently observed across a wide variety of oncogenes; 5% or more of the mutated samples carried MMs across nine oncogenes, particularly in PIK3CA (10% of samples) and EGFR (10%), suggesting that MMs are a relatively common phenomenon that targets various oncogenes (Fig. 1a). MMs in oncogenes were not attributed to hypermutator tumours, and their frequency was consistent across cohorts and samples with varying tumour purity (Extended Data Fig. 3a–c).

MNs in oncogenes may arise as a pair of driver–driver, driver–passenger, or passenger–passenger mutations. To identify genes recurrently affected by driver–driver MMs ( $q<0.01$ ), we modified the permutation framework used to detect significantly mutated genes, while accounting for mutational signature, expression and DNA replication time<sup>9</sup> (Extended Data Fig. 3d–f). In contrast with nonfunctional genes (NFGs), of which only 3% were significant, MMs were found frequently in TSGs (66% of the analysed genes). Intriguingly, the significant enrichment of putative driver

MNs was observed in 14 (23%) oncogenes (Fig. 1b, c and Supplementary Table 2). These genes ('MM<sup>+</sup> oncogenes') included components of the phosphatidylinositol-3-kinase (PI3K) pathway (namely PIK3CA, MTOR and PPP2R1A), members of the RAS family of GTPases (KRAS and NRAS), and receptor tyrosine kinases (EGFR, ERBB2 and so on) (Extended Data Fig. 3g).

The proportion of synonymous to total mutations indicates the strength and mode of natural selection<sup>10</sup>. Regardless of mutation status (single versus multiple), this proportion was low in TSGs and relatively high in NFGs, suggestive of positive and neutral selection, respectively (Extended Data Fig. 3h). For MM<sup>-</sup> oncogenes, the proportion of synonymous mutations was low in samples with single mutations, but significantly increased in MM<sup>+</sup> samples (namely those with MMs in the same oncogene), suggesting weakened selective pressure for secondary mutations (that is, enrichment of driver–passenger MMs) in these genes (Fig. 1d and Extended Data Fig. 3i). By contrast, in MM<sup>+</sup> oncogenes, the proportion of synonymous mutations remained low (less than 0.1) even in MM<sup>+</sup> samples, in which both of the mutations in each MM are considered to undergo positive selection equal to the positive selection on single mutations, suggesting that most MMs arising in these genes are a combination of putative driver–driver mutations.

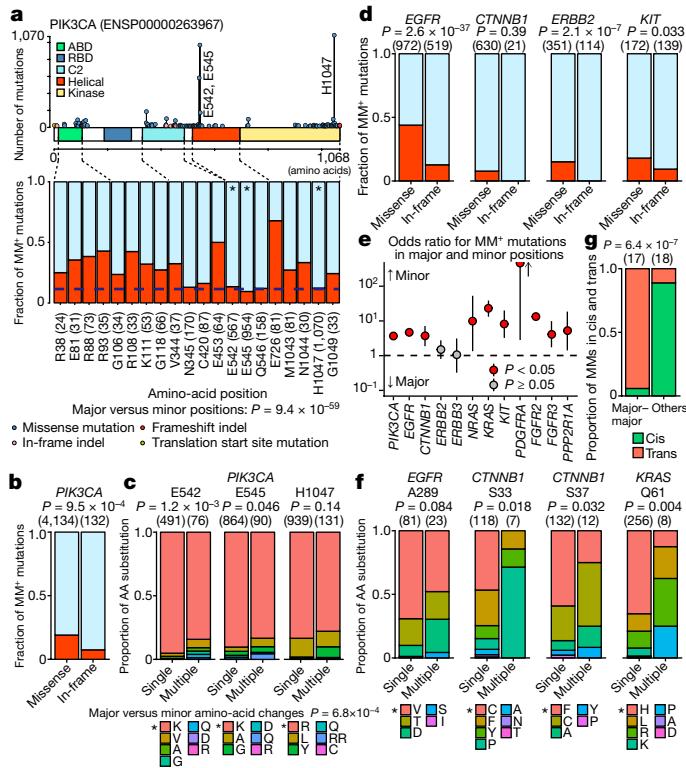
We investigated the allelic configuration of MMs by phasing from RNA sequencing (RNA-seq) or WES/WGS reads, which revealed that most MMs (78%) in oncogenes were present in cis with concordant allele frequencies, while the majority of MMs (72%) in TSGs were in trans (Fig. 1e, Extended Data Fig. 4a–c and Supplementary Table 3). Although several oncogenes were frequently affected by copy-number amplifications, MMs arose in cis even with concurrent amplification (Extended Data Fig. 4d, e). Notably, the proportion of MMs in cis was particularly high (86%) in MM<sup>+</sup> oncogenes, which was validated by Sanger sequencing of complementary DNA (cDNA) and long-read WGS in cell lines (Fig. 1e, Extended Data Fig. 4f–k).

## Selection pattern of MMs in PIK3CA

The presence of MMs in cis raised the possibility that an initial mutation may influence the clonal selection of subsequent mutations in MM<sup>+</sup> oncogenes, as they cause structural changes within the same molecule. Therefore, with the combined data of 40,002 primary samples from all five cohorts, we evaluated the mutational pattern of MMs in these genes (Extended Data Figs. 1a, 2f, 3c and Supplementary Table 4). To minimize the effect of confounding passenger mutations, we concentrated on the established hotspot and/or functionally relevant positions (1,679 positions for 14 MM<sup>+</sup> oncogenes) defined by knowledge-based databases and/or computational algorithms (Supplementary Tables 5–7; see Methods). Then, we classified them according to mutation type, position and amino-acid substitution, and compared their fraction of MM<sup>+</sup> mutations (namely those found in MM<sup>+</sup> samples) (Extended Data Fig. 1b).

We first examined PIK3CA, the most common oncogene in various cancers, which contained the largest number of MMs. In PIK3CA, MMs were more prevalent in missense mutations than in in-frame indels, and their frequency was quite different by mutational position (Fig. 2a, b and Supplementary Tables 5, 6). More than half of PIK3CA missense mutations were located in three major hotspots (E542, E545 and H1047), while almost all other mutations were in 275 minor hotspots (Extended Data Fig. 5a). Regardless of the cohort, MMs occurred more frequently in minor (32%) than major (11%) positions, particularly at E453 and E726, where 50% or more of the mutated cases possessed another mutation (Fig. 2a and Extended Data Fig. 5b). Moreover, even in major positions, less frequent (minor) amino-acid substitutions were more likely to be selected in samples with MMs (16–22%), compared with common (major) amino-acid substitutions (E542K, E545K and H1047R) (Fig. 2c and Supplementary Table 8). Together, these results clearly depict a substantial genetic difference between single and multiple mutations in PIK3CA.

Among PIK3CA MMs, major–minor combinations—such as E542–E726 and E726–H1047—were significantly enriched ( $q=0.0084$ ), followed by



**Fig. 2 | Biased selection for MMs across a wide range of oncogenes.** **a**, Distribution of mutations (top) and fraction of MM<sup>+</sup> mutations (bottom) for each hotspot/functional position (the top 20 are shown) in PIK3CA in recurrently mutated cancer types (defined as those with 20 or more hotspot/functional PIK3CA mutations) in primary samples from the total cohort. Asterisks indicate major positions (in which 10% or more of mutations were present in any of the recurrently mutated cancer types). The horizontal blue dotted line represents the mean value of major positions. Examined numbers of each mutation are in parentheses. ENSP00000263967 is the Ensembl database (<http://www.ensembl.org/>) identification code for PIK3CA protein. **b**, Fraction of MM<sup>+</sup> mutations for missense mutations and in-frame indels in PIK3CA. **c**, Proportion of each amino-acid (AA) substitution for each major position in PIK3CA for samples with single and multiple mutations. Asterisks indicate major amino-acid changes (occurring in more than 33% of samples with single mutations). **d**, Fraction of MM<sup>+</sup> mutations for missense mutations and in-frame indels across oncogenes (in which 100 or more mutations were present and 2% or more of them were indels). **e**, Comparison of the fraction of MM<sup>+</sup> mutations for major and minor positions across oncogenes (with more than 3 MM<sup>+</sup> samples). The y-axis shows odds ratios with 95% confidence intervals. See Supplementary Table 4 for sample size. **f**, Proportion of each amino-acid substitution for each major position (in which 100 or more mutations were present and 2% or more of them were from MM<sup>+</sup> samples) across oncogenes for samples with single and multiple mutations. Asterisks indicate major amino-acid changes. **b–d, f**, Values in parentheses indicate numbers of mutations examined. **g**, Proportion of MMs in cis and trans (with distances between mutations of less than 25 bp) by phasing from RNA-seq or WES/WGS in major-major combinations and others. Numbers in parentheses indicate numbers of MMs examined. **a–g**, Two-sided Fisher's exact test.

minor-minor combinations ( $q < 0.0001$ ; Extended Data Fig. 5c and Supplementary Table 9). In addition, mutational combinations involving different domains were overrepresented, particularly those involving the adaptor-binding domain (ABD) or C2 domain (such as R88–H1047 and E453–E545) (Extended Data Fig. 5d and Supplementary Table 10). Conversely, major-major combinations were substantially less common than expected, and among them those within the helical domain (E542–E545) were invariably present on different alleles (that is, in trans in the same clone or in different clones) (Extended Data Fig. 5c, e). On the basis of their allele frequencies, most of these mutational

combinations were estimated to occur in the same clone (concordant allele frequency), regardless of major or minor hotspots (Extended Data Fig. 5f). In the major–minor combinations showing discordant allele frequencies, mutations at major hotspots tended to have a higher allele frequency than at minor hotspots, implying that the mutations at major hotspots are likely to be earlier events causing intermediate clonal expansions (Extended Data Fig. 5g).

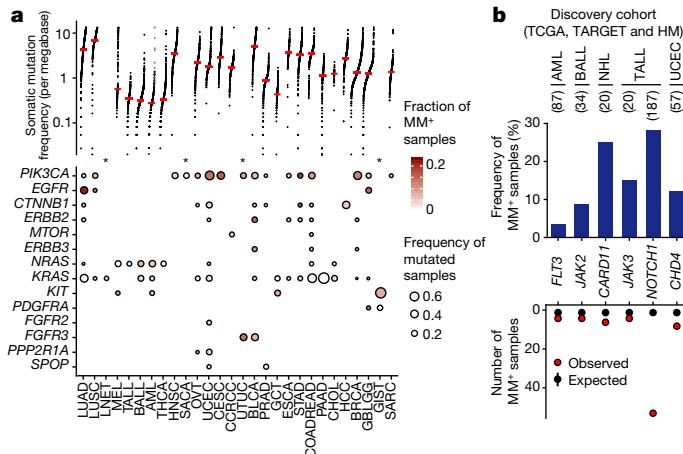
To understand the mechanism that generates MMs, we investigated the mutational signature caused by the cytidine deaminase APOBEC (C-to-G/T at TpCpX trinucleotides). Reflecting the underrepresentation of APOBEC signatures in minor hotspot mutations, we found that samples harbouring PIK3CA MMs showed a higher overall APOBEC activity but a lower proportion of APOBEC signature mutations in PIK3CA itself, compared with samples harbouring no or single PIK3CA mutations. This suggests that a single mutational process does not account for this phenomenon (Extended Data Fig. 5h, i).

## Widespread biased selection for MMs

The biased selection of PIK3CA MMs led us to assess the mutational spectrum of MMs in other MM<sup>+</sup> oncogenes. In MM<sup>+</sup> oncogenes, among 12,753 hotspot/functional missense mutations (87% of the total mutations), 79% resided in major hotspot positions (of which 36% showed less frequent (minor) amino-acid substitutions), whereas 21% were in minor positions (Extended Data Fig. 6a and Supplementary Table 4). Similar to the case of PIK3CA, missense mutations had, or tended to have, a higher fraction of MMs than did in-frame indels in other examined MM<sup>+</sup> oncogenes, most especially in EGFR and ERBB2 (Fig. 2d and Supplementary Tables 5, 6). Conspicuously, missense mutations in minor positions contained a larger fraction of MMs than did those in major positions in various MM<sup>+</sup> oncogenes, except for ERBB2/ERBB3 (Fig. 2e and Extended Data Fig. 6b). For instance, in KRAS almost no cases with major hotspot mutations (G12, G13 and Q61) harboured MMs, whereas a variable proportion of MMs was observed in minor hotspots. In CTNNB1, the fraction of MMs was considerably higher in minor than in major positions, even within the same functional site ( $\beta$ -TrCP-binding site: D32–S37). These findings were strengthened by combinatorial analysis, which showed a greater occurrence of major–minor and/or minor–minor combinations than expected (Extended Data Fig. 6c and Supplementary Table 9). Other genetic characteristics of PIK3CA MMs were also observed in these MM<sup>+</sup> oncogenes: minor amino-acid substitutions were more common in major hotspots in MM<sup>+</sup> oncogenes, including EGFR A289, CTNNB1 S33 and S37, and KRAS Q61 (Fig. 2f and Supplementary Table 8); and within MMs in close proximity—in contrast to other combinations mostly present in cis—almost all major–major combinations, such as NRAS/KRAS G12–G13 pairs, occurred on different alleles (Fig. 2g and Extended Data Fig. 6d). A substantial proportion of these MMs had discordant allele frequencies, and even those with concordant allele frequencies showed lower allele frequencies, probably suggesting that they arise as different subclones (Extended Data Fig. 6e, f). Overall, our findings demonstrate a skewed mutational pattern for MMs that is pervasive across various oncogenes.

## Lineage specificity of oncogenic MMs

The frequency of MMs remained consistent among recurrently mutated cancer types in several genes, such as EGFR and KRAS, whereas it varied extensively across cancer types in other genes, including PIK3CA and CTNNB1, independent of mutation burden (Fig. 3a and Extended Data Fig. 7a). Particularly, MMs in PDGFRA were observed in 12% of the mutated samples in glioma (glioblastoma/low-grade glioma (GBLGG)), but there was no MM<sup>+</sup> sample in gastrointestinal stromal tumour (GIST) (Extended Data Fig. 7b). In PIK3CA, the distribution of MMs resembled each other across cancer types (Extended Data Fig. 7c). By contrast, the distribution of MMs differed in EGFR between non-small-cell lung



**Fig. 3 | Cancer-type specificity and diversity of oncogenic MMs.** **a**, Somatic mutation frequency per megabase calculated from WES/WGS (top), and frequency of mutated samples and fraction of MM<sup>+</sup> samples for each MM<sup>+</sup> oncogene (bottom), across human cancer types (with more than 150 samples and 20 or more hotspot/functional mutations) in primary samples from the total cohort ( $n=40,002$ ). Top, each dot represents a sample and red horizontal lines indicate the median number in each cancer type. Asterisks indicate cancer types without WES/WGS data. BRCA, invasive breast carcinoma; CCRCC, clear cell renal-cell carcinoma; CESC, cervical squamous cell carcinoma; CHOL, cholangiocarcinoma; COADREAD, colorectal adenocarcinoma; ESCA, oesophageal adenocarcinoma; GCT, germ cell tumour; HCC, hepatocellular carcinoma; HNSC, head and neck squamous cell carcinoma; LNET, lung neuroendocrine tumour; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; MEL, melanoma; PAAD, pancreatic adenocarcinoma; PRAD, prostate adenocarcinoma; SACA, salivary carcinoma; SARC, soft tissue sarcoma; STAD, stomach adenocarcinoma; THCA, thyroid carcinoma. **b**, Frequency of MM<sup>+</sup> samples (top) and number of MM<sup>+</sup> samples (bottom) for six cancer-type-specific MM<sup>+</sup> oncogenes (with  $q < 0.01$  and three or more MMs) in the discovery cohort. Red and black circles indicate observed and expected (median with 95% confidence intervals) values. One-sided permutation test ( $n=10,000$ ) with Benjamini–Hochberg correction. Examined numbers are shown in parentheses.

cancer (NSCLC) and GBLGG: the fraction of MMs in major hotspots and minor positions in the kinase domain was almost equivalent between both tumours, whereas those involving minor positions in the extracellular domain were more common in NSCLC (Extended Data Fig. 7d, e), indicating substantial lineage specificity among oncogenic MMs.

In several genes, MMs were nearly exclusively identified in a certain cancer type or anatomically related ones, including FGFR2 and PPP2R1A mutations in gynaecological cancers (ovarian epithelial tumour (OVT) and/or uterine corpus endometrial carcinoma (UCEC)) and FGFR3 mutations in urothelial tumours (bladder urothelial carcinoma (BLCA) and upper tract urothelial carcinoma (UTUC)). For FGFR3 in urothelial tumours, missense mutations within the juxtamembrane region (codons 420–450) were observed only in MM<sup>+</sup> samples (Extended Data Fig. 8a). We then performed a permutation test for each cancer type from the discovery cohort with an extended list of 84 oncogenes. This analysis identified six cancer-type-specific MM<sup>+</sup> oncogenes that had been overlooked by our pan-cancer analysis—particularly those involving haematological neoplasms, such as NOTCH1, CARD11 and JAK3 (Fig. 3b, Extended Data Fig. 8b, c and Supplementary Tables 11, 12). Among these, NOTCH1 mutations in T-cell acute lymphoblastic leukaemia (TALL) showed the highest prevalence of MM<sup>+</sup> samples (28%), in contrast to chronic lymphocytic leukaemia (CLL), for which no MM<sup>+</sup> samples were observed (Extended Data Fig. 8d, e). In NOTCH1 MMs, missense mutations or in-frame indels in the HD domain significantly co-occurred with truncating mutations in the PEST domain ( $q=0.0050$  and  $q<0.0001$  for HD-N and HD-C, respectively; two-sided simulation test), both of which are gain-of-function mutations<sup>11</sup> (Extended Data

Fig. 8f). Consistent with the MM<sup>+</sup> oncogenes identified by our pan-cancer analysis, a biased mutational pattern was observed in cancer-type-specific MM<sup>+</sup> oncogenes: MMs were more common in missense mutations than in in-frame indels, and in minor than in major positions, in *FLT3* mutations in AML and/or *JAK2* mutations in B-cell acute lymphoblastic leukaemia (BALL) (Extended Data Fig. 8g, h).

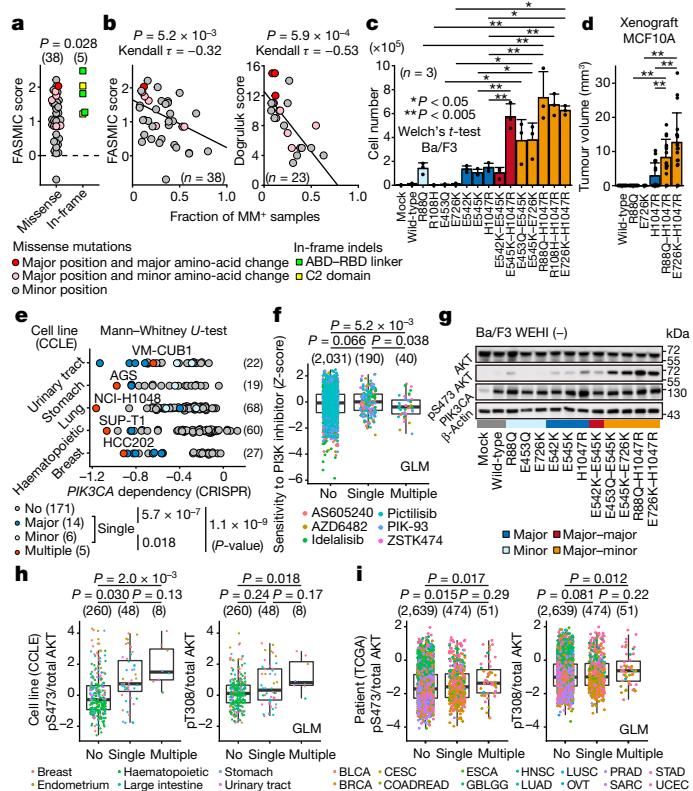
We next compared 40,002 primary and 20,952 metastatic cancers, which demonstrated nearly equivalent frequencies of MMs in MM<sup>+</sup> oncogenes, except for *EGFR* and *KIT* (Extended Data Fig. 9a, b). After exclusion of previously reported TKI-resistant mutations, including *EGFR* T790M and *KIT* V654A mutations in NSCLC and GIST, respectively<sup>5,6</sup>, the frequency of MMs in these genes was comparable between primary and metastatic samples (Extended Data Fig. 9c, d). Also, MMs in metastatic samples were enriched in missense mutations affecting minor hotspot positions, similarly to those in primary samples (Extended Data Fig. 9e–g). These findings suggest that, even when acquired TKI-resistant mutations are excluded, a biased mutational pattern for MMs is observed in metastatic samples.

## Functional relevance of MMs

To clarify the biological relevance of oncogenic MMs, we evaluated the functional activity scores for *PIK3CA* estimated from *in vitro* and/or *in vivo* assays of mutant-transduced cell lines<sup>12,13</sup>. Despite their low frequency, in-frame indels had a higher score than missense mutations, suggesting a negative association between functionality and MM fraction (Fig. 4a). Among missense mutations, functional activity scores were inversely correlated with the fraction of MMs, suggesting the selection of functionally weak mutations in MMs (Fig. 4b). In line with this suggestion, Ba/F3 cells transduced with *PIK3CA* major hotspot mutants exhibited increased growth, whereas those transduced with minor hotspot mutants (except R88Q) did not, compared with wild-type-transduced cells in the absence of cytokine (Fig. 4c and Extended Data Fig. 10a). Notably, major–minor double mutants markedly enhanced proliferation compared with single mutants. A major–major double mutant involving different domains (E545K–H1047R) had a similar effect, whereas one within the same domain (E542K–E545K) exhibited no synergistic activity. Moreover, double mutants further augmented tumour growth *in vivo* compared with single mutants following xenotransplantation of MCF10A cells (Fig. 4d).

These results suggest that individually suboptimal mutations can confer enhanced oncogenic potential in MMs, which might explain their skewed mutational pattern. Notably, analysis of CRISPR–Cas9 loss-of-function screens in Cancer Cell Line Encyclopedia (CCLE) cell lines<sup>14</sup> revealed that most cell lines with *PIK3CA* MMs had the highest relative dependency on this gene within each cancer type (Fig. 4e). Furthermore, cells harbouring *PIK3CA* MMs exhibited stronger *AKT1* dependency and higher sensitivity to PI3K inhibitors than those with no or single *PIK3CA* mutations, pointing to the potential value of MMs as a predictive marker for targeted therapies (Fig. 4f and Extended Data Fig. 10b, c).

In terms of their underlying molecular mechanism, Ba/F3 cells expressing *PIK3CA* major hotspot mutants displayed increased AKT activation (that is, phosphorylation) than those expressing wild-type or minor hotspot mutants, but major–minor double mutants showed much greater AKT activation than single mutants (Fig. 4g and Extended Data Fig. 10d, e). A similar synergistic effect was observed for major–major double mutants affecting different domains, but not for those within the same domain. Moreover, *PIK3CA* double mutants strongly enhanced mTOR, p70S6K, PRAS40 and GSK-3β phosphorylation downstream of AKT. These findings were supported by reverse-phase protein array (RPPA) data from CCLE cell lines, with elevated AKT phosphorylation (at S473 and T308) observed in MM<sup>+</sup> samples (Fig. 4h and Extended Data Fig. 10f). In addition, TCGA patients showed a greatly increased activation of AKT in MM<sup>+</sup> tumours, although a range of activity was evident within each cancer type (Fig. 4i and Extended Data Fig. 10g).



**Fig. 4 | PIK3CA suboptimal mutations confer enhanced oncogenicity and downstream pathway activation in combination.** **a**, Functional activity score for PIK3CA in the Functional Annotation of Somatic Mutations in Cancer database (FASMIC; <https://ibl.mdanderson.org/fasmic/>) according to mutation type. Two-sided Brunner–Munzel test. **b**, Association between fraction of MM<sup>+</sup> samples and PIK3CA functional activity score, in FASMIC<sup>13</sup> (left) and ref.<sup>12</sup> (right). Two-sided Kendall's rank correlation test. **a**, **b**, Each dot represents a mutation (present in five or more samples), shaped and coloured by type, position and amino-acid change. **c**, Proliferation of Ba/F3 cells expressing mock control or wild-type, single-mutant or double-mutant PIK3CA without WEHI-3-conditioned medium for three days. Two-sided Welch's *t*-test on log-transformed values. **d**, Xenograft tumour volumes (at eight weeks post-injection) resulting from MCF10A cells expressing mock control or wild-type, single-mutant or double-mutant PIK3CA. See Source Data for sample size. Two-sided Welch's *t*-test. **e**, **f**, Data represent means + s.d. **e**, Dependency of 196 CCLE cell lines on PIK3CA, with cell lines coloured by MM status, in the DepMap CRISPR–Cas9 knockout data (see Methods). Two-sided Mann–Whitney *U*-test. **f**, Box plots showing sensitivity to six PI3K inhibitors (*Z*-scores) for 417 CCLE cell lines, according to MM status. **g**, Immunoblot analysis of AKT, AKT with phosphorylated S473 (pS473 AKT), PIK3CA and β-actin in Ba/F3 cells expressing mock control or wild-type, single-mutant or double-mutant PIK3CA without WEHI-3-conditioned medium. Representative of three independent experiments. See Supplementary Fig. 1 for source images. **h**, **i**, Box plots showing levels of phosphorylated S473 (pS473; left) and T308 (pT308; right) in RPPA data relative to total AKT protein expression for 316 CCLE cell lines (with cancer types for which there are cell lines harbouring PIK3CA MMs) (**h**) and in 3,164 TCGA patients (with purity of 50% or more, in recurrently mutated cancer types) (**i**) according to MM status. Each dot represents a sample, coloured by cancer type. **f**, **h**, **i**, Box plots show medians (lines), interquartile ranges (IQR; boxes) and  $\pm 1.5 \times$  IQR (whiskers). GLM, generalized linear model (see Methods). Examined numbers are shown in parentheses.

Among genetic alterations involved in the PI3K pathway, PIK3CA mutations tended to co-occur with PTEN mutations, but be mutually exclusive with PIK3R1 mutations and PIK3CA amplifications in UCEC, which had the highest frequency of the mutated samples (Extended Data Fig. 11a, b). These associations were augmented by PIK3CA MMs,

suggesting that a second mutation can further enhance the functional interplay among driver alterations of the PI3K pathway.

We then investigated the structural mechanism of MM-mediated PIK3CA (p110 $\alpha$  subunit) activation by means of molecular-dynamics simulations of the synergistic mutants R88Q–H1047R, which show the strongest activity in vitro. Examination of the overall structure and various residue–residue contacts confirmed the predicted single-mutant-induced conformational changes. First, the R88–D746 salt bridge between the ABD and kinase domains was rendered unstable by the R88Q mutation, which promoted rotation of the iSH2 domain, contributing to exposure of the kinase domain. Second, the H1047R mutation distorted the orientation of the kinase domain, broadening substrate accessibility<sup>15,16</sup> (Extended Data Fig. 11c, d and Supplementary Table 10). Unexpectedly, although R88Q and H1047R single mutants slightly affected the R38–D743 salt bridge, the R88Q–H1047R double mutant caused its cleavage, leading to further detachment of the interface between the ABD and kinase domains and subsequent kinase exposure (Extended Data Fig. 11e–g). Therefore, a coordinated structural alteration might underlie the enhancement of downstream pathway activation by PIK3CA MMs.

We also found MM-induced enhanced functional activity in another gene: CRISPR–Cas9 screens in the CCLE revealed that a cell line with MMs in NOTCH1 had the highest relative dependency on NOTCH1, among cell lines from haematopoietic and lymphoid tissue (Extended Data Fig. 11h). Single missense mutations in the NOTCH1 HD domain increased NOTCH1 transcriptional activity compared with the wild type; but these mutations in combination with a truncating mutation further augmented its transcriptional activity (Extended Data Fig. 11i). Together, these observations reinforce the idea that MMs in the same oncogene cooperate to potentiate its tumour-promoting activity.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2175-2>.

- Kohsaka, S. et al. A method of high-throughput functional evaluation of EGFR gene variants of unknown significance in cancer. *Sci. Transl. Med.* **9**, eaan6566 (2017).
- Madsen, R. R. et al. Oncogenic PIK3CA promotes cellular stemness in an allele dose-dependent manner. *Proc. Natl Acad. Sci. USA* **116**, 8380–8389 (2019).
- Garraway, L. A. & Lander, E. S. Lessons from the cancer genome. *Cell* **153**, 17–37 (2013).
- Vogelstein, B. et al. Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).
- Heinrich, M. C. et al. Molecular correlates of imatinib resistance in gastrointestinal stromal tumors. *J. Clin. Oncol.* **24**, 4764–4774 (2006).
- Kobayashi, S. et al. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **352**, 786–792 (2005).
- Pao, W. et al. Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Med.* **2**, e73 (2005).
- Soh, J. et al. Oncogene mutations, copy number gains and mutant allele specific imbalance (MASI) frequently occur together in tumor cells. *PLoS One* **4**, e7464 (2009).
- Lawrence, M. S. et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
- Martincorena, I. et al. Universal patterns of selection in cancer and somatic tissues. *Cell* **171**, 1029–1041 (2017).
- Weng, A. P. et al. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science* **306**, 269–271 (2004).
- Dogruluk, T. et al. Identification of variant-specific functions of PIK3CA by rapid phenotyping of rare mutations. *Cancer Res.* **75**, 5341–5354 (2015).
- Ng, P. K. et al. Systematic functional annotation of somatic mutations in cancer. *Cancer Cell* **33**, 450–462 (2018).
- Ghandi, M. et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503–508 (2019).
- Gkeka, P. et al. Investigating the structure and dynamics of the PIK3CA wild-type and H1047R oncogenic mutant. *PLOS Comput. Biol.* **10**, e1003895 (2014).
- Zhang, M., Jang, H. & Nussinov, R. The mechanism of PI3K $\alpha$  activation at the atomic level. *Chem. Sci.* **10**, 3671–3680 (2019).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

# Article

## Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

### Data set preparation

As the discovery data set, we obtained BAM files and mutational data in Mutation Annotation Format (MAF) (compiled by the MC3 Working Group for TCGA) for WES/WGS from TCGA (phs000178.v10.p8)<sup>17–19</sup>, TARGET (phs000218.v20.p7)<sup>20–22</sup>, and HM (AML (phs001657.v1.p1)<sup>23</sup> and NHL from the University of Iowa/Mayo Clinic Lymphoma Specialized Program of Research Excellence (SPORE) (phs000450.v3.p1)<sup>24</sup>) cohorts from the Genomic Data Commons (GDC) (<https://gdc.cancer.gov/>), the TARGET Data Matrix (<https://ocg.cancer.gov/programs/target/data-matrix/>), or the cBioPortal for Cancer Genomics (<https://www.cbioperl.org/>) (Extended Data Figs. 1a, 2a). As the additional data set, mutation calls for targeted-sequencing data from MSK-IMPACT and the Dana Farber Cancer Institute (DFCI) Oncopanel of the American Association for Cancer Research Project GENIE (Release 5.0; syn7222066)<sup>25</sup> and FM (Release 1.0; phs001179.v1.p1)<sup>26</sup> were downloaded from Synapse (<https://www.synapse.org/>) and the GDC Data Portal, respectively.

As synonymous mutations were excluded from the original mutational data for the HM AML cohort, mutation calling was performed as described with some modifications<sup>27</sup> (see Supplementary Methods). Genomic coordinates of mutations from alignments to the human reference genome GRCh38 were converted to GRCh37 (hg19) using LiftOver for the FM cohort. Then, the mutation calls were annotated to gene transcripts in Ensembl (release 95), and a single canonical effect per mutation was reported using Variant Effect Predictor (version 95.3) and vcf2maf (version 1.6.16). Germline variants had already been excluded in each cohort before downloading from the above sources. Briefly, matched normal samples were used in TCGA, TARGET, HM NHL and GENIE MSK-IMPACT cohorts. Further filtering for TCGA and HM NHL was performed with a pool of unmatched normal samples. Finally, all cohorts used single-nucleotide polymorphism (SNP) databases, such as Exome Aggregation Consortium (ExAC) and The Genome Aggregate Database, to further remove putative germline variants.

To eliminate probable false-positive MMs, we developed DNVChecker, which detects multiple SNVs occurring in the same codon, evaluates their allelic status using the corresponding BAM files, and changes them into a single alteration if they exist in cis (Extended Data Fig. 2c). We applied this to the discovery cohort and curated 13,482 DNVs and 812 TNVs. As more than 95% of the multiple SNVs in the same codon were DNVs/TNVs in the discovery cohort, all of them were converted to single alterations in the additional cohort, for which BAM files were unavailable. Moreover, we removed missense mutations (mostly SNVs) located within 3 bp of indels, as almost all of them were considered as a single genetic event (Extended Data Fig. 2e). On the basis of mutant allele frequencies obtained from MAF files or calculated from BAM files using SAMtools (version 1.4.1) mpileup, mutations with low allele frequencies (less than 0.05), together with non-coding mutations, were removed.

Samples with missing data, derived from xenografts, and duplicated in the same patient were also excluded from the analysis. Hypermutator samples (with 500 or more and 20 or more coding mutations per exome (for the discovery cohort) and per targeted region (for the additional cohort), respectively) were also removed. In the remaining samples of the discovery cohort, 88 samples showed microsatellite instability (defined as having an MSIsensor score of 4 or more<sup>17</sup>) or pathogenic *POLE/POLD1* mutations<sup>28</sup>, whose effect on MMs was independently evaluated. Only untreated sample data were used in the discovery data set. Although information on prior treatment was not available for the additional data set, we confirmed that the overall frequency and distribution of MMs were consistent across cohorts (Extended Data Figs. 3c, 5a, b).

The total cohort consisted of 40,002 primary and 20,952 metastatic samples representing more than 150 cancer types, which had 1,224,150 (synonymous and nonsynonymous) mutations (Extended Data Figs. 1a, 2b). Among the primary samples, the discovery cohort ( $n = 11,043$ ) included 924,518 mutations, of which 4,472 mutations arose in 14 MM<sup>+</sup> oncogenes, while the additional cohort ( $n = 28,959$ ) included 15,467 mutations in these oncogenes (Extended Data Fig. 2f). The metastatic samples, consisting mainly of the FM and GENIE cohorts, harboured 149,190 mutations, which contained 11,556 mutations in 14 MM<sup>+</sup> oncogenes (Extended Data Fig. 9a). Cancer type was classified on the basis of OncoTree codes (<https://www.cbioperl.org/oncotree/>), with slight modifications (Supplementary Table 1). Independent sets of AML<sup>29</sup>, BALL<sup>30–33</sup> and CLL<sup>34</sup> patients were analysed for *FLT3*, *JAK2* and *NOTCH1*, respectively.

### Oncogenes, TSGs and NFGs

To characterize MMs in oncogenes, we compiled a high-confidence list of 60 oncogenes, 35 TSGs and 129 NFGs (Supplementary Table 2). Oncogenes were selected if they were: (i) listed as ‘oncogenes’ in ref.<sup>35</sup>; or (ii) manually rescued as well described oncogenes (including *JAK2*, *PIK3R2*, *PTPN11*, and *SPOP*<sup>4,18</sup>). Putative TSGs (*CUL3*, *POLE* and *MED12*), target genes of aberrant somatic hypermutation (*BCL2*, *HIST1H1C* and so on) and oncogenes specific to haematological neoplasms (*ABL1*, *FLT3* and so on) were excluded from the pan-cancer oncogene list, although the latter genes were assessed as cancer-type-specific oncogenes. TSGs were selected if they were listed as ‘tsg’ in PANCAN from the TCGA Pan-Cancer Atlas project<sup>19</sup> and described as ‘TSG’ in ref.<sup>4</sup>. NFGs included genes frequently affected by passenger hotspot mutations and olfactory genes described in ref.<sup>36</sup> and ref.<sup>9</sup>, respectively. Cancer-type-specific oncogenes were selected from those listed as ‘oncogene’ or ‘possible oncogene’ in a certain cancer type from the TCGA Pan-Cancer Atlas project and manually curated. Paediatric and haematological cancer-type-specific oncogenes were added from the literature. In total, we compiled an extended list of 84 cancer-type-specific oncogenes, including 38 genes listed in the pan-cancer oncogene list (Supplementary Table 11).

### Identification of MM<sup>+</sup> oncogenes by permutation

To identify genes significantly affected by driver mutations, a permutation test is widely used, where the expected number of samples with mutations in gene X (the gene of interest) is estimated by permuting all coding (synonymous and nonsynonymous) mutations randomly across the coding region in all samples. Statistical significance is determined by comparing the observed number of samples with nonsynonymous mutations and the expected distribution in gene X. Here, to identify genes significantly affected by putative driver–driver MMs, we modified the permutation framework, such that the expected number of samples with MMs in gene X (that is, with additional gene X mutations) was estimated by permuting all coding mutations other than gene X mutations in samples harbouring gene X mutations (Extended Data Fig. 3d). For the permutation of SNVs, we considered: (i) sequence composition and mutational signature; (ii) expression; and (iii) DNA replication time, all of which are known to affect mutational frequency. Copy-number alterations were not taken into consideration, as they did not affect the frequency of MMs in our analysis (Extended Data Fig. 4d).

To account for sequence composition and mutational signature, permutations were restricted within each of the 64 trinucleotide contexts (2 types of mutated base (C or T)  $\times$  2 types of transcriptional strand  $\times$  4 types of 5' base  $\times$  4 types of 3' base). In addition, all coding sequences were divided into 225 bins according to expression (mean expression of 1,156 CCLE cell lines) and DNA replication time<sup>9</sup> (15 bins each), and the synonymous mutation rate was calculated in each bin and used to weight the sampling probability (Extended Data Fig. 3e, f). For example, a C-to-T substitution at ACG was moved to another ACG site of the same strand from the coding region randomly with the weighted

probabilities. Then, the permuted SNVs were classified into nonsynonymous or synonymous mutations, depending on the reading frame. Indels and other mutation types, such as DNVs/TNVs, were separately moved to a random position without weighting. The expected number of samples with additional gene X mutations—including nonsynonymous mutations and indels, but not synonymous mutations—was estimated, and its distribution (10,000 permutations) was compared with the observed number of samples with MMs. The advantage of the permutation test is that gene length and mutation burden (total number of mutations per sample) are considered: samples with high mutation burden have more mutations to be randomized, and genes with long coding sequences have a higher probability to obtain additional mutations. In pan-cancer analysis, 60 oncogenes, 35 TSGs and 129 NFGs were analysed in the discovery cohort, and multiple testing was adjusted for using the Benjamini–Hochberg method. To confirm the results, we estimated the selective pressure for MMs in these genes. Both synonymous and nonsynonymous mutations were collected, and the proportion of synonymous to total mutations was compared between samples with single and multiple mutations. In cancer-type-specific analysis, 200 gene–cancer-type pairs, including 46 oncogenes that had not been analysed in pan-cancer analysis, were evaluated.

### Hotspot/functional mutation analysis

We first filtered out loss-of-function mutations, such as splice-site, nonsense, and frameshift mutations, unless they were listed as ‘hotspot/functional’ mutations (such as *NOTCH1* truncating mutations). Then, we divided the remaining mutations into ‘hotspot/functional’ and ‘non-hotspot/functional’ ones by position, and further classified the former into missense mutations and in-frame indels (Extended Data Fig. 1). The positions of hotspot/functional mutations were defined according to four knowledge-based databases (OncoKB<sup>37</sup>, PMKB<sup>38</sup>, Cancer Genome Interpreter (CGI)<sup>39</sup> and Clinical Interpretations of Variants in Cancers (CIVIC)<sup>40</sup>) and four computational algorithms (3D Hotspots<sup>41</sup>, Cancer Hotspots<sup>42,43</sup>, HotMAPS<sup>44</sup> and CHASMplus<sup>45</sup>) (Supplementary Tables 5–7). For cancer-type-specific mutations, functionally validated mutation positions, such as *JAK2* I682, were also included (Supplementary Tables 13–15). Among 19,154 mutations in 14 MM<sup>+</sup> oncogenes in primary samples from the total cohort (after excluding 482 loss-of-function and 303 synonymous mutations), 14,645 mutations were present in recurrently mutated cancer types (defined as those with 20 or more hotspot/functional mutations in the gene), among which 13,708 mutations (94%) were in 1,679 hotspot/functional positions. These hotspot/functional missense mutational positions were further subdivided into major and minor positions: major ones were defined as those in which 10% or more of mutations were present in any of the recurrently mutated cancer types, and others were classified into minor ones. Amino-acid substitutions in major positions (excluding those frequently affected by major–major combinations in trans) were also subdivided into major and minor changes, which were defined as those present in 33% or more and less than 33% of cases with single mutations, respectively (Supplementary Table 8). Then, the fraction of MM<sup>+</sup> mutations—namely those found in samples harbouring MMs in the same oncogene (namely MM<sup>+</sup> samples)—was compared according to type, position and amino-acid change, in the recurrently mutated cancer types.

### Enrichment of mutational combinations

Enrichment of a specific pair of mutations were assessed by Monte-Carlo simulation according to hotspot/functional position and functional domain. For analysis of hotspot/functional position, only combinations consisting of a pair of missense mutations in hotspot/functional positions were analysed. The expected number of combinations was estimated by randomly sampling two mutations (without replacement) on the basis of the mutation frequency in cases with single mutations (with 10,000 iterations). Statistical significance was determined by comparing the observed number of combinations and

the expected distribution, and adjusted for multiple testing by the Benjamini–Hochberg method.

### Determination of cis and trans phase of MMs

To determine allelic configurations of MMs, we developed CisChecker, which first identifies MMs within the same gene from mutation call data, then extracts sequencing reads encompassing both mutational positions from the corresponding BAM file, and classifies them into those containing both mutant alleles (cis reads), one mutant and one reference allele (trans reads), or both reference alleles (reference reads) (Extended Data Fig. 4a). MMs supported by two or more cis reads and no trans read were considered as cis, and vice versa as trans. Otherwise, the allelic status was assessed by a permutation test, where each mutation was randomly permuted in each position among sequencing reads encompassing both mutational positions. Then, the expected numbers of cis and trans reads were estimated by 10,000 permutations and compared with the observed numbers of cis and trans reads, respectively.

Using the MAF and RNA-seq BAM files, we applied this algorithm to 1,417 MMs (283 and 1,134 MMs in oncogenes and TSGs, respectively) from 1,082 samples in the discovery cohort, among which 220 MMs (16%) were evaluable. Among them, the allelic status for 142 MMs could be assessed by WES/WGS, and almost all of them (more than 99%) were validated. On the basis of this finding, 120 MMs from 101 samples whose RNA-seq data were not available were analysed only by WES/WGS, among which 27 MMs (23%) were evaluable. In the combined data (247 MMs), allelic configuration was compared according to gene category, mutational combination, and distance between mutations. Representative MMs were visualized with the Integrated Genomics Viewer (version 2.4.10). The difference in allele frequency between MMs (combinations of missense mutations only) in the same gene was assessed by Fisher’s exact test (with Benjamini–Hochberg correction) for the number of mutant and reference reads, and those with significantly different allele frequencies (with *q*-values of less than 0.05 and a mutant allele frequency difference of greater than 0.10) were considered as ‘discordant’, and otherwise as ‘concordant’. For *PIK3CA* major–minor combinations showing discordant allele frequencies, we used a binomial test to evaluate whether the proportion of major and minor dominant MMs (based on mutant allele frequency) deviated from the expected distribution in which the order of major and minor mutations was random (that is, the probabilities of major and minor dominant MMs were equal).

### Association with clinical and genetic information

Tumour purity inferred by ABSOLUTE (TCGA\_mastercalls.abs\_tables\_JSedit.fixed.txt), In Silico Admixture Removal (ISAR)-corrected Genomic Identification of Significant Targets in Cancer (GISTIC) copy-number data (ISAR\_GISTIC.all\_thresholded\_by\_genes.txt) and batch-corrected RPPA data (TCGA-RPPA-pancan-clean.txt) by the TCGA Pan-Cancer Atlas project<sup>19</sup> were obtained from the GDC.

### CCLE cell line data

The mutation call data (depmap\_19Q1\_mutation\_calls\_v2.csv) for 1,601 cell lines and batch-corrected genome-wide CRISPR–Cas9 knockout screen data (gene\_effect\_corrected.csv for Public 19Q1) for 558 cell lines, RPPA data (CCLE\_RPPA\_20181003.csv) for 899 cell lines, and drug-sensitivity data (v17.3\_fitted\_dose\_response.xlsx) for 1,065 cell lines were obtained from the DepMap (<https://depmap.org/portal/>), CCLE (<https://portals.broadinstitute.org/ccle/>), and Genomics of Drug Sensitivity in Cancer (GDSC; <https://www.cancerrxgene.org/>) databases, respectively. Mutations with allele frequencies of 0.05 or more were assessed and cell lines (with less than 5,000 coding mutations) were classified according to *PIK3CA* or *NOTCH1* mutation status. Dependency on *PIK3CA*, *AKT1* and *NOTCH1* was examined by ranking the CERES score (a measure of gene dependency) in each cancer type (Mann–Whitney *U*-test).

# Article

## RPPA and drug-sensitivity analysis

For RPPA, levels of AKT with phosphorylated S473 and T308 relative to total AKT protein expression were compared in 14 recurrently mutated cancer types (TCGA) and 6 cancer types for which there are *PIK3CA*-MM-harbouring cell lines (CCLE), using the generalized linear model (GLM) with cancer type as a covariate. For drug-sensitivity analysis, we used at least three *PIK3CA*-MM-harbouring cell lines to test six PI3K inhibitors, namely AS605240 (GDSC drug identification code 224), AZD6482 (156), Idelalisib (238), Pictilisib (1,058), PIK-93 (303) and ZSTK474 (223), by GLM, using drug and cancer type as covariates.

## PIK3CA functional activity score

The biological activity of *PIK3CA* hotspot/functional missense mutations and in-frame indels (five or more samples) was evaluated on the basis of two previous functional studies<sup>12,13</sup>. In ref. <sup>12</sup>, five different functional assays—including growth-factor- and insulin-free survival, colony formation for MCF10A cells, interleukin-3-less survival for Ba/F3 cells, and tumorigenesis for immortalized mammary epithelial cells—were performed; we scored these results according to their functional impact (no, weak, intermediate and strong phenotype for 0, 1, 2 and 3, respectively) and summed them to yield a total score ranging from 0–15 for each mutant ( $n = 23$ ). For the cell viability data (v1) for Ba/F3 cells obtained from the FASMIC database (<https://ibl.mdanderson.org/fasmic/>)<sup>13</sup>, the cell number of each mutant was averaged in available replicates, normalized to wild-type *PIK3CA*, and log-transformed ( $n = 43$ ). A functional activity score was calculated as the arithmetic mean at well controlled time points (with wild-type *PIK3CA* relative to negative control (mCherry and/or Luc) value of 2.5 or less). Correlation between these functional scores and MM fraction was examined using Kendall's rank correlation coefficient.

## Cell lines, plasmid constructs and lentiviral transduction

Ba/F3-CL1, WEHI-3 and 293T cell lines were obtained from the RIKEN Cell Bank, the HEC-1 cell line from the JCRB Cell Bank, and the MCF10A, BT-20, NCI-H1048, SUP-T1 and HRT-18 cell lines from ATCC. Lenti-X 293T cells were purchased from TaKaRa. Cell lines were authenticated by the providers using karyotype, isoenzymes, and/or microsatellite profiling. They were cultured according to the providers' instructions, and routinely tested for mycoplasma infection. According to the International Cell Line Authentication Committee (ICLAC) register (<https://iclac.org/>), cross-contamination has been reported in BT-20 cells, although authentic stocks apparently do exist. We used the BT-20 cell line authenticated by ATCC using short tandem repeat (STR) profiling analysis. This cell line was selected because of the limited number of available cell lines harbouring MMs in *PIK3CA*.

Plasmids containing cDNAs encoding human wild-type (catalogue number 81736), R108H (82875), E453Q (82844), E545K (82881), E726K (82845) and H1047R (82824) *PIK3CA* were obtained from Addgene. R88Q, E542K and E542K–E545K mutants were constructed from the wild type by site-directed polymerase chain reaction (PCR)-based mutagenesis using PrimeSTAR MAX (TaKaRa). The PCR primers are listed in Supplementary Table 16. E545K–H1047R, E453Q–E545K, E545K–E726K, R88Q–H1047R, R108H–H1047R and E726K–H1047R double mutants were constructed from single-mutant vectors using standard restriction-enzyme-mediated cloning. These *PIK3CA* cDNAs were cloned into CSII-EF-RFA-IRES2-Venus (RDBO4389; from H. Miyoshi, Riken) using Gateway LR clonase II enzyme mix (Thermo Fisher Scientific). Lenti-X 293T cells were transiently transfected with lentiviral vectors, pMD2.G and psPAX2 (catalogue numbers 12259 and 12260; Addgene) using polyethylenimine 'max' (Polysciences). The viral supernatant was collected 48 h later and used for infection of Ba/F3-CL1 cells with RetroNectin (TaKaRa). After 48 h, the infected cells were harvested, sorted and subjected to culture.

Plasmids containing wild-type human *NOTCH1* cDNA (pFN21AE3300) were obtained from the Kazusa DNA Research Institute. *NOTCH1* L1600P, L1678P, Q2416\* (stop codon), L1600P–Q2416\* and L1678P–Q2416\* mutations were artificially synthesized (FASMAC) and cloned into pcDNA3 (Invitrogen) using standard restriction-enzyme-mediated cloning.

## Cell proliferation and viability assay

Ba/F3-CL1 cells were seeded in 12-well plates at  $1 \times 10^5$  cells per well and incubated for 72 h without WEHI-3 supernatant. Cell number and viability were determined using trypan blue (Nacalai Tesque) staining and cell counter model R1 (Olympus).

## In vivo xenograft tumour model

All mouse experiments were approved by the Animal Ethics Committee of the National Cancer Center and strictly adhered to its guidelines. Female BALB/c-nu/nu mice (six weeks old) were obtained from CLEA and maintained under pathogen-free conditions. Mice were subcutaneously injected with  $5 \times 10^5$  MCF10A cells in 1/1 phosphate-buffered saline (PBS)/matrigel (BD Biosciences), and tumour volumes were measured eight weeks after transplantation. The maximum tumour diameter permitted under the relevant animal protocols is 20 mm, and this limit was not exceeded in any experiment.

## Statistical analysis

Statistical analyses were performed with R3.6.0 software (The R Foundation for Statistical Computing). Comparison of categorical and continuous data was performed using two-sided Fisher's exact test and Brunner–Munzel test, respectively, unless otherwise specified. For functional assays, statistical significance was assessed by two-sided Welch's *t*-test, unless otherwise specified.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

Our findings are supported by data that are available from public online repositories, or data that are publicly available upon request from the data provider. See the 'Data set preparation' section above for details. Long-read WGS data for cell lines have been deposited in the European Genome-phenome Archive (EGA; <https://www.ebi.ac.uk/ega/>) under accession number EGAS00001003763. Data generated here are available as Source Data files accompanying Fig. 4 and Extended Data Figs. 10, 11.

## Code availability

Source codes for CisChecker, DNVChecker and the permutation test are available at <https://github.com/nccmo/CisChecker>, <https://github.com/nccmo/DNVChecker> and <https://github.com/nccmo/Permutation-test>, respectively.

17. Ding, L. et al. Perspective on oncogenic processes at the end of the beginning of cancer genomics. *Cell* **173**, 305–320 (2018).
18. Sanchez-Vega, F. et al. Oncogenic signaling pathways in The Cancer Genome Atlas. *Cell* **173**, 321–337 (2018).
19. Bailey, M. H. et al. Comprehensive characterization of cancer driver genes and mutations. *Cell* **173**, 371–385 (2018).
20. Ma, X. et al. Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. *Nature* **555**, 371–376 (2018).
21. Mullighan, C. G. et al. Deletion of *IKZF1* and prognosis in acute lymphoblastic leukemia. *N. Engl. J. Med.* **360**, 470–480 (2009).
22. Pugh, T. J. et al. The genetic landscape of high-risk neuroblastoma. *Nat. Genet.* **45**, 279–284 (2013).
23. Tyner, J. W. et al. Functional genomic landscape of acute myeloid leukaemia. *Nature* **562**, 526–531 (2018).

24. Chapuy, B. et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat. Med.* **24**, 679–690 (2018).
25. AACR Project GENIE Consortium. AACR Project GENIE: powering precision medicine through an international consortium. *Cancer Discov.* **7**, 818–831 (2017).
26. Hartmaier, R. J. et al. High-throughput genomic profiling of adult solid tumors reveals novel insights into cancer pathogenesis. *Cancer Res.* **77**, 2464–2475 (2017).
27. Kataoka, K. et al. Integrated molecular analysis of adult T cell leukemia/lymphoma. *Nat. Genet.* **47**, 1304–1315 (2015).
28. Campbell, B. B. et al. Comprehensive analysis of hypermutation in human cancer. *Cell* **171**, 1042–1056 (2017).
29. Papaemmanuil, E. et al. Genomic classification and prognosis in acute myeloid leukemia. *N. Engl. J. Med.* **374**, 2209–2221 (2016).
30. Gu, Z. et al. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat. Genet.* **51**, 296–307 (2019).
31. Li, J. F. et al. Transcriptional landscape of B cell precursor acute lymphoblastic leukemia based on an international study of 1,223 cases. *Proc. Natl Acad. Sci. USA* **115**, E11711–E11720 (2018).
32. Steeghs, E. M. P. et al. JAK2 aberrations in childhood B-cell precursor acute lymphoblastic leukemia. *Oncotarget* **8**, 89923–89938 (2017).
33. Forero-Castro, M. et al. Mutations in TP53 and JAK2 are independent prognostic biomarkers in B-cell precursor acute lymphoblastic leukaemia. *Br. J. Cancer* **117**, 256–265 (2017).
34. Puente, X. S. et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature* **526**, 519–524 (2015).
35. Bielski, C. M. et al. Widespread selection for oncogenic mutant allele imbalance in cancer. *Cancer Cell* **34**, 852–862 (2018).
36. Buisson, R. et al. Passenger hotspot mutations in cancer driven by APOBEC3A and mesoscale genomic features. *Science* **364**, eaaw2872 (2019).
37. Chakravarty, D. et al. OncoKB: a precision oncology knowledge base. *JCO Precis. Oncol.* **1**, 1–16 (2017).
38. Huang, L. et al. The cancer precision medicine knowledge base for structured clinical-grade mutations and interpretations. *J. Am. Med. Inform. Assoc.* **24**, 513–519 (2017).
39. Tamborero, D. et al. Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med.* **10**, 25 (2018).
40. Griffith, M. et al. CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nat. Genet.* **49**, 170–174 (2017).
41. Gao, J. et al. 3D clusters of somatic mutations in cancer reveal numerous rare mutations as functional targets. *Genome Med.* **9**, 4 (2017).
42. Chang, M. T. et al. Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity. *Nat. Biotechnol.* **34**, 155–163 (2016).
43. Chang, M. T. et al. Accelerating discovery of functional mutant alleles in cancer. *Cancer Discov.* **8**, 174–183 (2018).
44. Tokheim, C. et al. Exome-scale discovery of hotspot mutation regions in human cancer using 3D protein structure. *Cancer Res.* **76**, 3719–3731 (2016).
45. Tokheim, C. & Karchin, R. CHASMplus reveals the scope of somatic missense mutations driving human cancers. *Cell Syst.* **9**, 9–23 (2019).

**Acknowledgements** We acknowledge support from the Japan Society for the Promotion of Science (JSPS) KAKENHI (grant numbers 17K19592, 18K06594 and 15H05912) and National Cancer Center Research and Development Funds (30-A-1), together with many other funding bodies, organizations and individuals (see Supplementary Note).

**Author contributions** Y. Saito and K.K. designed the study. Y. Saito, Y.K. and K.K. analysed sequencing data. M.B.M., H.T., T.K., S. Miyano and Y. Shiraishi assisted with the collection and analysis of sequencing data. S.S. and M.T. performed sequencing experiments. J.K. and K.Y. performed immunoblots, capillary-based immunoassays, cell proliferation assays and in vivo xenograft assays. J.K. and Y.K. performed luciferase assays. M.A., S. Matsumoto, Y.I. and Y.O. performed molecular-dynamics simulations. Y. Saito and K.K. generated figures and tables and wrote the manuscript. K.K. led the entire project. All authors participated in discussions and interpretation of the data and results.

**Competing interests** The authors declare no competing interests.

**Additional information**

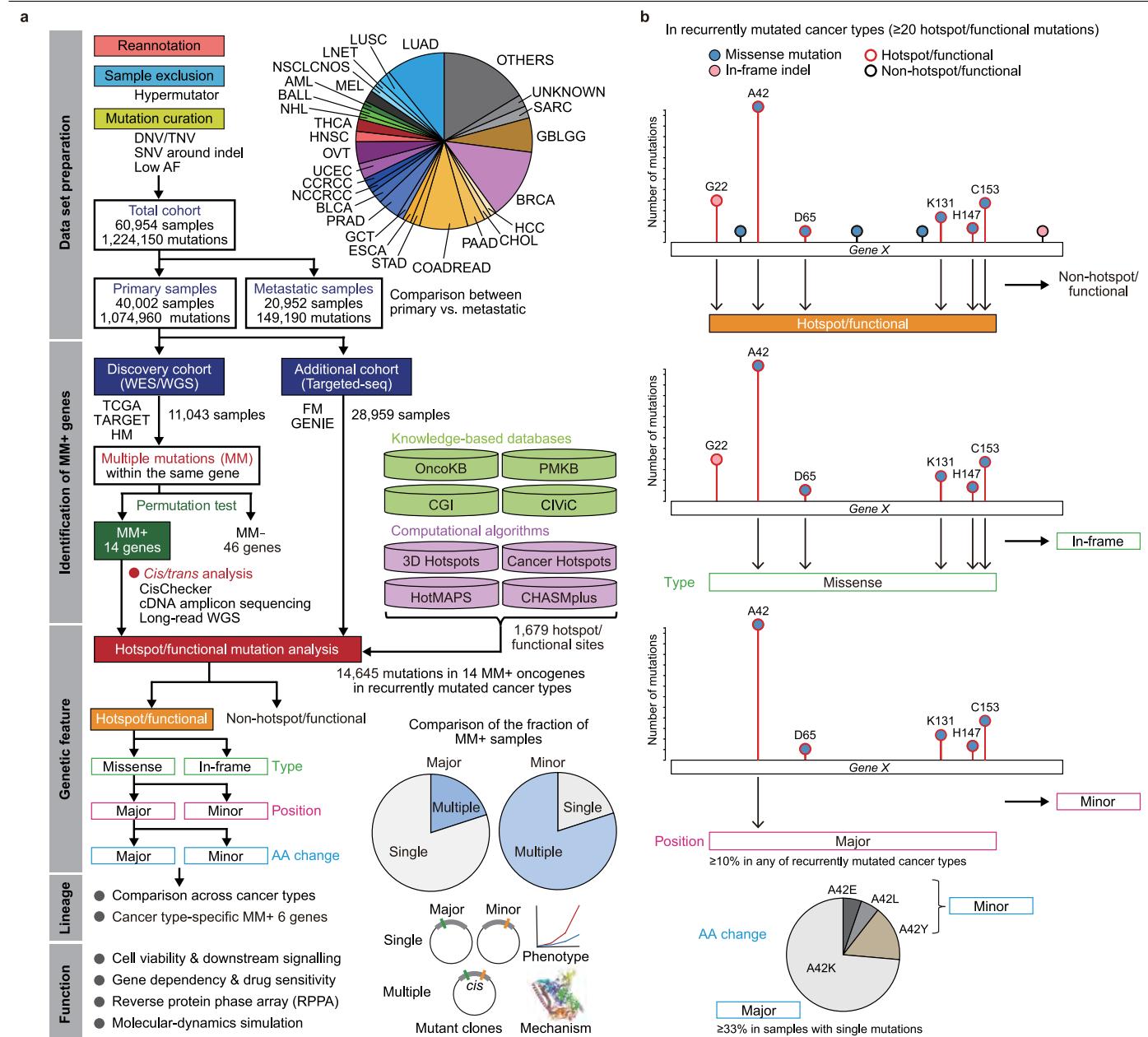
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2175-2>.

**Correspondence and requests for materials** should be addressed to K.K.

**Peer review information** *Nature* thanks Mark Lackner, Inigo Martincorena and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

# Article



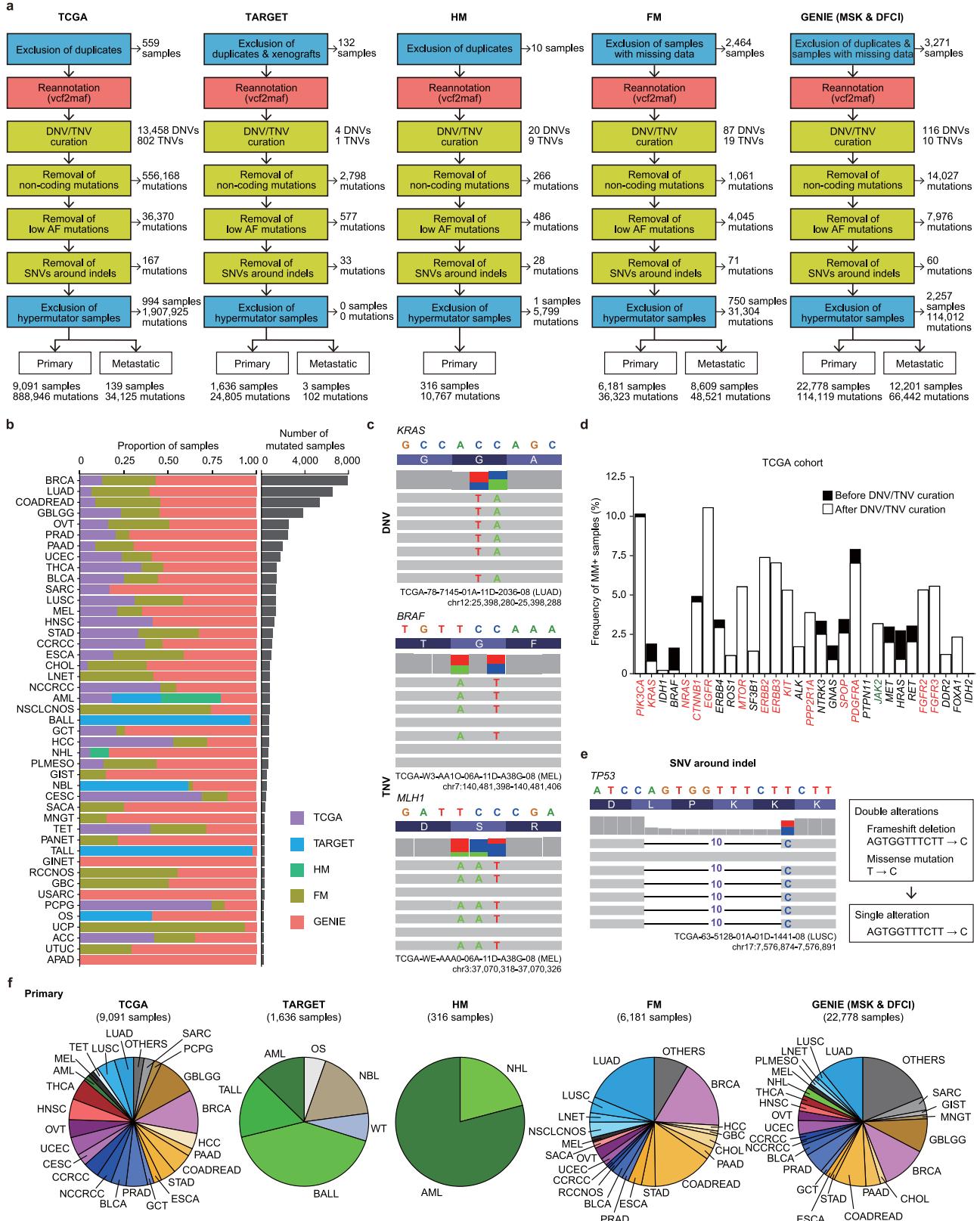
## Extended Data Fig. 1 | Workflow showing discovery and assessment of oncogenic MMs.

**a**, Data set preparation, identification of MM<sup>+</sup> oncogenes, and analysis of genetic features, lineage specificity and functional role.

Distribution of cancer types (Supplementary Table 1) for the total cohort ( $n=60,954$ ). Cancer types with less than 1% frequency were combined as

'OTHERS'. AF, allele frequency; NCCRCC, non-clear cell renal-cell carcinoma;

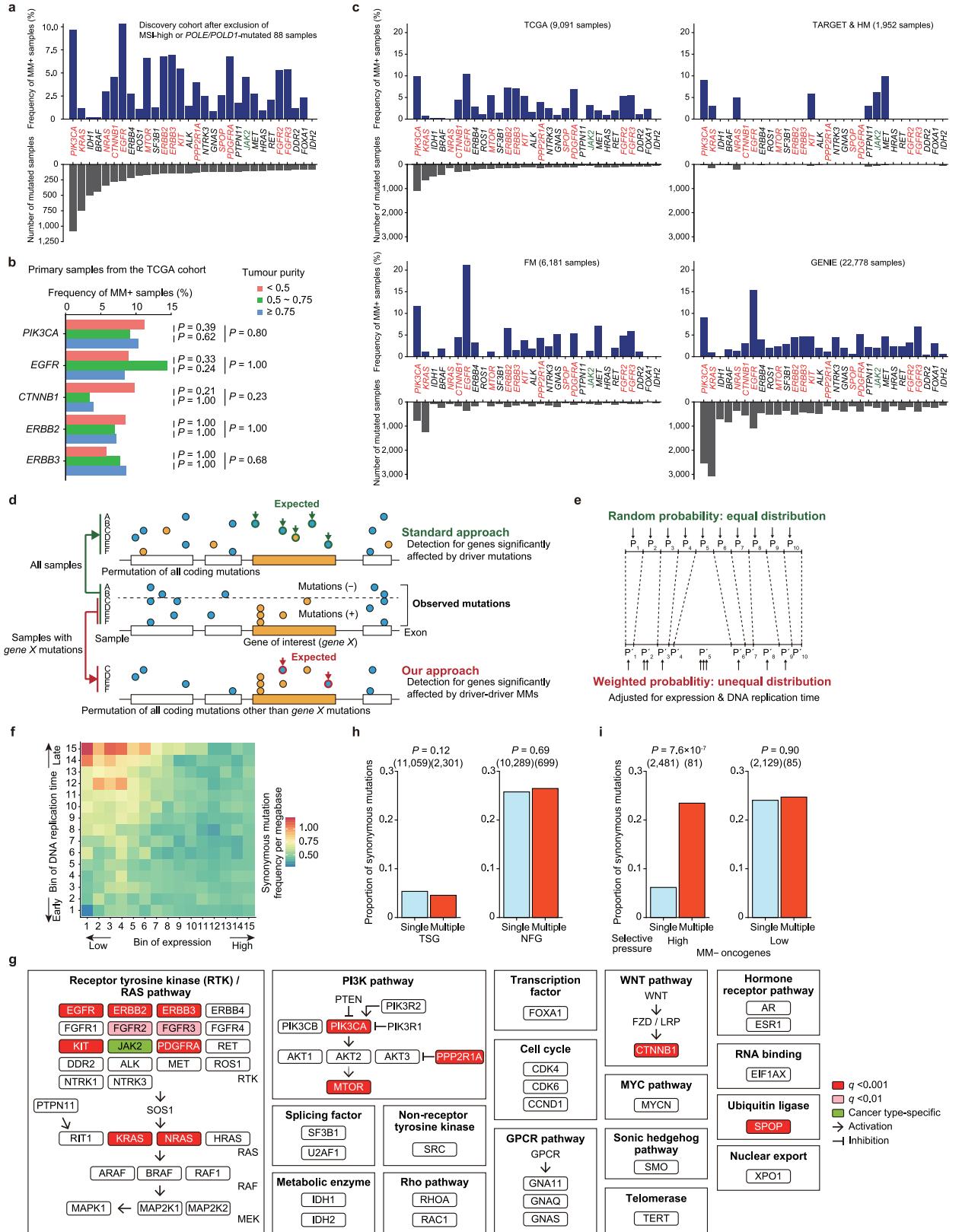
NSCLCNOS, NSCLC not otherwise specified. **b**, Classification of hotspot/functional mutations according to mutation type, position and amino-acid change.



**Extended Data Fig. 2** | See next page for caption.

## Article

**Extended Data Fig. 2 | Dataset preparation for the TCGA, TARGET, HM, FM and GENIE cohorts.** **a**, Steps involved in reannotation, sample exclusion and mutation curation in each cohort. **b**, Number of mutated samples (right) and proportion of samples in each cohort (left) for 45 cancer types (with 200 or more samples). ACC, adrenocortical carcinoma; APAD, appendiceal adenocarcinoma; GBC, gallbladder cancer; GINET, gastrointestinal neuroendocrine tumours; MNGT, meningothelial tumour; NBL, neuroblastoma; OS, osteosarcoma; PANET, pancreatic neuroendocrine tumour; PCPG, pheochromocytoma and paraganglioma; PLMESO, pleural mesothelioma; RCCNOS, renal-cell carcinoma not otherwise specified; TET, thymic epithelial tumour; UCP, undifferentiated carcinoma of the pancreas, USARC, uterine sarcoma/mesenchymal (see also Supplementary Table 1). **c**, DNVs and TNVs arising in the same codon (considered as a single genomic event) in representative cases. The TCGA identification code for the sample is shown below each chart. **d**, Frequency of MM<sup>+</sup> samples with or without DNV and TNV curation for 30 oncogenes (as in Fig. 1a) in primary samples from TCGA. **e**, Removed SNVs around indels. In this representative case, SNVs (T-to-C) and frameshift deletions (AGTGGTTTCTT-to-C) were independently miscalled, but considered to be a single genomic event. **f**, Distribution of cancer types for primary samples included in the TCGA, TARGET, HM, FM and GENIE cohorts. Cancer types with less than 1% frequency were combined as ‘OTHERS’ in each cohort. WT, Wilms’ tumour.

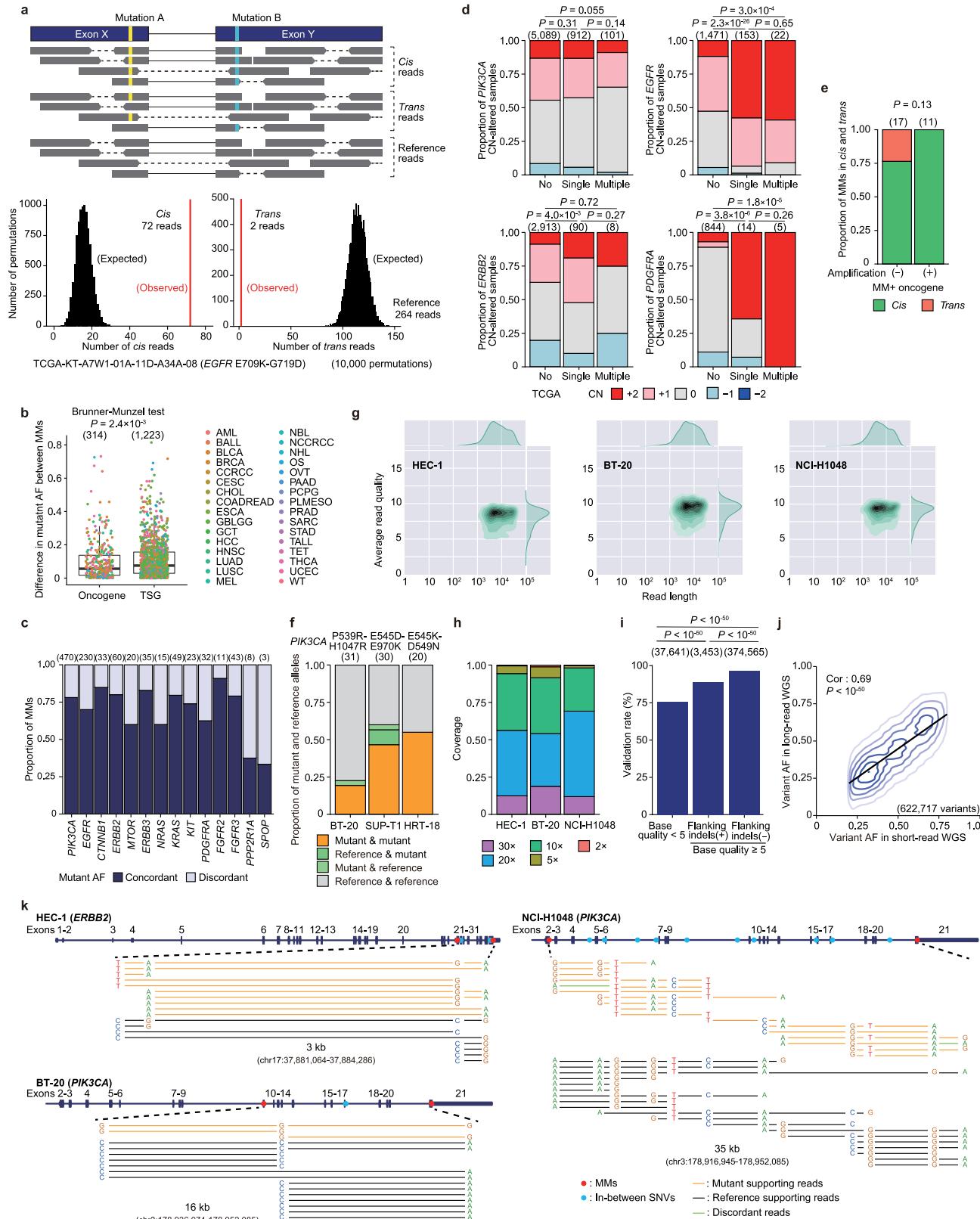


**Extended Data Fig. 3** | See next page for caption.

## Article

**Extended Data Fig. 3 | Identification of MM<sup>+</sup> oncogenes by a permutation test across cancers.** **a**, Number of mutated samples and frequency of MM<sup>+</sup> samples for 30 oncogenes (as in Fig. 1a) after excluding microsatellite instability (MSI)-high or *POLE/POLD1*-mutated samples in the discovery cohort. **b**, Frequency of MM<sup>+</sup> samples in *PIK3CA*, *EGFR*, *CTNNB1*, *ERBB2* and *ERBB3* according to tumour purity in primary samples from the TCGA cohort ( $n=8,699$ ). **c**, Number of mutated samples and frequency of MM<sup>+</sup> samples for 30 oncogenes (as in Fig. 1a) in the TCGA, TARGET, HM, FM and GENIE cohorts. **d**, Representation of the permutation-based framework. In the standard approach, to identify genes significantly affected by driver mutations, the expected number of samples with mutations in gene X (the gene of interest; green) is estimated by permuting all coding mutations randomly across the coding region in all samples (for example, samples A–F). Statistical significance is determined by comparing the observed number of samples with nonsynonymous mutations and the expected distribution in gene X. In our approach, to identify genes significantly affected by putative driver–driver MMs, the expected number of samples with MMs in gene X (red) is estimated by permuting all coding mutations other than gene X mutations in samples

harbouring gene X mutations (samples C–F). Statistical significance is determined by comparing the observed number of samples with MMs and the expected distribution in gene X. **e**, In the random-choice model, mutations are moved to another position with equal probability ( $P$ ), whereas in the weighted-choice model, mutations are moved with unequal probability ( $P'$ ), reflecting expression and DNA replication time. **f**, Synonymous mutation frequency per megabase according to expression and DNA replication time. **g**, Pathways related to 60 oncogenes analysed here. MM<sup>+</sup> oncogenes identified in pan-cancer and cancer type-specific analyses are indicated in red ( $q < 0.001$ )/pink ( $q < 0.01$ ) and green, respectively. **h**, Proportion of synonymous to total mutations according to MM status in TSGs and NFGs. **i**, Proportion of synonymous to total mutations according to MM status in MM<sup>+</sup> oncogenes under high and low selective pressure (that is, oncogenes in which the proportion of synonymous to total mutations is less than and more than 15% in samples with single mutations, respectively). The proportion of synonymous mutations was substantially increased in MM<sup>+</sup> samples, even in MM<sup>-</sup> oncogenes under high selective pressure. **b**, **h**, **i**, Two-sided Fisher's exact test. The numbers examined are shown in parentheses.



**Extended Data Fig. 4** | See next page for caption.

# Article

## Extended Data Fig. 4 | Allelic configuration (cis versus trans) of oncogenic

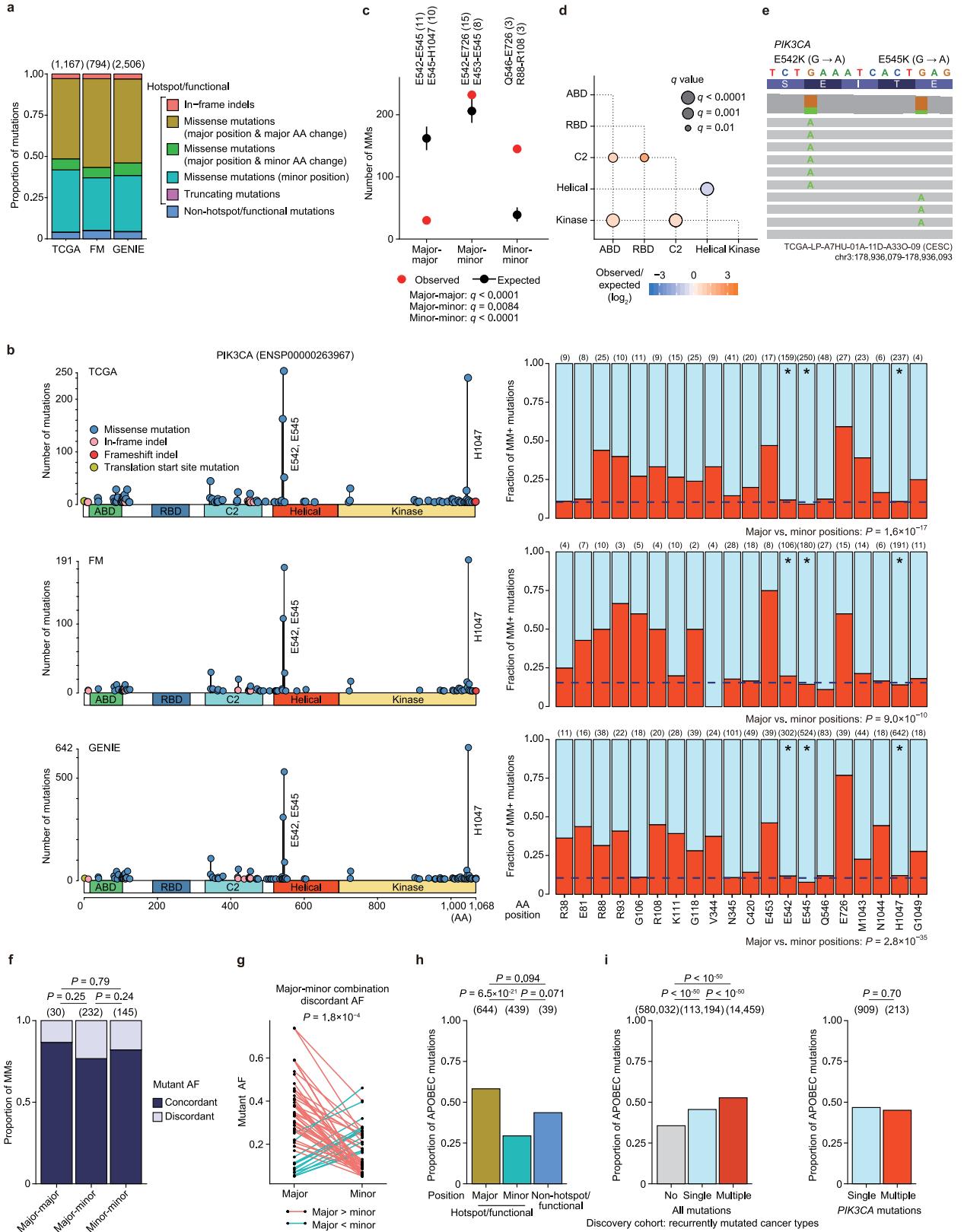
**MMs.** **a**, Top, sequencing reads encompassing MMs were classified into those containing both mutant alleles (cis reads), one mutant and one reference allele (trans reads), and both reference alleles (reference reads). Bottom, one-sided permutation test ( $n=10,000$ ) for the allelic configuration (cis versus trans) of MMs. In this representative case, the observed numbers of cis (left) and trans (right) reads (red) were significantly higher and lower, respectively, than the expected distribution (black); thus, this example is considered to be cis.

**b**, Difference in mutant allele frequency (AF) between MMs across 60 oncogenes and 35 TSGs in the discovery cohort. Each dot represents an MM, coloured by cancer type. Two-sided Brunner–Munzel test. Box plots show medians (lines), interquartile ranges (IQRs; boxes) and  $\pm 1.5 \times$  IQRs (whiskers). **c**, Proportion of MMs (combinations of missense mutations only) showing concordant or discordant allele frequencies in MM<sup>+</sup> oncogenes in primary samples from the total cohort. **d**, Fraction of *PIK3CA*, *EGFR*, *ERBB2* and *PDGFRA* copy-number (CN) alterations according to MM status in recurrently mutated cancer types (defined as those with 20 or more hotspot/functional mutations) in primary samples from TCGA. **e**, Proportion of MMs in cis and trans (with distances between mutations of 25 bp or more) by phasing from RNA-seq or

WES/WGS in MM<sup>+</sup> oncogenes with and without concurrent CN amplification of the mutated gene. **f**, Allelic configuration (cis versus trans) assessed by cDNA amplicon sequencing for *PIK3CA* P539R–H1047R, E545D–E970K and E545K–D549N mutations in BT-20, SUP-T1 and HRT-18 cell lines, respectively.

Proportions of mutant and reference alleles are shown. **b–f**, Examined numbers are shown in parentheses. **g**, Density plot illustrating the distribution of read length and average read quality for each of three long-read WGS samples. **h**, Percentage of bases covered by at least  $\times 2$ ,  $\times 5$ ,  $\times 10$ ,  $\times 20$  and  $\times 30$  sequencing reads for three long-read WGS samples. **i**, Validation rate of SNV calling from long-read WGS according to base quality and/or flanking indels. Examined read numbers are shown in parentheses. **d**, **e**, **i**, Two-sided Fisher's exact test.

**j**, Density plot showing the correlation between variant allele frequencies in short-read and long-read WGS in positions with coverage of at least  $\times 40$  and at least  $\times 20$ , in short-read and long-read WGS, respectively. Two-sided Pearson's correlation test. **k**, Phasing of MMs using long-read WGS reads. Positions of MMs (red) and in-between SNVs (blue) according to their genomic position (top) and long-read WGS reads between them (bottom). Reads supporting both mutant alleles and both reference alleles present in cis are shown in orange and black, respectively; discordant reads are shown in green.



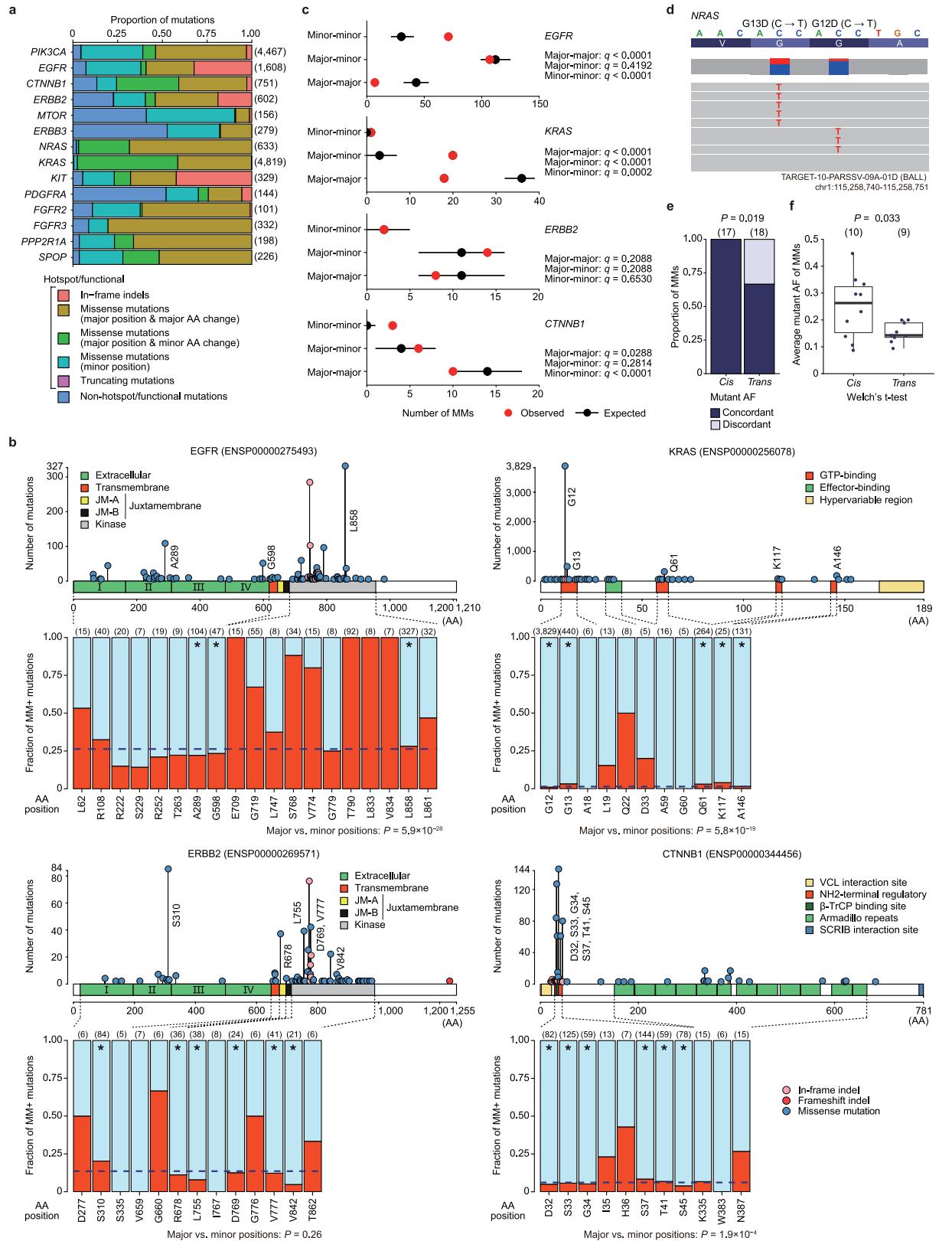
**Extended Data Fig. 5** | See next page for caption.

## Article

### Extended Data Fig. 5 | Comparison of *PIK3CA* MMs among cohorts. a,

Proportion of *PIK3CA* mutations according to type, position and amino-acid change in recurrently mutated cancer types (defined as those with 20 or more hotspot/functional mutations) in primary samples from the TCGA, FM and GENIE cohorts. **b**, Distribution of mutations and fraction of MM\* mutations for each hotspot/functional position. Asterisks indicate major positions (in which 10% or more of mutations were present in any of the recurrently mutated cancer types). The horizontal blue dotted lines represent the mean values of major positions. **c**, Number of MMs according to mutational combinations ( $n=407$ ). Red and black circles indicate, respectively, observed and expected (median with 95% confidence intervals) values. Representative combinations and observed numbers are also shown. **d**, Significant pairwise associations ( $q < 0.01$ ) with observed/expected ratios among functional domains ( $n=471$ ). Orange and blue colours depict co-occurring (observed number of MMs

significantly higher than expected) and mutually exclusive (lower than expected) associations. **e**, **f**, Two-sided simulation test ( $n=10,000$ ) with Benjamini–Hochberg correction. **e**, Major–major combinations on different alleles within close proximity (E542–E545). **f**, Proportion of MMs showing concordant or discordant allele frequencies according to mutational combinations. **g**, Order of major and minor hotspot mutations in major–minor combinations showing discordant allele frequencies ( $n=54$ ). Two-sided binomial test. **h**, Proportion of mutations with APOBEC signature (C-to-G/T at TpCpX trinucleotides) according to hotspot/functional position in *PIK3CA* in the discovery cohort. **i**, Proportion of mutations with APOBEC signature according to MM status in all coding (left) and *PIK3CA* (right) mutations in the discovery cohort. **b**, **f**, **h**, **i**, Two-sided Fisher’s exact test. Examined numbers are shown in parentheses.



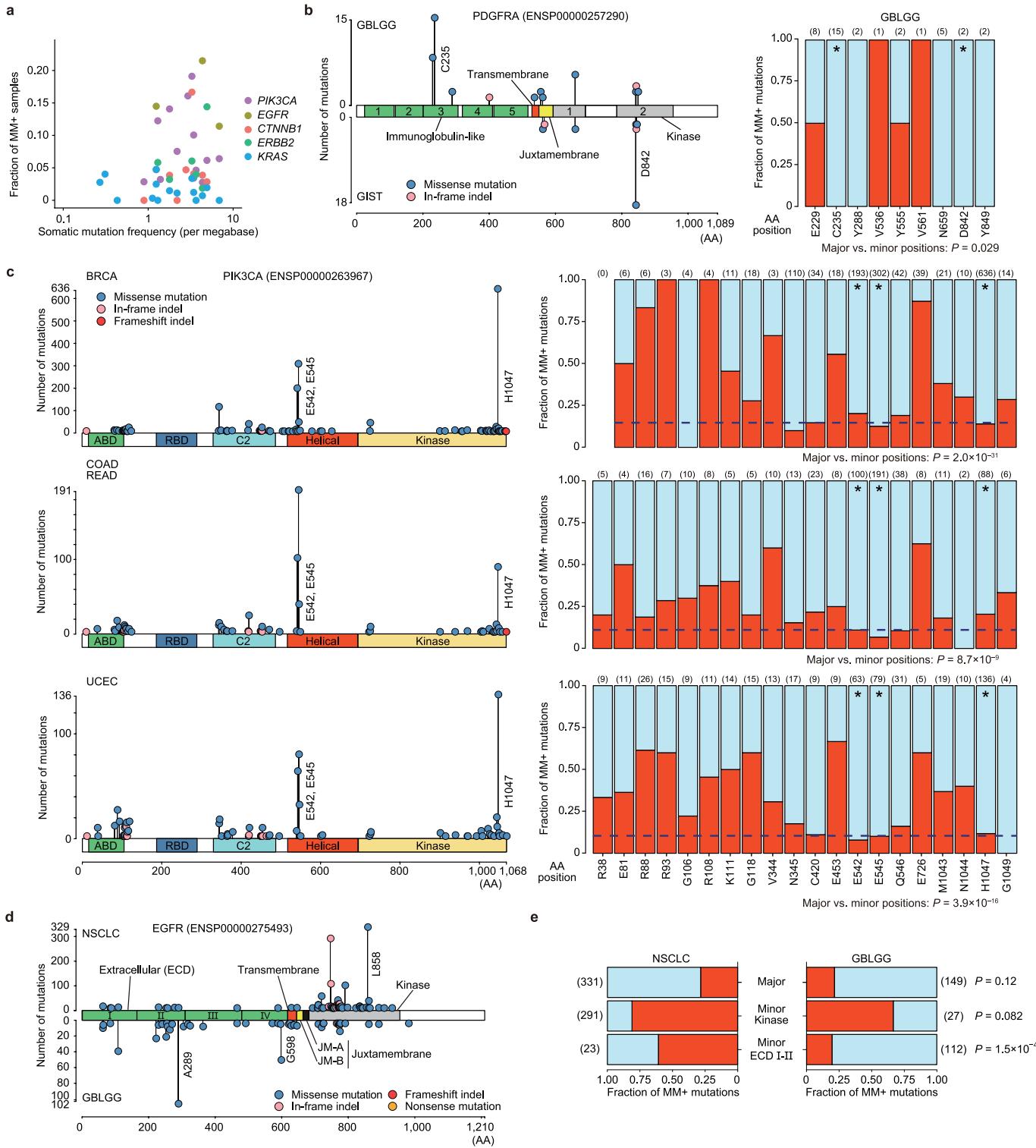
**Extended Data Fig. 6** | See next page for caption.

# Article

## Extended Data Fig. 6 | Genetic features of MMs in a variety of oncogenes.

**a**, Proportion of mutations according to type, position and amino-acid change across MM<sup>+</sup> oncogenes in recurrently mutated cancer types (defined as those with 20 or more hotspot/functional mutations) in primary samples from the total cohort. **b**, Distribution of mutations and fraction of MM<sup>+</sup> mutations for each hotspot/functional position in *EGFR*, *KRAS*, *ERBB2* and *CTNNB1*. Positions showing five mutations or more (and within the top 20 for *EGFR*) are shown in the bar plots. Asterisks indicate major positions (in which 10% or more of mutations were present in any of the recurrently mutated cancer types). The horizontal blue dotted lines represent the mean values of major positions. **c**, Number of MMs according to mutational combinations in *EGFR* ( $n=185$ ), *KRAS* ( $n=39$ ), *ERBB2* ( $n=24$ ) and *CTNNB1* ( $n=19$ ). Red and black circles indicate,

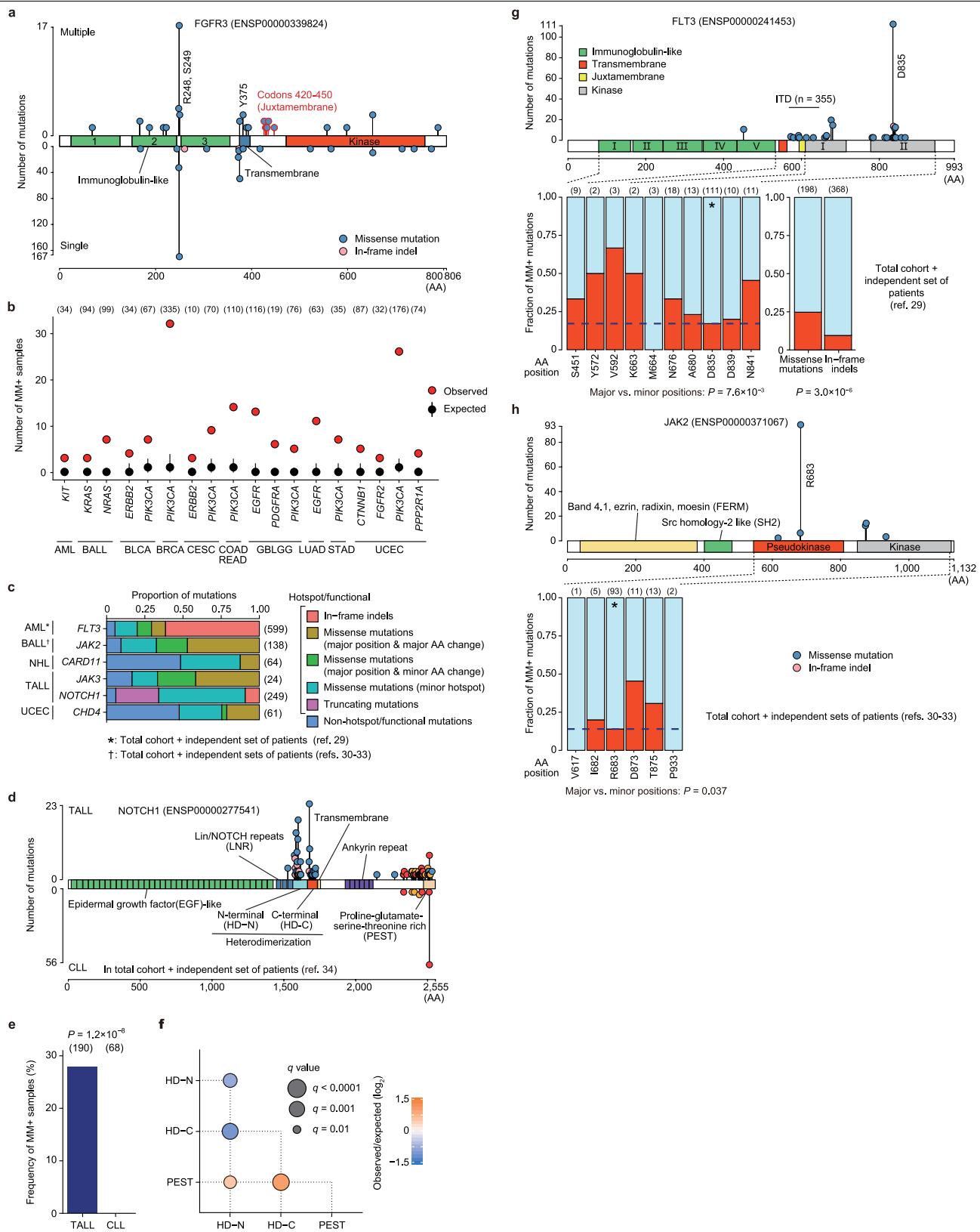
respectively, observed and expected (median with 95% confidence intervals) values. Two-sided simulation test ( $n=10,000$ ) with Benjamini–Hochberg correction. **d**, Major–major combinations on different alleles within close proximity (*NRAS* G12–G13). **e**, Proportion of MMs showing concordant or discordant allele frequencies (with distance between mutations of less than 25 bp) according to allelic configuration (cis versus trans). **f**, Average mutant allele frequency of MMs showing concordant allele frequencies present in copy-number-neutral region according to allelic configuration (cis versus trans). Box plots show medians (lines), interquartile ranges (IQRs; boxes) and  $1.5 \times$  IQRs (whiskers). Two-sided Welch's *t*-test. Examined numbers are shown in parentheses.



**Extended Data Fig. 7 | Similarities and differences of oncogenic MMs across cancer types.** **a**, Correlation between somatic mutation frequency per megabase (median value shown) and fraction of MM<sup>+</sup> samples according to MM<sup>+</sup> oncogenes in primary samples from the total cohort. Each dot represents a cancer type, coloured by gene. **b**, Distribution of mutations and/or fraction of MM<sup>+</sup> mutations for each hotspot/functional position in PDGFRA for GBLGG and GIST in primary samples from the total cohort. Asterisks indicate major positions (in which 10% or more of mutations were present in any of the recurrently mutated cancer types (defined as those with 20 or more hotspot/

functional mutations)). **c**, Distribution of mutations and fraction of MM<sup>+</sup> mutations for each hotspot/functional position in PIK3CA for BRCA, COADREAD and UCEC in primary samples from the total cohort. Asterisks indicate major positions. The horizontal blue dotted lines represent the mean values of major positions. **d**, **e**, Distribution of hotspot/functional mutations (**d**) and fraction of MM<sup>+</sup> mutations occurring at major positions and minor positions in the ECD I-II and kinase (exon 18–21) domains (**e**) in EGFR for NSCLC and GBLGG in primary samples from the total cohort. **b**, **c**, **e**, Two-sided Fisher's exact test. Examined numbers are shown in parentheses.

# Article

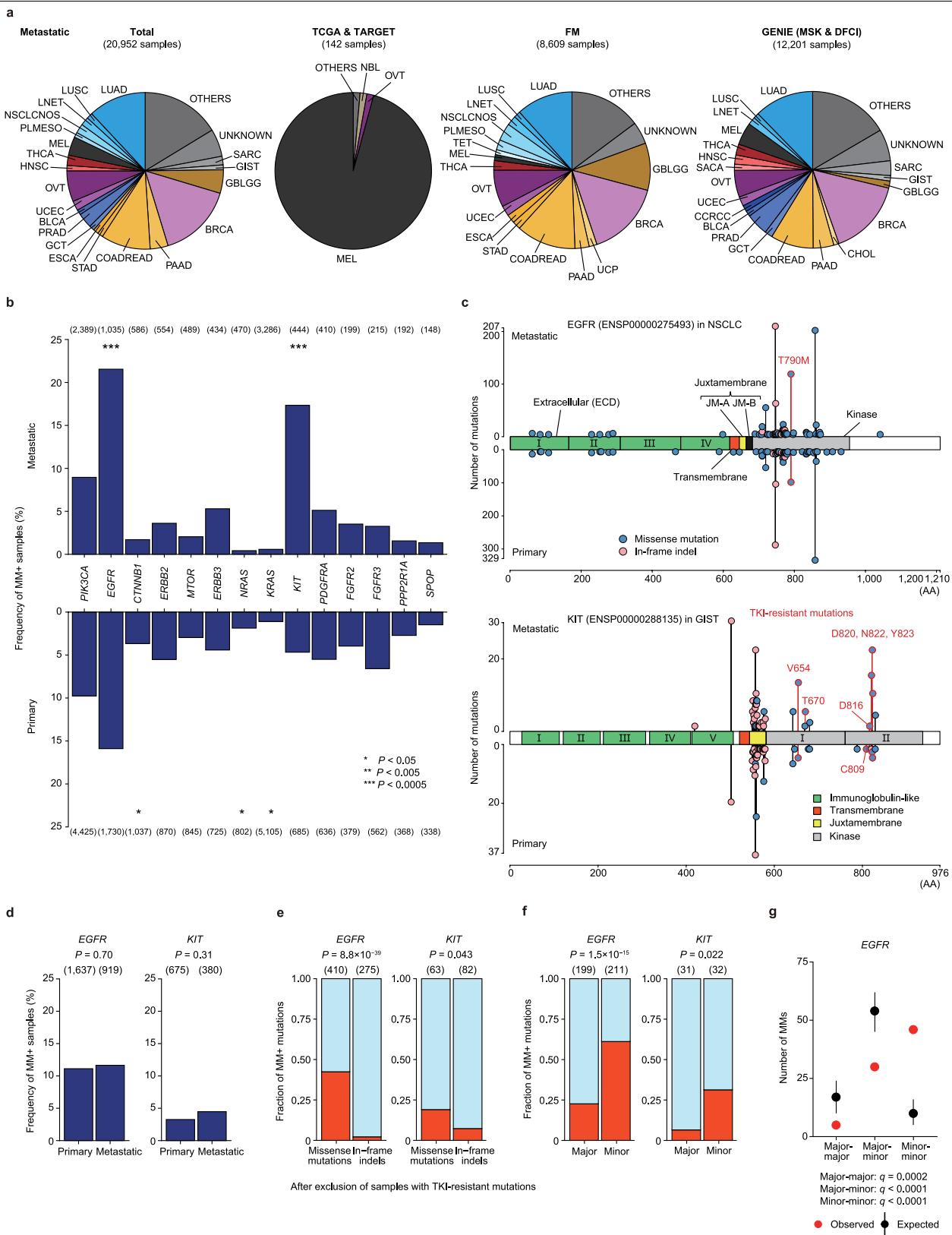


**Extended Data Fig. 8** | See next page for caption.

**Extended Data Fig. 8 | Cancer-type-specific oncogenic MMs.** **a**, Distribution of *FGFR3* mutations in samples with single and multiple mutations in recurrently mutated cancer types (defined as those with 20 or more hotspot/functional mutations) in primary samples from the total cohort. Mutations specific to MM<sup>+</sup> samples are indicated in red. **b**, Cancer-type-specific analysis of MM<sup>+</sup> oncogenes identified by pan-cancer analysis. Shown are numbers of MM<sup>+</sup> samples for MM<sup>+</sup> oncogenes in the corresponding cancer type from the discovery cohort (with  $q < 0.01$  and three or more MMs). Red and black circles indicate, respectively, observed and expected (median with 95% confidence intervals) values. One-sided permutation test ( $n = 10,000$ ) with Benjamini–Hochberg correction. **c**, Proportion of mutations according to type, position and amino-acid change, in six cancer-type-specific MM<sup>+</sup> oncogenes in the corresponding cancer type. Asterisk, AML, and dagger, BALL were analysed in primary samples from the total cohort and (an) independent set(s) of patients<sup>29–33</sup>. **d**, **e**, Distribution of hotspot/functional mutations (**d**) and frequency of MM<sup>+</sup> samples (**e**) in *NOTCH1* for TALL and CLL in primary samples

from the total cohort. An independent cohort<sup>34</sup> was also analysed for CLL. **f**, Significant pairwise associations ( $q < 0.01$ ) with observed/expected ratios among functional domains ( $n = 65$ ). Orange and blue colours depict co-occurring (observed number of MMs significantly higher than expected) and mutually exclusive (lower than expected) associations. Two-sided simulation test ( $n = 10,000$ ) with Benjamini–Hochberg correction. **g**, Fraction of MM<sup>+</sup> mutations for missense mutations and in-frame indels (consisting mainly of internal tandem duplications, ITDs) as well as distribution of mutations and fraction of MM<sup>+</sup> mutations for each hotspot/functional position (the top ten are shown) in *FLT3* for AML in primary samples from the total cohort and an independent set of patients. **h**, Distribution of mutations and fraction of MM<sup>+</sup> mutations for each hotspot/functional position in *AK2* for BALL in primary samples from the total cohort and independent sets of patients. **g**, **h**, Asterisks indicate major positions (in which 10% or more of all mutations were present). **e**, **g**, **h**, Two-sided Fisher's exact test. Examined numbers are shown in parentheses.

# Article

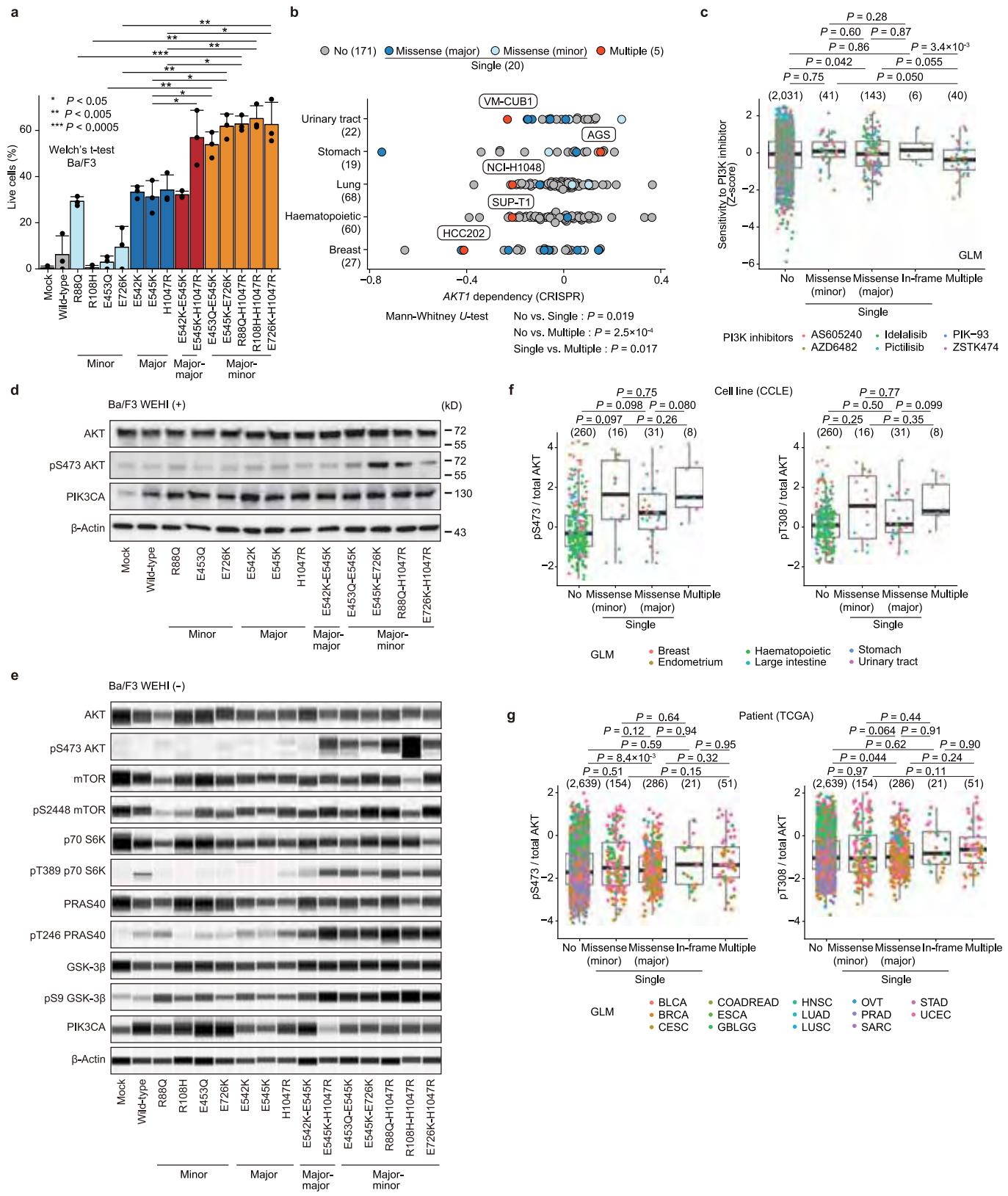


Extended Data Fig. 9 | See next page for caption.

**Extended Data Fig. 9 | Comparison of oncogenic MMs between primary and metastatic samples across cancers.** **a**, Distribution of cancer types for metastatic samples included in the total, TCGA and TARGET, FM and GENIE cohorts. Cancer types with less than 1% frequency were combined as ‘OTHERS’ in each cohort. Cancer-type abbreviations are listed in Supplementary Table 1. **b**, Comparison of the frequency of MM<sup>+</sup> samples in primary and metastatic samples across 14 MM<sup>+</sup> oncogenes in the total cohort. **c**, Distribution of *EGFR* and *KIT* hotspot/functional mutations in primary and metastatic samples in NSCLC and GIST, respectively. Acquired TKI-resistant mutations are indicated in red. **d**, Frequency of MM<sup>+</sup> samples for *EGFR* and *KIT* in primary and metastatic samples. **e,f**, Fraction of MM<sup>+</sup> mutations for missense mutations and in-frame

indels (**e**) and major and minor hotspots (**f**) in *EGFR* and *KIT* in recurrently mutated cancer types in metastatic samples. **b,d–f**, Two-sided Fisher’s exact test. **g**, Number of MMs according to mutational combinations in *EGFR* in recurrently mutated cancer types in metastatic samples ( $n=81$ ). Red and black circles indicate, respectively, observed and expected (median with 95% confidence intervals) values. Two-sided simulation test ( $n=10,000$ ) with Benjamini–Hochberg correction. **d–g**, After exclusion of acquired TKI-resistant mutations, including *EGFR* T790M in NSCLC and *KIT* V654, T670, C809, D816, D820, N822 and Y823 missense mutations in GIST. Examined numbers are shown in parentheses.

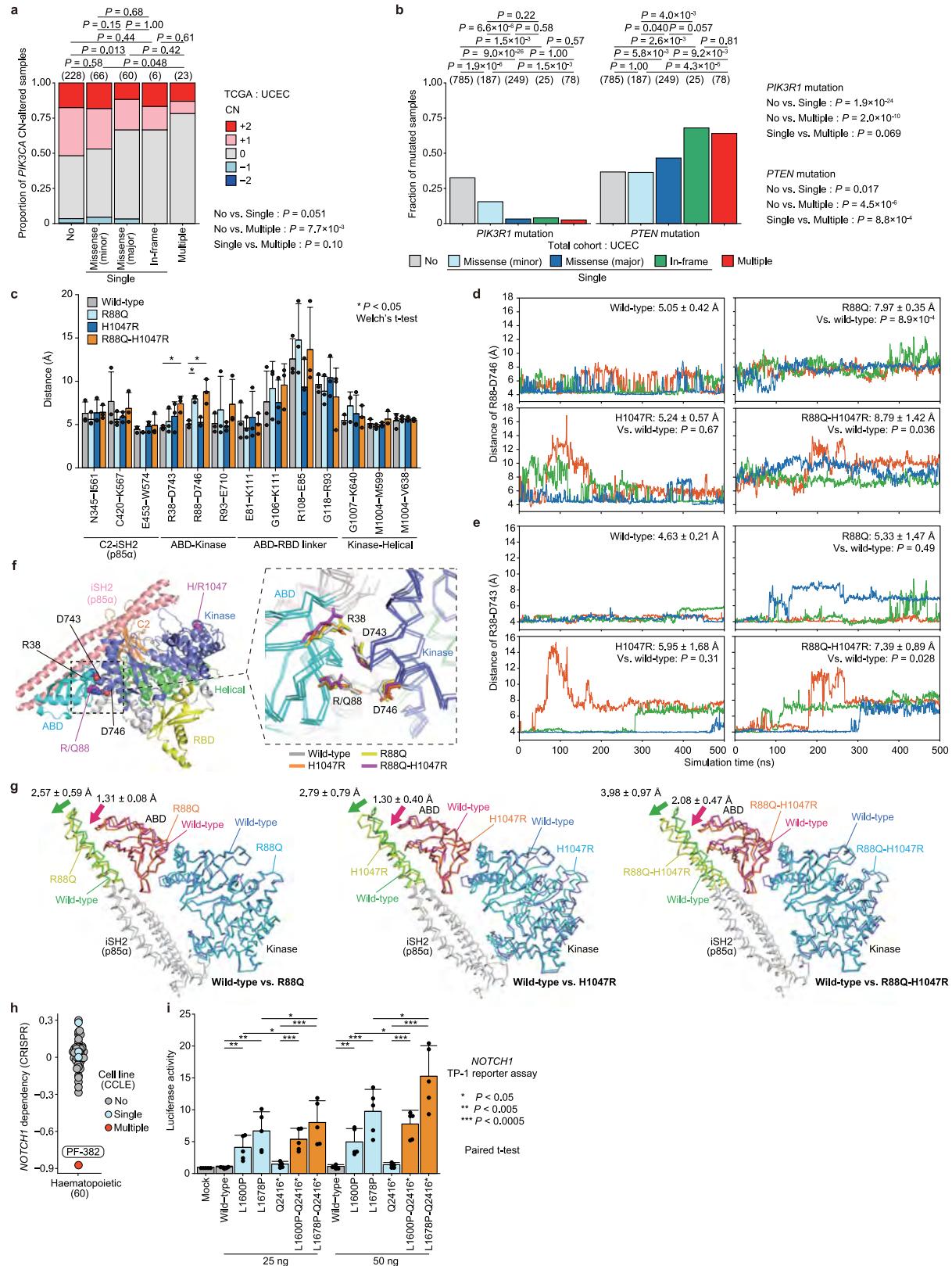
# Article



Extended Data Fig. 10 | See next page for caption.

**Extended Data Fig. 10 | Biological relevance of PIK3CA MMs.** **a**, Viability of Ba/F3 cells expressing mock control or wild-type, single-mutant or double-mutant *PIK3CA* without WEHI-3-conditioned medium for three days ( $n=3$ ). Data represent means + s.d. Two-sided Welch's *t*-test. **b**, Dependency of 196 CCLE cell lines on *AKT1*, with cell lines coloured by *PIK3CA* MM status, in the DepMap CRISPR–Cas9 knockout data (see Methods). Five cancer types for which there are cell lines harbouring *PIK3CA* MMs were analysed by two-sided Mann–Whitney *U*-test. **c**, Box plots showing sensitivity to six PI3K inhibitors (Z-scores) for 417 CCLE cell lines, according to MM status. Five cancer types for which there are cell lines harbouring *PIK3CA* MMs were analysed. **d**, Immunoblot analysis of AKT, pS473 AKT, PIK3CA and β-actin in Ba/F3 cells expressing mock control or wild-type, single-mutant or double-mutant *PIK3CA* in the presence of WEHI-3-conditioned medium. Representative of three independent experiments. **e**, Capillary-based immunoassay of AKT, pS473 AKT, mTOR, pS2448 mTOR, p70S6K, pT389 p70S6K, PRAS40, pT246 PRAS40, GSK-3β, pS9 GSK-3β, PIK3CA and β-actin in Ba/F3 cells expressing mock control or wild-type, single-mutant or double-mutant *PIK3CA* without WEHI-3-conditioned medium. Representative of two independent experiments. **d, e**, See Supplementary Fig. 1 for source images. **f, g**, Box plots showing levels of phosphorylated S473 (pS473; left) and T308 (pT308; right) in RPPA data relative to total AKT protein expression for 316 CCLE cell lines (with cancer types for which there are cell lines harbouring *PIK3CA* MMs) (**f**) and in 3,164 TCGA patients (with purity of 50% or more, in recurrently mutated cancer types) (**g**) according to MM status. Each dot represents a sample, coloured by cancer type. **c, f, g**, Single mutations were classified into missense mutations in major and minor hotspot positions, in-frame indels, and non-hotspot/functional mutations, some of which are not shown owing to small numbers. Box plots show medians (lines), interquartile ranges (IQRs; boxes) and  $1.5 \times$  IQRs (whiskers). GLM, generalized linear model (see Methods). Examined numbers are shown in parentheses.

# Article



**Extended Data Fig. 11** | See next page for caption.

**Extended Data Fig. 11 | Molecular mechanism underlying the enhanced functional activity of oncogenic MMs.** **a**, Fraction of *PIK3CA* copy-number (CN) alterations according to *PIK3CA* MM status for 385 UCEC patients in TCGA. **b**, Fraction of *PIK3R1* and *PTEN* mutations according to *PIK3CA* MM status for 1,339 UCEC patients in the total cohort. **a, b**, Two-sided Fisher's exact test. Single mutations were classified into missense mutations in major and minor hotspot positions, in-frame indels, and non-hotspot/functional mutations, some of which are not shown owing to small numbers. Examined numbers are shown in parentheses. **c**, The distance between key residues of the p110 $\alpha$ /p85 $\alpha$  complex<sup>16</sup>. Data represent means  $\pm$  s.d. (250–500 ns). **d**, Distance between R88 and D746 atoms. **e**, Distance between R38 and D743 atoms. **d, e**, Data represent means  $\pm$  s.d. (250–500 ns). **c–e**, Two-sided Welch's *t*-test. **d, e**, Different colours show independent simulations. **f**, Overall structure of wild-type PIK3CA (left) and a close-up view of the mean structure at the ABD–kinase interface (right). **g**, Mean backbone structures of R88Q, H1047R and R88Q–H1047R mutants, showing the orientation of the ABD, kinase and iSH2 (p85 $\alpha$ ) domains, superimposed on those of the wild type. Movements of the ABD and iSH2 domains induced by each mutant are indicated by arrows, along with the backbone root mean square deviation (r.m.s.d.) of the ABD (residues 16–105; magenta) and iSH2 (490–540; green) domains from the wild-type structure. **c–g**, Molecular-dynamics simulations for wild-type and each mutant PIK3CA (combined results from three independent 500-ns simulations are shown). **h**, Dependency of 60 CCLE cell lines from haematopoietic and lymphoid tissue on *NOTCH1*, with cell lines coloured by MM status, in the DepMap CRISPR–Cas9 knockout data. **i**, Luciferase assays of TP-1 reporter activity in 293T cells transfected with mock control or the indicated amounts of wild-type, single-mutant or double-mutant *NOTCH1* vectors ( $n=5$ ). Data represent means  $\pm$  s.d. Two-sided paired *t*-test on log-transformed values.

Corresponding author(s): Keisuke Kataoka

Last updated by author(s): Feb 5, 2020

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
  - Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted
  - Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

## Data collection

ImageQuant LAS 4000 control software version 1.2  
Molecular Operating Environment (MOE) program version 2016.8  
GROMACS version 2019.1  
Compass software version 4.1.0  
VICTOR Nivo Control Software version 3.0.2

## Data analysis

LiftOver (<https://genome.ucsc.edu/cgi-bin/hgLiftOver>)  
Variant Effect Predictor version 95.3  
vcf2maf version 1.6.16  
SAMtools version 1.4.1  
Minimap2 version 2.14  
NanoQC version 0.9.0  
Integrative Genomics Viewer (IGV) version 2.4.10  
Python version 2.7.15  
R version 3.6.0  
Genomon version 2.6.2  
Custom script for the permutation test (<https://github.com/nccmo/Permutation-test>)  
CisChecker (<https://github.com/nccmo/CisChecker/>)  
DNVChecker (<https://github.com/nccmo/DNVChecker>)  
pyMOL version 2.2.0  
Gnuplot version 4.6.2

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The findings of this study are supported by data that are available from public online repositories and data that are publicly available upon request of the data provider. See Methods for detail. Long-read WGS data of cell lines have been deposited in the European Genome-phenome Archive (EGA) under accession EGAS00001003763. Data generated in the current study are available as Source Data files that accompany Fig.4 and Extended Data Figs.10-11.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](http://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size      No statistical methods were used to predetermine sample size. Sample size was determined by the availability of sequencing data. We enrolled all samples from public datasets.

Data exclusions      Samples with missing data, derived from xenografts, and duplicated in the same patient were excluded from the analysis. Hypermutator samples [with  $\geq 500$  and  $\geq 20$  coding mutations per exome (for the discovery cohort) and targeted region (for the additional cohort), respectively] were also removed from the analysis to minimize the effect of confounding passenger mutations. The specific exclusion criteria were not pre-established, although the cutoff values were determined based on previous publications (example: PMID: 29056346). No data were excluded for in vivo and in vitro experiments.

Replication      No experimental replication was performed for sequencing experiments. Molecular dynamics simulations were performed three times with different velocities. For in vivo and in vitro experiments, all attempts at replication were successful with biological replicates performed on separate cohorts of animals/cells.

Randomization      Samples were assigned to each group based on the number of mutations (samples with no mutation, single mutation, or multiple mutations in the gene of interest). Therefore, randomization was not required.

Blinding      Blinding was not relevant to the study, as there was no control and treatment arms involved.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
	<input type="checkbox"/> <input checked="" type="checkbox"/> ChIP-seq
	<input checked="" type="checkbox"/> <input type="checkbox"/> Flow cytometry
	<input checked="" type="checkbox"/> <input type="checkbox"/> MRI-based neuroimaging

## Antibodies

### Antibodies used

Rabbit anti-PI3 Kinase p110 $\alpha$  (C73F8) (Cell Signaling Technology, #4249, dilution 1:1000 [IB] or 1:50 [CI])  
 Rabbit anti- $\beta$ -Actin (13E5) (Cell Signaling Technology, #4970, dilution 1:2000 [IB] or 1:50 [CI])  
 Rabbit anti-Akt (pan) (C67E7) (Cell Signaling Technology, #4691, dilution 1:1000 [IB] or 1:50 [CI])  
 Rabbit anti-phospho-Akt (Ser473) (D9E) (Cell Signaling Technology, #4060, dilution 1:2000 [IB] or 1:50 [CI])  
 Rabbit anti-mTOR (7C10) (Cell Signaling Technology, #2983, dilution 1:50 [CI])

Rabbit anti-phospho-mTOR (Ser2448) (D9C2) (Cell Signaling Technology, #5536, dilution 1:50 [CI])  
 Rabbit anti-p70 S6 Kinase (Cell Signaling Technology, #9202, dilution 1:50 [CI])  
 Rabbit anti-phospho-p70 S6 Kinase (Thr389) (Cell Signaling Technology, #9205, dilution 1:50 [CI])  
 Rabbit anti-PRAS40 (Cell Signaling Technology, #2610, dilution 1:50 [CI])  
 Rabbit anti-phospho-PRAS40 (Thr246) (C77D7) (Cell Signaling Technology, #2997, dilution 1:50 [CI])  
 Rabbit anti-GSK-3β (27C10) (Cell Signaling Technology, #9315, dilution 1:50 [CI])  
 Rabbit anti-Phospho-GSK-3β (Ser9) (5B3) (Cell Signaling Technology, #9323, dilution 1:50 [CI])  
 IB, immunoblot; CI, capillary-based immunoassay.

## Validation

Antibodies with prior manufacturer validation were used. Validation statements at manufacturer's website were as follows:

Rabbit anti-PI3 Kinase p110α (C73F8) (Cell Signaling Technology, #4249): Western blot analysis of extracts from HeLa cells and neonatal mouse brain using PI3 Kinase p110α (C73F8) Rabbit mAb. Currently over 224 citations.

Rabbit anti-β-Actin (13E5) (Cell Signaling Technology, #4970): Western blot analysis of cell extracts from various cell lines (NIH/3T3, HeLa, PAE, A431) using beta-Actin (13E5) Rabbit mAb. Currently over 1,861 citations.

Rabbit anti-Akt (pan) (C67E7) (Cell Signaling Technology, #4691): Western blot analysis of recombinant Akt1, Akt2 and Akt3 proteins, and extracts from various cell lines (HeLa, NIH/3T3, C6, COS), using Akt (pan) (C67E7) Rabbit mAb. Currently over 1,851 citations.

Rabbit anti-phospho-Akt (Ser473) (D9E) (Cell Signaling Technology, #4060): Western blot analysis of extracts from PC-3 cells, untreated or LY294002/wortmannin-treated, and NIH/3T3 cells, serum-starved or PDGF-treated, using Phospho-Akt (Ser473) (D9E) XP® Rabbit mAb (upper) or Akt (pan) (C67E7) Rabbit mAb #4691 (lower). Currently over 3,926 citations.

Rabbit anti-mTOR (7C10) (Cell Signaling Technology, #2983): Western blot analysis of extracts from 293, A431, COS, C6, and C2C12 cells, using mTOR (7C10) Rabbit mAb. Currently over 970 citations.

Rabbit anti-phospho-mTOR (Ser2448) (D9C2) (Cell Signaling Technology, #5536): Western blot analysis of extracts from serum-starved NIH/3T3 cells, untreated or insulin-treated (150 nM, 5 minutes), alone or in combination with λ-phosphatase, using Phospho-mTOR (Ser2448) (D9C2) XP® Rabbit mAb (upper) or mTOR (7C10) Rabbit mAb #2983. Currently over 650 citations.

Rabbit anti-p70 S6 Kinase (Cell Signaling Technology, #9202): Western blot analysis of extracts from HeLa, NIH/3T3, PC12 and COS-7 cells using p70 S6 Kinase Antibody. Currently over 939 citations.

Rabbit anti-phospho-p70 S6 Kinase (Thr389) (Cell Signaling Technology, #9205): Western blot analysis of HeLa, COS, C6 and 3T3 cells, serum-starved overnight, then treated with insulin, I-phosphatase or 20% serum as indicated. Upper panel probed with Phospho-p70 S6 Kinase (Thr389) Antibody #9205; lower panel probed with p70 S6 Kinase Antibody #9202. Currently over 904 citations.

Rabbit anti-PRAS40 (Cell Signaling Technology, #2610): Western blot analysis of extracts from various cell types (MCF-7, 293, HeLa, A204, RD, 3T3, RAW, KNRK, NBTII, and COS7) using PRAS40 Antibody. Currently over 40 citations.

Rabbit anti-phospho-PRAS40 (Thr246) (C77D7) (Cell Signaling Technology, #2997): Western blot analysis of extracts from serum starved H3255, Mkn45 and NIH/3T3 cells, untreated or treated with either Gefitinib (1 μM, 3 hours), Su11274 (1 μM, 3 hours) or insulin (150 nM, 15 minutes), using Phospho-PRAS40 (Thr246) (C77D7) Rabbit mAb (upper) or PRAS40 (D23C7) Rabbit mAb #2691 (lower). Currently over 119 citations.

Rabbit anti-GSK-3β (27C10) (Cell Signaling Technology, #9315): Western blot analysis of extracts from HeLa, NIH/3T3, COS, C6 and 293 cells using GSK-3β (27C10) Rabbit mAb. Currently over 666 citations.

Rabbit anti-Phospho-GSK-3β (Ser9) (5B3) (Cell Signaling Technology, #9323): Western blot analysis of extracts from NIH/3T3 cells, λ-phosphatase- or PDGF-treated, using Phospho-GSK-3β (Ser9) (5B3) Rabbit mAb (upper) or GSK-3β (27C10) Rabbit mAb #9315 (lower). Currently over 294 citations.

See <https://www.citeab.com/> for details of relevant citations.

## Eukaryotic cell lines

### Policy information about [cell lines](#)

#### Cell line source(s)

Ba/F3-CL1, WEHI-3, and 293T cell lines were obtained from the RIKEN Cell Bank, HEC-1 cell line from the JCRB Cell Bank, and MCF10A, BT-20, NCI-H1048, SUP-T1, and HRT-18 cell lines from ATCC. Lenti-X 293T cells were purchased from TaKaRa.

#### Authentication

All cell lines were authenticated by the providers using karyotype, isoenzymes, and/or microsatellite profiling (short tandem repeat or simple sequence length polymorphism).

#### Mycoplasma contamination

We confirmed that all cell lines were negative for mycoplasma contamination using MycoAlert™ Mycoplasma Detection Kit (Lonza, LT07-318).

#### Commonly misidentified lines (See [ICLAC](#) register)

According to ICLAC register, cross-contamination was reported in BT-20 in 1976 (Reference PubMed ID: 6451928), although authentic stocks apparently do exist. We used the BT-20 cell line authenticated by ATCC using short tandem repeat (STR) profiling analysis. This cell line was selected because of the limited number of available cell lines harboring multiple mutations in PIK3CA.

## Animals and other organisms

### Policy information about [studies involving animals: ARRIVE guidelines](#) recommended for reporting animal research

#### Laboratory animals

Female BALB/c-nu/nu mice (6 weeks old)

#### Wild animals

The study did not involve wild animals.

#### Field-collected samples

The study did not involve samples collected from the field.

Ethics oversight

All mouse experiments were approved by the Animal Ethics Committee of the National Cancer Center and strictly adhered to its guidelines.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Article

# Phase and context shape the function of composite oncogenic mutations

<https://doi.org/10.1038/s41586-020-2315-8>

Received: 7 September 2019

Accepted: 6 April 2020

Published online: 27 May 2020

 Check for updates

Alexander N. Gorelick<sup>1,2</sup>, Francisco J. Sánchez-Rivera<sup>3</sup>, Yanyan Cai<sup>4</sup>, Craig M. Bielski<sup>1,2</sup>, Evan Biederstedt<sup>2</sup>, Philip Jonsson<sup>5</sup>, Allison L. Richards<sup>5</sup>, Neil Vasan<sup>1,6</sup>, Alexander V. Penson<sup>1,2</sup>, Noah D. Friedman<sup>1,2</sup>, Yu-Jui Ho<sup>3</sup>, Timour Baslan<sup>3</sup>, Chaitanya Bandlamudi<sup>5</sup>, Maurizio Scaltriti<sup>4</sup>, Nikolaus Schultz<sup>2,5,7</sup>, Scott W. Lowe<sup>3,8</sup>, Ed Reznik<sup>2,5</sup> & Barry S. Taylor<sup>1,2,5,7</sup>

Cancers develop as a result of driver mutations<sup>1,2</sup> that lead to clonal outgrowth and the evolution of disease<sup>3,4</sup>. The discovery and functional characterization of individual driver mutations are central aims of cancer research, and have elucidated myriad phenotypes<sup>5</sup> and therapeutic vulnerabilities<sup>6</sup>. However, the serial genetic evolution of mutant cancer genes<sup>7,8</sup> and the allelic context in which they arise is poorly understood in both common and rare cancer genes and tumour types. Here we find that nearly one in four human tumours contains a composite mutation of a cancer-associated gene, defined as two or more nonsynonymous somatic mutations in the same gene and tumour. Composite mutations are enriched in specific genes, have an elevated rate of use of less-common hotspot mutations acquired in a chronology driven in part by oncogenic fitness, and arise in an allelic configuration that reflects context-specific selective pressures. *cis*-acting composite mutations are hypermorphic in some genes in which dosage effects predominate (such as *TERT*), whereas they lead to selection of function in other genes (such as *TP53*). Collectively, composite mutations are driver alterations that arise from context- and allele-specific selective pressures that are dependent in part on gene and mutation function, and which lead to complex—often neomorphic—functions of biological and therapeutic importance.

To study the pattern, prevalence and function of composite mutations (hereafter defined as two or more distinct somatic mutations in the same gene and tumour specimen) in cancer, we analysed the germline blood and matched tumour tissue of 31,359 patients with cancer for whom prospective clinical sequencing was performed to guide treatment decisions for advanced and metastatic disease (Fig. 1a, Extended Data Fig. 1a, Supplementary Table 1).

### Selection for composite mutations

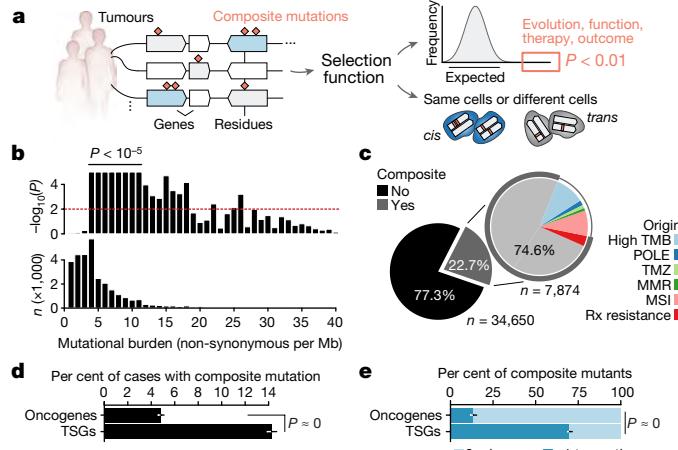
In total, 22.7% ( $n = 7,874$ ) of tumours contained at least one composite mutation—which is 56% more frequent than would be expected by chance, when controlling for gene content and mutational burden ( $P < 10^{-5}$ ) (Extended Data Fig. 1b, c, Methods, Supplementary Table 2). Significantly more composite mutations arose than would be expected in cases of modest mutational burden (4–11 mutations per megabase (Mb), about 44% of all tumours,  $P < 10^{-5}$ ) (Fig. 1b, Extended Data Fig. 1d), an enrichment that decreased in tumours of increasing mutational burden. As positive selection cannot be easily distinguished from the predominantly neutral effect of increasing mutational burden, tumours with a high mutational burden were considered to be biologically distinct and were excluded from analysis (Fig. 1c, Methods).

Finally, we also found that known mechanisms of localized hypermutation explain few composite mutations overall (Extended Data Fig. 2).

Composite mutations in tumour-suppressor genes affected a greater proportion of cases than those in oncogenes (14.2% versus 4.8% of all cases;  $P < 10^{-308}$ , two-sided McNemar's test) (Fig. 1d). Furthermore, 70% of composite mutations in tumour-suppressor genes consisted of one or more truncating variants, compared to only 13% for oncogenes (Fig. 1e); this suggests that biallelic loss drives the enrichment for composite mutations in tumour-suppressor genes. Lineage-specific patterns of driver mutations in individual cancer genes were, in part, reflected in the pattern of composite mutations (Fig. 2a, Extended Data Fig. 3a). This included a higher burden of composite mutations in *PIK3CA* in breast cancers, *APC* in colorectal cancers, *CDK12* in prostate cancers and *EGFR* in both lung cancers and gliomas, among others. By contrast, not all frequently mutated genes—such as *KRAS* in multiple cancers or *VHL* in renal cell carcinomas—had frequent composite mutations, which often evolve serial genetic changes by other means (such as allelic imbalance and/or loss of heterozygosity).

We next sought to determine whether individual cancer genes were enriched or depleted for composite mutations, controlling for the determinants of their background mutation rate<sup>9</sup> (Methods). In total, 34 genes were significantly enriched for composite mutations ( $Q < 0.01$ )

<sup>1</sup>Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>2</sup>Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>3</sup>Cancer Biology and Genetics Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>4</sup>Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>5</sup>Marie-Josée and Henry R. Kravis Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>6</sup>Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>7</sup>Weill Cornell Medical College, New York, NY, USA. <sup>8</sup>Howard Hughes Medical Institute, New York, NY, USA. <sup>✉</sup>e-mail: reznike@mskcc.org; taylorb@mskcc.org



**Fig. 1 | Composite mutations in human cancers.** **a**, Schematic of the discovery and characterization of composite mutations. **b**, Top, statistically significant enrichment ( $P < 10^{-5}$ ) for composite mutations in tumours of increasing tumour mutational burden. Nominal  $P$  based on one-sided permutation tests for enrichment (100,000 permutations) applied independently to the subset of tumours with each indicated tumour mutational burden (bottom, number of cases),  $n = 30,505$  biologically independent tumour samples with tumour mutational burden  $\leq 40$  nonsynonymous exonic mutations per Mb. **c**, Proportion of composite mutations including the fraction ascribed to mutational processes associated with hypermutation. MMR, mismatch repair; MSI, microsatellite instability; POLE, DNA-polymerase- $\epsilon$ -associated hypermutation; Rx resistance, acquired resistance to therapy; TMB, tumour mutational burden; TMZ, temozolamide-associated hypermutation; cases excluded from analysis unless otherwise noted. **d**, Percentage of cases with composite mutations by cancer gene function.  $P < 10^{-308}$  (numeric limit, two-sided McNemar's test;  $n = 29,507$  patients). TSG, tumour-suppressor gene. **e**, Types of composite mutations by cancer-gene function ( $P < 10^{-308}$ , numeric limit, two-sided Fisher's exact test;  $n = 5,954$  composite mutations). Error bars in **d**, **e** are 95% binomial confidence intervals.

(Fig. 2b, Supplementary Table 3), including both tumour-suppressor genes (such as *APC*, *TP53*, *PTEN* and *MAP3K1*) and oncogenes, the most significant of which was *PIK3CA* (9.9% of all mutations in *PIK3CA* were composite, 95% confidence interval 9.0–10.9) (Extended Data Fig. 3b). Other frequently mutated oncogenes were not enriched for composite mutations; these included *IDH1*, which reflects the requirement for heterozygosity in IDH-mutant cells to sustain adequate production of D-2-hydroxyglutarate<sup>10</sup>, and *KRAS*, which may reflect selection against further detrimental oncogenic RAS activation<sup>8,11</sup>. Mutational recurrence alone cannot, therefore, predict whether a cancer gene is enriched for composite mutations.

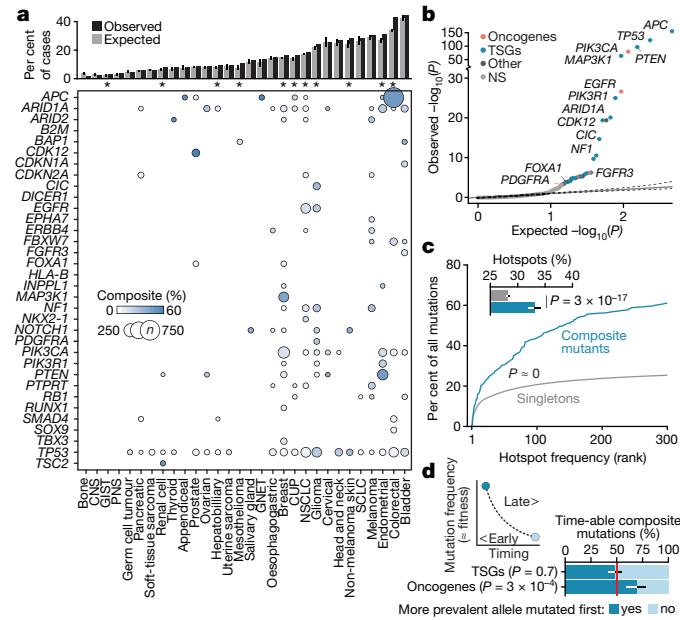
Consistent with their selection, composite mutants were 2.5-fold more likely than individual mutations to include a hotspot–residues that are mutated in cancer more often than would be expected in the absence of selection<sup>12,13</sup> ( $P < 10^{-308}$ , two-sample Z-test for equal proportion) (Fig. 2c). Composite mutations notably lacked the hotspots of greatest positive selection (for example, *KRAS*<sup>G12</sup> and *BRAF*<sup>V600</sup>) but were instead prevalent among less common hotspots, which suggests that the selective pressure is greatest for weakly functional alleles. On the basis of differences in their clonality, in 69% of cases the more prevalent hotspot mutation (at the population level) preceded the less prevalent mutation in oncogenes (95% confidence interval 59–78%) (Fig. 2d), consistent with a model in which the less prevalent allele synergizes with a more-potent initial hotspot mutation. Tumour-suppressor genes exhibited no such temporal ordering, which reflects how prevalence is poorly correlated with fitness for predominantly loss-of-function mutations. Together, these data indicate a strong mutant-allele-specific selective pressure for composite mutations that evolve along a chronology driven in part by oncogenic fitness.

## Phase and function

The elevated rate of likely driver mutations in composite mutants led us to investigate their allelic configuration. We combined sequencing read support with clonality to phase mutations, and thereby ensured that composite mutations arose in the same tumour cell population. Among evaluable composite mutants, 67% and 19% ( $n = 977$  and 275) arose in *cis* (on the same allele) and *in trans* (on different alleles), respectively, and 14% ( $n = 210$ ) were indeterminate. In part, the higher rate of *cis* mutants reflected reduced sensitivity for detecting *trans* mutations from the short-read sequencing used here, an effect we controlled for in subsequent analyses (Methods). Tumour-suppressor genes were substantially more likely to contain composite mutations *in trans*, especially those with two truncating mutations that were consistent with biallelic inactivation (71% *in trans*,  $n = 79$  of 111). By contrast, composite-mutant oncogenes with two missense mutations were largely *cis*-acting (91%,  $n = 243$  of 268;  $P = 3 \times 10^{-33}$ , two-sided Fisher's exact test) (Fig. 3a). Composite mutations that involved silent mutations exhibited no such difference in phase among these genes, which suggests that the *cis*-mutant enrichment in oncogenes reflects selective pressure. Notably, although not precluding resistance *in trans*<sup>14</sup>, all the secondary resistance mutations that we identified arose *in cis*<sup>15–17</sup> ( $n = 18$ ;  $P = 0.02$ , two-sided Fisher's test) (Fig. 3b, Extended Data Fig. 4), which suggests that exogenous selective pressures drive—in part—the phase of composite mutations.

Despite these patterns, extensive variability existed in the phase of composite mutations in individual cancer genes (Fig. 3c). *EGFR*, *TERT* and *PIK3CA* had the highest percentage of *cis* composite mutations among oncogenes (88–97%). Prevalent *cis*-acting composite mutations were observed even among canonical tumour-suppressor genes, comprising 77.1% of all composite mutations in these genes. Here, *TP53* was notable: 43% of all phase-able composite mutations ( $n = 70$  of 163) in this gene were *cis*-acting, and enriched in a cluster of residues near the C-terminal end of the DNA binding domain (E287, E285, E271 and R280) (Fig. 3d). Although short-read sequencing technologies restrict phasing to variants within close physical proximity and potentially overestimate the prevalence of *cis* mutations, these data are nevertheless inconsistent with conventional loss of function via biallelic inactivation and may suggest a broader functional effect of composite mutations in *TP53* and other tumour-suppressor genes.

To assess the phenotypic consequence of *cis*-acting composite mutations in the DNA binding domain of *TP53*, we developed an isogenic system for acute reconstitution of *TP53*. As E287D was the most significant mutated residue enriched in composite mutants, we focused on a representative *TP53*<sup>R280T/E287D</sup> *cis* composite mutant. To model the effect of this composite mutation in the lineage of affected tumours, we transduced *Kras*<sup>G12D</sup> *Trp53*<sup>−/−</sup> mouse lung cancer cells with GFP-labelled retroviral constructs that encode complementary (c)DNAs for wild-type *Trp53*, *Trp53*<sup>R277T</sup>, *Trp53*<sup>E284D</sup> and *cis* *Trp53*<sup>R277T/E284D</sup> (orthologous to human wild-type *TP53*, *TP53*<sup>R280T</sup>, *TP53*<sup>E287D</sup> and *cis* *TP53*<sup>R280T/E287D</sup>, respectively), after which we selected GFP-expressing cells and performed RNA sequencing (Fig. 3e, Extended Data Fig. 5a, Supplementary Table 5). *Trp53* mRNA expression was stable and robust, whereas in *Trp53*<sup>−/−</sup>, *Trp53*<sup>R277T</sup> and *Trp53*<sup>R277T/E284D</sup> cells there was a decrease in p21 (encoded by *Cdkn1a*) induction, a surrogate marker of p53 functionality (Extended Data Fig. 5b, c). *Trp53*<sup>E284D</sup> cells transcriptionally resembled *Trp53*<sup>+/+</sup> cells, whereas *Trp53*<sup>R277T</sup> cells resembled *Trp53*<sup>−/−</sup> cells (Extended Data Fig. 5d). By contrast, *Trp53*<sup>R277T/E284D</sup> cells had a mixed transcriptional phenotype, bearing a dominant differential expression signature that was equivalent to the one induced in either *Trp53*<sup>R277T</sup> or *Trp53*<sup>−/−</sup> cells while retaining a *Trp53*<sup>E284D</sup>-like downregulation of the AP-1 transcription factor program (Fig. 3f, Extended Data Fig. 5e). These data correlated with human tumour genomics, in which the null-like *TP53*<sup>R280T</sup> mutation was common but *TP53*<sup>E287D</sup> mutation was rare and nearly always arose as a composite mutation (Extended Data Fig. 5f).

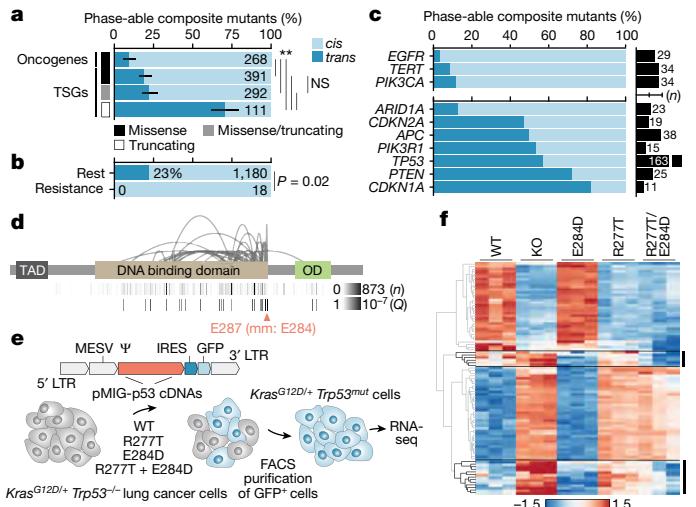


**Fig. 2 | Gene- and residue-specific selective pressure for composite mutations.** **a**, Prevalence of composite mutations by affected gene and lineage (cancer types of  $\geq 100$  and  $\geq 5$  total and composite-mutant cases,  $n = 31,563$  samples). Top, percentage of cases with composite mutations, and the expected value on the basis of cohort size and mutational burden. Expected values are the mean percentage of 10,000 random permutations for each lineage. Asterisks denote cancer types with a significantly greater proportion of affected cases than expected by chance (false-discovery-rate (FDR)-adjusted  $P < 0.01$ ). Bars are 95% confidence intervals. CNS, central nervous system; CUP, cancer of unknown primary; GIST, gastrointestinal stromal tumour; GNET, gastrointestinal neuroendocrine tumour; NSCLC, non-small-cell lung cancer; PNS, peripheral nervous system; SCLC, small-cell lung cancer. **b**, The significance of enrichment for composite mutations in cancer genes (FDR-adjusted  $P$  values from one-sided binomial test for enrichment,  $n = 26,997$ ; light grey is not significant (NS)). **c**, Hotspot mutation use among composite and singleton mutations by decreasing population-level frequency ( $P < 10^{-308}$ , numeric limit, two-sided Mann–Whitney  $U$ -test,  $n = 93,616$  and 2,920 singleton and composite missense mutations, respectively, in 25,037 patients). Inset, the percentage of all missense mutations, comprising composite and singleton mutants that arose at individually significant mutational hotspots.  $P$ , two-sided two-sample Z-test for equal proportions,  $n = 105,297$  total single-nucleotide variants, error bars are 95% binomial confidence intervals. **d**, Right and left are the proposed and observed temporal order, respectively, of the acquisition of two functional variants in composite mutations in oncogenes (from mutation clonality). Tumour-suppressor genes shown as a negative control.  $P$ , two-sided binomial test, error bars in all panels are 95% binomial confidence intervals ( $n = 336$  evaluable composite mutations).

A second *cis*-acting composite mutant ( $Trp53^{R277K/E282K}$ , orthologous to human  $TP53^{R280K/E285K}$ ) also promoted a transcriptional program distinct from its constituent mutations (Extended Data Fig. 5g). Importantly, the  $TP53^{R277T/E284D}$  mutation was not associated with increased growth in vitro or survival in vivo compared to the individual mutations (Extended Data Fig. 5h, i). Collectively, these data suggest that *cis*-acting  $TP53$  composite mutations tune mutant p53 transcriptional phenotypes, which leads to a selection of function that is absent from null-like single  $TP53$  mutations.

## Conditionally dependent mutant alleles

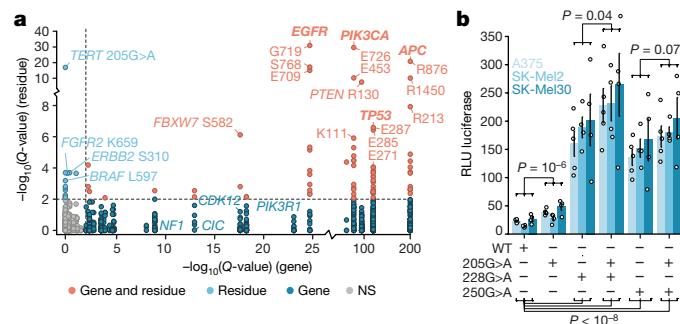
The residue-specific transcriptional phenotypes of  $TP53$ -composite mutants suggest broader allele-specific selection among composite mutations. We therefore identified individual alleles that exhibit an



**Fig. 3 | *cis*- and *trans*-acting composite mutants.** **a–c**, The phase of composite mutations by their type and affected cancer gene ( $P$  for asterisked comparisons from left to right are  $4 \times 10^{-4}$ ,  $2 \times 10^{-5}$ ,  $3 \times 10^{-33}$ ,  $8 \times 10^{-24}$  and  $3 \times 10^{-19}$ , and not significant was 0.3, two-sided Fisher's exact test,  $n = 1,062$  evaluable composite mutations; error bars are 95% binomial confidence intervals) (**a**); association or not with acquired therapy resistance ( $P$ , two-sided Fisher's exact test,  $n = 1,198$  evaluable composite mutations) (**b**); and affected individual oncogenes and tumour-suppressor genes (top and bottom, respectively, known and predicted functional mutations in  $\geq 10$  phase-able tumours; number of cases with phase-able composite mutations as indicated) (**c**). **d**, The pattern of  $TP53$  composite mutations with arcing lines indicating the position of pairs of mutations in  $\geq 2$  tumours; height corresponds to recurrence. At the bottom, the number of mutated cases at each individual residue and the  $Q$  of significance (FDR-adjusted  $P$  value from one-sided binomial test) for each residue as arising in composite are shown. TAD, transactivation domain; OD, oligomerization domain. mm, mouse. **e**, Schematic of the experimental workflow for generating isogenic cells for phenotypic comparison of  $TP53$  mutations. **f**, Heat map of the top 30 differentially expressed genes between  $Trp53^{R277T}$ ,  $Trp53^{E284D}$ , and  $Trp53^{R277T/E284D}$ -mutant cells.

excess of composite mutations (Methods). In total, 86 mutant residues in 24 cancer genes were enriched for arising as composite variants ( $Q < 0.01$ ) (Fig. 4a, Supplementary Table 4). Nearly 70% of these mutations occurred in only 4 genes ( $TP53$ ,  $PIK3CA$ ,  $APC$  and  $EGFR$ ), with few reaching saturation for discovery at the current cohort size, and 56% also arising as individually significant hotspot mutations<sup>13</sup> (Fig. 2b, Extended Data Fig. 6). As with  $TP53$ , several tumour-suppressor genes had mutant-allele-specific enrichment that may suggest selection for something other than conventional loss of function. In  $PIK3CA$ , mutations that are enriched in composite mutants (in residues E726, E453, K111, R108 and R93) were nearly always *in cis* when phase-able, and often arose through APOBEC-associated mutagenesis (Extended Data Fig. 7). Notably, composite  $PIK3CA$  mutations drive elevated PI3K activity, downstream signalling, cell proliferation and tumour growth, and may increase sensitivity to PI3K inhibitors<sup>18</sup>, confirming that—in addition to introducing passenger mutations—APOBEC and other mutational processes create numerous functional driver mutations.

Multiple significant residues appeared to be conditional alleles—rarely arising without a second *cis* activating mutation (Extended Data Fig. 8a). Among these were  $EGFR$ -mutant residues (E709, V834 and L833)<sup>19</sup> and the  $TERT$  promoter mutation 205G>A (Fig. 4a).  $TERT$  promoter mutations are common in human cancer<sup>20</sup> and create novel GABPA binding sites that promote aberrant telomerase activity<sup>21</sup>. The 205G>A mutation was the sixth most common  $TERT$  promoter mutant, and exclusively arose *in cis* ( $n = 13$  of 13) with either the highly prevalent 228G>A or 250G>A hotspots, which—despite their frequency—were never together in composite (Extended Data Fig. 8b). To test whether



**Fig. 4 | Mutant-allele-specific enrichment for composite mutations.** **a,** Enrichment significance of individual mutant residues arising in composite mutations ( $n=1,821$  distinct mutant sites tested;  $n=155,241$  variants overall) compared to significance of composite enrichment among genes ( $Q$  for mutant sites is FDR-adjusted one-sided Fisher's exact test; for  $Q$  for genes, refer to Fig. 2b). Genes in bold label each of the residues beneath. **b,** The degree of *TERT* expression induced by transient transfection of the indicated mutations individually, or as *cis* composite, in three melanoma cell lines (A375, SK-Mel2 and SK-Mel30). The mean and s.e.m. (error bars) across  $n=4$  or 5 replicates per allele.  $P$ , two-way analysis of variance assessing expression as a function of genotype and baseline expression of each cell line (Methods); at the bottom,  $P<10^{-8}$  values from left to right are  $3\times10^{-9}$ ,  $1\times10^{-9}$ ,  $2\times10^{-9}$  and  $2\times10^{-11}$ . RLU, relative luminescence unit.

the 205G>A mutation synergizes with existing promoter mutations to enhance *TERT* expression, we expressed constructs with a luciferase reporter engineered to contain various *TERT* promoter mutations alone or as *cis*-composite mutants in three melanoma cell lines (A375, SK-Mel2 and SK-Mel30). *TERT*<sup>205G>A</sup> induced modest *TERT* expression compared to wild type, but less than *TERT*<sup>228G>A</sup> or *TERT*<sup>250G>A</sup> alone. Consistently, *TERT*<sup>205G>A</sup> creates a novel motif to which GABPA binds with lower affinity than those motifs created by canonical *TERT* hotspots (Extended Data Fig. 8c). The selective pressure for *TERT*<sup>205G>A</sup> is therefore probably based on the cooperativity of tandem motifs generated by this mutation, and canonical promoter hotspots bound by GABPA heterotetramer complexes<sup>21</sup>. When expressing *TERT*<sup>205G>A</sup> as a *cis* composite with either *TERT*<sup>228G>A</sup> or *TERT*<sup>250G>A</sup> (thereby modelling the 205G>A-mutant human tumours), *TERT* expression increased relative to either mutation alone (Fig. 4b). These data suggest that 205G>A is hypermorphic, driving modestly elevated *TERT* expression that is weakly selected for and therefore does not arise as an individual hotspot mutation but is instead a conditionally dependent composite allele.

Our results indicate that composite mutations are driver alterations with a selective advantage that appears to be primarily determined by their allelic configuration and context. No single model explains the context-dependent phenotypic consequences of composite mutations. In some cancer genes with dosage-dependent function, *cis*-acting composite mutants are additive and arise predominantly in weakly oncogenic alleles and genes (for example, *PIK3CA*<sup>22–24</sup>). This suggests an evolutionary model in which the second mutation arises through selection for hypermorphic activity beyond the level sufficient for activation by the first allele. In genes (such as *TP53*) with manifold phenotypic consequences, *cis* mutants seem to drive functional innovation.

With these mutations, the evolutionary advantage consistent with our results is via tuning of the subtle phenotypic differences that are conferred by the asymmetric combination of the output of individual mutations. Mutant cancer genes must ultimately be considered—both biologically and clinically—in their allelic context, with implications for our understanding of cancer gene function, malignant phenotypes and therapy.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2315-8>.

- Vogelstein, B. et al. Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).
- Garraway, L. A. & Lander, E. S. Lessons from the cancer genome. *Cell* **153**, 17–37 (2013).
- Cairns, J. Mutation selection and the natural history of cancer. *Nature* **255**, 197–200 (1975).
- Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
- Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
- Hyman, D. M., Taylor, B. S. & Baselga, J. Implementing genome-driven oncology. *Cell* **168**, 584–599 (2017).
- Knudson, A. G., Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl Acad. Sci. USA* **68**, 820–823 (1971).
- Bielski, C. M. et al. Widespread selection for oncogenic mutant allele imbalance in cancer. *Cancer Cell* **34**, 852–862.e4 (2018).
- Lawrence, M. S. et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
- Jin, G. et al. Disruption of wild-type IDH1 suppresses d-2-hydroxyglutarate production in IDH1-mutated gliomas. *Cancer Res.* **73**, 496–501 (2013).
- Mueller, S. et al. Evolutionary routes and KRAS dosage define pancreatic cancer phenotypes. *Nature* **554**, 62–68 (2018).
- Chang, M. T. et al. Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity. *Nat. Biotechnol.* **34**, 155–163 (2016).
- Chang, M. T. et al. Accelerating discovery of functional mutant alleles in cancer. *Cancer Discov.* **8**, 174–183 (2018).
- Intlekofer, A. M. et al. Acquired resistance to IDH inhibition through *trans* or *cis* dimer-interface mutations. *Nature* **559**, 125–129 (2018).
- Hidaka, N. et al. Most T790M mutations are present on the same *EGFR* allele as activating mutations in patients with non-small cell lung cancer. *Lung Cancer* **108**, 75–82 (2017).
- Gainor, J. F. et al. Molecular mechanisms of resistance to first- and second-generation ALK inhibitors in ALK-rearranged lung cancer. *Cancer Discov.* **6**, 1118–1133 (2016).
- Kobayashi, S. et al. *EGFR* mutation and resistance of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **352**, 786–792 (2005).
- Vasan, N. et al. Double *PIK3CA* mutations in *cis* increase oncogenicity and sensitivity to PI3Kα inhibitors. *Science* **366**, 714–723 (2019).
- Chen, Z. et al. *EGFR* somatic doublets in lung cancer are frequent and generally arise from a pair of driver mutations uncommonly seen as singlet mutations: one-third of doublets occur at five pairs of amino acids. *Oncogene* **27**, 4336–4343 (2008).
- Huang, F. W. et al. Highly recurrent *TERT* promoter mutations in human melanoma. *Science* **339**, 957–959 (2013).
- Bell, R. J. A. et al. The transcription factor GABP selectively binds and activates the mutant *TERT* promoter in cancer. *Science* **348**, 1036–1039 (2015).
- Berenjeno, I. M. et al. Oncogenic *PIK3CA* induces centrosome amplification and tolerance to genome doubling. *Nat. Commun.* **8**, 1773 (2017).
- Kinross, K. M. et al. An activating *Pik3ca* mutation coupled with *Pten* loss is sufficient to initiate ovarian tumorigenesis in mice. *J. Clin. Invest.* **122**, 553–557 (2012).
- Madsen, R. R. et al. Oncogenic *PIK3CA* promotes cellular stemness in an allele dose-dependent manner. *Proc. Natl Acad. Sci. USA* **116**, 8380–8389 (2019).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

### Prospective sequencing cohort

Somatic mutation data consisted of 34,650 tumour and matched normal specimens from 31,359 patients with prospectively characterized solid cancers. All patients provided written informed consent and were prospectively sequenced as part of their active care at Memorial Sloan Kettering Cancer Center (MSKCC) between January 2014 and April 2019 as part of an Institutional-Review-Board-approved research protocol (NCT01775072). Details of patient consent, sample acquisition, sequencing and mutational analysis have previously been published<sup>25,26</sup>. In brief, matched tumour and blood specimens for each patient were sequenced using MSK-IMPACT, a custom hybridization capture-based next-generation sequencing assay. All samples were sequenced with one of three incrementally larger versions of the assay encompassing 341, 410, and 468 cancer-associated genes, respectively. The study cohort consisted of tumours samples with one of 429 distinct subtypes of cancer. For the purposes of grouping histological subtypes into primary cancer diagnosis, we used the OncoTree structured classification of disease (<http://oncotree.mskcc.org>). Histologic subtypes of fewer than 50 tumour samples were aggregated into a miscellaneous category and nonsolid tumour types were excluded from the study cohort (as well as from analyses of The Cancer Genome Atlas (TCGA) data), resulting in a final cohort of 41 distinct types of tumour.

### Mutational data and annotation

Somatic nonsynonymous substitutions and small insertions and deletions (indels) were identified with a clinically validated pipeline, as previously described<sup>26,27</sup>. Each mutation was classified as probably functional if it was previously reported as a mutational hotspot<sup>12,13</sup>. Truncating variants were considered probably functional if they arose in known tumour-suppressor genes, on the basis of gene function curated by OncoKB<sup>28</sup>. Finally, any additional somatic mutations that did not satisfy the aforementioned criteria were similarly annotated as probably functional if previously curated via literature mining by OncoKB as oncogenic, probably oncogenic or predicted to be oncogenic<sup>28</sup>.

For all composite mutants in which one or both mutations were a known therapeutic target or known resistance mutation as defined by OncoKB levels 1 to 4, R1 or R2 alterations (annotation as of April 2019), each mutation was manually reviewed and classified as a likely resistance mutation on the basis of the cancer type of the affected tumour sample, the existence of known resistance mutations to commonly used targeted therapies indicated for the given cancer type and—if available—review of the clinical histories of affected patients. Composite mutations in which one mutation was an established second-site mutation (for example, *EGFR<sup>T790M</sup>* in non-small cell lung cancer<sup>17</sup> and *AR* mutations in prostate cancer that mediate resistance to anti-androgen therapy) were always classified as resistance mutations. Notably, composite mutations in only 3.4% of cases in this advanced and post-treatment cohort have been associated with therapy resistance, indicating that prior therapy exposure alone cannot explain their prevalence. However, as detailed clinical histories including previous lines of treatment and response phenotypes were not available for all patients, a small number of composite mutations are probably misclassified as non-resistance-associated.

### Mutational burden classification

Tumour samples were classified as hypermutated if they contained either MSI MMR deficiency, *POLE*-mediated ultra-mutation, or TMZ-induced hypermutation<sup>29</sup>. MSI was considered present for any tumour with an MSISensor<sup>30</sup> score of greater than or equal to 10, as

previously clinically validated<sup>31</sup>. Tumour samples with *POLE*, MMR and TMZ-induced hypermutation were identified by mutational signature decomposition analysis. In brief, in each tumour specimen with 20 or more substitutions, the proportion of mutations attributable to each of 30 known somatic mutational signatures were calculated on the basis of a basin-hopping algorithm (<https://github.com/mskcc/mutation-signatures>)<sup>32</sup>. This method uses the distribution of 96 unique trinucleotides generated by 6 possible C- or T-normalized single-nucleotide substitutions (that is, C>A, C>G, C>T, T>A, T>C or T>G) and their 5'- and 3'-adjacent bases to estimate the fraction of mutations attributed to each mutational signature in each specimen. Tumour specimens for which at least 20% of its substitutions were attributed to *POLE* (signatures 10 or 14), TMZ (signature 11), or MMR (signatures 6, 15, 20, 21 or 26) were classified as hypermutated.

To classify tumour specimens with a high mutational burden compared to the majority of cancers of that type, but that otherwise lack one of these known mechanisms of hypermutation, we performed in each individual cancer type of greater than 50 tumour specimens one-dimensional *k*-means clustering of the mutational burden of all tumours (nonsynonymous exonic mutations per Mb). Between 1 and 9 clusters were inferred to best describe the distribution of mutational burden per cancer type. The cluster of lowest mutational burden centred at 20+ mutations per Mb and accounting for <10% of the samples in tumour type established the threshold for high mutational burden, and all tumour specimens in this cluster or those clusters with higher mutational burden were considered to be of high mutational burden.

### Composite mutation identification and annotation

For the purposes of this analysis, a composite mutation was the occurrence of two or more somatic mutations to the same gene within a single sequenced tumour specimen. Carriers of pathogenic germline variants with a second somatic mutation were not considered here. We identified composite mutations as arising owing to somatic hypermutation or high mutational burden of unknown aetiology (as defined in ‘Mutational burden classification’), or a mechanism of resistance to targeted therapy per the aforementioned annotation (‘Mutational data and annotation’) in nonhypermutated tumours. Any composite mutation arising in a hypermutated tumour was considered separately, and excluded from primary analyses unless otherwise noted. All composite mutations that did not meet these criteria were analysed further.

### Testing of population-, gene- and residue-specific composite mutation enrichment

Multiple somatic mutations will accumulate in a gene in the absence of selection at a rate that correlates with the mutational burden and mutational mechanisms of a given tumour. Using a permutation-based framework, we simulated the burden of composite mutations for a given tumour mutation burden. In brief, the true number of tumour specimens containing a composite mutation was calculated ( $n^{\text{true}}$ ). We assembled an  $m \times 2$  matrix, in which  $m$  is the total number of nonsynonymous somatic mutations in our cohort. Each row in the matrix identified the sample and the gene in which a particular mutation arose. We constructed a null distribution by randomly permuting the second column of this matrix 100,000 times, thereby preserving the mutation burden of each gene and each tumour specimen. Upon each iteration, the number of tumour specimens containing a composite mutation was reassessed. An empirical  $P$  value was calculated as the fraction of permutations satisfying  $n_i \geq n^{\text{true}}$ . We used the same procedure for assessing the enrichment of composite mutations for tumour samples in ranges of specific mutational burdens.

To test for enrichment or depletion for composite mutations within cancer types (in cancer types with more than 50 profiled tumours), we used a modified permutation analysis controlling for the underlying gene-specific tendency for mutated genes within each cancer type to contain a composite. To do so, we defined a mutation event to be

## Article

a tumour-sample-mutated gene tuple. A mutation event ( $s, g$ ) occurs when a tumour sample  $s$  was found to contain one or more mutations to a gene,  $g$ . Then, we implemented a permutation analysis that shuffles mutations across samples in a manner that preserves (1) gene mutation burden, (2) tumour sample mutation burden and (3) the total number of mutation events that were observed in that cancer type using the permatswap function in the R package vegan<sup>33</sup>. This final constraint enforces that the number of non-zero entries in the mutation event matrix (the binary matrix of patients and genes) remains constant for each permutation. This constraint is particularly relevant in cancer types that have a mutation burden that is dominated by genes that are depleted of composite mutations (for example, KRAS in pancreatic cancer and BRAF or KRAS in thyroid cancer).

We evaluated the enrichment of composite mutations in each gene by modelling composite mutation burden as a function of genomic covariates, testing the likelihood of the observed number of composite mutations (corresponding to the probability of observing this burden of composite mutations by chance) using a binomial test. To parametrize  $\hat{p}$  (the background rate of composite mutations in the absence of selection for each gene  $g$ ), we estimated the expected number of composite mutated samples in a gene  $\hat{n}_c$  from the total number of samples with an observed mutation in the gene  $n_s$ , such that  $\hat{p}^g = \hat{n}_c^g/n_s^g$ . Dropping the superscript for clarity,  $\hat{n}_c$  was estimated for each gene using negative binomial regression to model the observed number of composite-mutant samples in a gene  $n_c$  as a function of the global background rate of composite mutations across all genes, adjusted for multiple covariates per gene, including its replication timing  $r$ , coding sequence length  $l$ , the per cent of GC content  $g$  and the chromatin state of the gene  $h$ . Coding sequence length and per cent of GC content were obtained from the Biomart community portal<sup>34</sup> for Ensembl human reference genome GRCh37. For the purposes of statistical testing, the noncoding promoter region of *TERT* was added as a distinct unit (gene) for which we computed distinct values of per cent GC content and length for the region targeted by the MSK-IMPACT assay design. Replication timing and chromatin state for each gene were obtained from previous estimates<sup>9</sup>. Additional covariates included the version of the MSK-IMPACT assay in which the gene was introduced  $i$ , and the average total DNA copy number of the gene across its mutated samples  $t$ . As the composite mutation rate for a gene depends on both the number of composite mutant tumours and the number of samples mutated (that is, the exposure for the count of composite mutants), an offset term was added to the model that represents the log-transformed number of tumour samples containing mutations in the gene of interest. The observed number of composite mutant tumours for a gene was therefore modelled as  $n_c \sim NB(r + l + g + h + i + t + \text{offset}(\log(n_s)))$ . Using this model, we predicted the number of composite mutant tumours for each gene arising by chance,  $\hat{n}_c$ , calculating the expected fraction of samples with a composite mutation (out of the total number of mutated samples) in each gene  $\hat{p}$ . We then used a binomial test to evaluate the null hypothesis that for each gene the observed number of composite mutations arose owing to random chance. Here, we modelled the incidence of composite mutations per gene using a binomial distribution, and calculated the probability of  $n_s$  tumour specimens containing composite mutations in  $n_c$  tumour specimens by chance given  $\hat{p}$ :

$$\Pr(X \geq n_c) = \sum_{i=n_c}^{n_s} \binom{n_s}{i} \hat{p}^i (1 - \hat{p})^{n_s - i}$$

Our parameterization  $\hat{p}$  was estimated using nonsynonymous mutations, including those under positive selection in cancer (for example, hotspots), which may reduce overall model sensitivity. We therefore evaluated multiple alternative parameterizations of  $\hat{p}$ , including using (1) nonsynonymous mutational data that exclude known hotspot mutations under selection and (2) only synonymous mutations. Neither

alternative parameterization produced a qualitatively distinct result for genes originally detected as significantly enriched, but did increase the overall sensitivity of the test. To ensure appropriate control for potential false-positive findings, we leveraged the parameterization from the complete dataset on nonsynonymous mutational data. Moreover, we observed no difference in the rate of synonymous mutations among genes that were either enriched for composite mutations or not ( $P=0.2$ , Mann-Whitney  $U$ -test), indicating there was little evidence for the accumulation of variants in the absence of selective pressure.

Finally, all unique individual mutant residues present in five or more nonhyperpermuted cases, excluding known or likely resistance mutations, were also assessed for the significance of their enrichment for arising as composite mutations. All missense, nonsense, splice-site and translation start-site mutations at a given residue were included, as were unique mutant positions in the promoter of *TERT* and in-frame indels spanning known hotspots of clustered indels<sup>13</sup>. For each residue in a given gene, we assessed whether it arose as part of a composite mutation significantly more often than all other mutant residues in the same gene using a right-sided Fisher's exact test. Mutant residues were considered significant at FDR-adjusted  $P < 0.01$  ('Statistical analyses and figures').

### Attributing mutations to mutagenic processes

We attributed the individual variants that comprise composite mutations to a mutational origin using 1 of 30 established mutational signatures<sup>35,36</sup>. Mutational signature decomposition in each tumour was performed as described in 'Mutational burden classification' and a signature was considered present if it accounted for five or more substitutions in the affected specimen (to ensure high-confidence decompositions in targeted sequencing data with comparatively fewer mutations relative to broader-scale sequencing). Multiple signatures of the same aetiology were merged by combining the frequency distribution of trinucleotide contexts (APOBEC signatures 2 and 13; MMR signatures 6, 15, 20, 21 and 26; and smoking-associated signatures 4, 18, 24 and 29). A substitution was attributed to a mutational signature present in a given case if the probability weight of the relevant trinucleotide exceeded 10%. For a substitution attributed to multiple signatures present in an affected tumour, it was attributed to the signature that was most frequently associated with the affected cancer type. To adjust for the nonspecificity of trinucleotide context probabilities for smoking-associated signatures, C>A mutations—regardless of trinucleotide context—were considered smoking-associated in tumours for which mutational signature decomposition identified a smoking signature (in oesophageal squamous and adenocarcinomas; head and neck squamous; hepatobiliary; hepatocellular; lung squamous, adenocarcinoma, and adeno-squamous, oral cavity and renal cell carcinoma)<sup>37</sup>. Substitutions of a trinucleotide context of insufficient probability in any signature in an affected tumour were considered of ambiguous origin and not attributable, and those mutations that could be attributed to ageing and another signature present in a given tumour were considered nonseparable and classified as being of multiple signatures.

Finally, we also considered several additional mechanisms that can drive site-specific mutation rates as potential sources of composite mutations<sup>38–40</sup>. First, we estimated the mutation rate within 1 kb up- and downstream of all nucleosome dyads (obtained from <https://bitbucket.org/bbglab/nucleosome-periodicity/src/master/>) mapping to regions sequenced in the MSK-IMPACT panels. Having fit a spline to the mutation rate distribution, we calculated the full-width-half-maximum distances from the dyad and compared the rate of singleton and composite mutations within this region (Extended Data Fig. 2b). We conducted a similar analysis on the potential effect of active coding transcription factor binding sites (TFBSs) on composite mutations. We obtained the positions of active TFBSs in coding regions of the genome via integration with DNase I hypersensitive binding sites in human melanocytes following an established procedure<sup>39</sup>. The mutation rate within 1 kb

of these active TFBSSs were inferred using TCGA cutaneous melanoma samples from the TCGA MC3 dataset to increase the total number of mutations among melanoma samples. We then assessed the proximity of singleton and composite mutations to the elevated mutation rate at TFBS sites as described for nucleosome dyads (Extended Data Fig. 2).

To investigate the effect of APOBEC3A-mediated mutagenesis, we obtained the position of the optimal stem-loop DNA structure favoured by APOBEC3A from published sources<sup>40</sup>. We investigated the overlap of such optimal sites with those mutant alleles that were enriched for arising as a composite mutation. In total, only 1 of 86 significant residues enriched for arising as a composite mutation was at the position of the optimal APOBEC3A substrate (*ARID1A*<sup>S2264</sup>). Finally, we compared the rate of composite mutations involving known hotspot mutations as described in ‘Mutational data and annotation’ with those derived from an orthogonal method optimized to reduce false-positive mutations due to site-specific mutagenesis<sup>41</sup>. Controlling for overlapping gene content, there was no difference between the proportion of composite mutations involving hotspot mutations based on the origin of the hotspot mutations (per cent and 95% confidence interval are 9.6 (9.2–10) versus 10 (9.6–10.5),  $P = 0.2$ , two-sample Z-test), indicating that no excess of false-positive hotspots due to site-specific mutagenesis are driving the results described here.

### Phasing composite mutations

The allelic configuration of composite mutations (phase)—*in cis* (arising on the same allele) or *in trans* (arising on different alleles)—was inferred primarily from sequencing read support. In brief, for each pair of somatic mutations in a composite mutant, all reads spanning the relevant loci were re-aligned to the reference genome (hg19) by pairwise sequence alignment using a Needleman–Wunsch algorithm<sup>42</sup>. The numbers of unique reads that supported both alleles being wild type (AB), both alleles being mutant (ab) or a mixture of mutant and wild-type alleles (aB or Ab) were subsequently tabulated. For the purposes of the present study, composite mutations were classified as *in cis* when: (1) three or more spanning reads supported both mutant alleles ( $ab \geq 3$ ) and (2) at least one of these variants was called by two or fewer spanning reads that called the other variant as wild type (that is,  $aB \leq 2$  or  $Ab \leq 2$ , or both). Composite mutations were classified as *in trans* when: (1) each variant was supported by three or more reads that were simultaneously wild type for its partner mutation ( $aB \geq 3$  and  $Ab \geq 3$ ), (2) two or fewer reads called both mutant alleles ( $AB \leq 2$ ) and (3) the mutations arose in the same tumour cell population on the basis of their cancer cell fractions (CCFs, see ‘Assessing cellular context and molecular timing’). There is an inherent difference in the sensitivity of detection for *cis* and *trans* variants, specifically that *trans* variants must satisfy at least two read-support positive criteria ( $aB \geq 3$  and  $Ab \geq 3$ ) and are required to be in the same cell, whereas *cis* variants require only a single positive criterion ( $ab \geq 3$ ) without any constraint of evidence for arising in the same cell. This difference in sensitivity for detection probably explains—to some extent—the increased number of *cis* relative to *trans* composite mutations. To determine the effect of this sensitivity bias, we also phased variants with at least one synonymous mutation. We observed no difference in the rate of synonymous composite mutations in oncogenes versus tumour-suppressor genes (5% versus 7%,  $P = 0.2$ , Mann–Whitney *U*-test), in contrast to the significant difference in nonsynonymous composite mutations (14% versus 35%,  $P < 10^{-6}$ ). To control for differences in the sensitivity of detection of *cis* and *trans* mutations, analyses of the effects of allelic configuration on composite mutations compared the relative fraction of *cis* and *trans* mutations between two defined groups (for example, oncogenes versus tumour-suppressor genes).

We additionally inferred the phase of select composite mutants associated with therapeutic resistance mutations in regions of clonal loss of heterozygosity (LOH or copy-neutral LOH). Composite mutants spanned by LOH were assumed to be *in cis* if the spanning locus had a minor copy number of zero and a total copy number of one or more

(LOH via heterozygous loss, copy-neutral LOH or the latter combined with subsequent genomic gains) inferred from the aforementioned purity-corrected integer copy number data from FACETS. These must also have arisen in the same tumour cell population as estimated from CCFs (as described in ‘Assessing cellular context and molecular timing’) and their observed mutant allele frequencies were approximately equal to the expected mutant allele frequencies for a given copy number state in a *cis* allelic configuration (95% confidence intervals of the observed mutant allele frequency overlap the expected mutant allele frequency of the given copy number configuration, controlling for tumour purity). Composite mutations that did not satisfy any of the aforementioned conditions were not able to be unambiguously phased.

As with other short-read sequencing data, our phasing approach is limited by the requirement that any two mutations arise within sufficient physical proximity in the genome to be spanned by common aligned sequencing reads. Although the higher depth of sequencing coverage in our targeted clinical sequencing platform (about 700-fold median in the tumour samples) does increase the likelihood of sequencing a fragment of tumour DNA encompassing both somatic mutations, and improves the quantification of CCFs by reducing measurement error<sup>8</sup>, this limitation cannot be overcome with short-read sequencing.

### Assessing cellular context and molecular timing

We estimated the clonality of all somatic mutations in each affected tumour specimen (the CCF) using the FACETS framework, as described previously<sup>8</sup>. To ensure conservative estimates, all somatic mutations were conservatively assumed to have arisen on the major (more common) allele, thus minimizing the possibility of overestimating the CCF. To determine whether the constituents of a composite mutation arose in the same cell, we defined a criterion based on the confidence intervals of the CCF. Specifically, if the sum of the lower bounds of the 95% confidence intervals for each mutation CCF summed to greater than 1, the two somatic mutations in the same gene and tumour specimen were considered to exist within the same cancer cell population. If either of the two somatic mutations were clonal (the upper bound of the 95% confidence interval overlapped 1), then both mutations were considered to have arisen in the same tumour cell population.

We inferred the chronological order of two somatic mutations in each composite mutation on the basis of their estimated CCFs. Any mutations previously associated with acquired resistance to targeted therapies were excluded, as these will arise after the originating sensitizing lesion and skew results. Only composite mutations determined to arise in the same tumour cell population (based on the sum of CCFs) were considered and required previous evidence establishing both mutations as candidate functional driver mutations individually. The 95% confidence intervals of the CCFs of both mutations were inferred as previously described<sup>43</sup>. If the lower bound of the 95% confidence interval was greater than the upper bound of the 95% confidence interval for a second variant, then the first mutation was determined to have a greater clonality, and therefore to have arisen first in the tumour. Similarly, if the upper bound of the 95% confidence interval of a mutation was less than the lower bound of the other mutation in the composite, it was considered to have arisen second. If the 95% confidence intervals of CCFs of the two mutations in the composite overlapped, or if there was not sufficient evidence that the two mutations existed in the same cancer cell population in the affected tumour specimen, we considered their chronology to be indeterminate.

### *TP53* composite mutation analysis and validation studies

For the generation of MSCV-p53-IRES-GFP constructs (pMIG-p53 cDNAs), methods were as follows. Fragments encoding wild-type, single- or composite-mutant *Trp53* (mouse orthologue to human *TP53*) cDNAs were obtained from IDT or SGI-DNA, and cloned into pMIG (Addgene no. 9044) using standard restriction enzyme-based methods. In brief, *Trp53* cDNAs were amplified using primers that add BgIII and EcoRI restriction

# Article

sites on the 5' and 3' regions, respectively, and subsequently digested and cloned into linearized pMIG backbone containing BglII and EcoRI cloning overhangs. All constructs were sequence-verified using Sanger sequencing. Primer sequences are available in Supplementary Table 5.

HEK293T (ATCC CRL-3216) cells were obtained from ATCC. Mouse *Kras*<sup>G12D/+</sup> *Trp53*<sup>-/-</sup> lung adenocarcinoma cells were provided by the Jacks Laboratory<sup>44</sup>. All cells were maintained in a humidified incubator at 37 °C with 5% CO<sub>2</sub>, and grown in DMEM supplemented with 10% FBS and 100 IU/ml penicillin–streptomycin. For virus production, 7.5 million HEK293T cells were plated in 15-cm plates the day before transfection. The following day, cells were transfected with 10 µg pMIG-p53 cDNA (or pMIG-empty as control) and 10 µg of pCL-Eco (Addgene no. 12371) using 50 µl of Lipofectamine 2000 (ThermoFisher). Twenty-four hours later, transfection medium was replaced with fresh DMEM. Two rounds of virus were collected (at 48 and 72 h after transfection), pooled and kept at 4 °C until used for cell transduction. One million *Kras*<sup>G12D/+</sup> *Trp53*<sup>-/-</sup> lung adenocarcinoma cells were seeded in 10-cm plates and immediately transduced with retroviral supernatants and 8 µg/ml polybrene. Cells were grown for 48 h before purifying using fluorescence-activated cell sorting (FACS). All transductions were done in triplicate. Following transduction, stable GFP<sup>+</sup> populations were purified by FACS on a FACSAria (BD Biosciences). One hundred and twenty hours after transduction, total RNA was isolated using the RNeasy Mini Kit (Qiagen) following standard manufacturer protocols.

Purified polyA mRNA was subsequently fragmented, and first- and second-strand cDNA synthesis was performed using standard Illumina mRNA TruSeq library preparation protocols. Double-stranded cDNA was subsequently processed for TruSeq dual-index Illumina library generation. For sequencing, pooled multiplexed libraries were sequenced on NextSeq instrumentation in high-output mode, generating approximately 12 million 76-bp single-end reads per replicate condition. The resulting RNA sequencing data were analysed by first trimming adaptor sequences using Trimmomatic<sup>45</sup>. Sequencing reads were aligned to GRCm38.p5 (mm10) using STAR<sup>46</sup>, and genome-wide transcript quantification was performed using featureCounts<sup>47</sup>. After removing transcripts with fewer than 8 aligned reads (low undetected expression at given library size,  $n=9,848$  transcripts retained), differentially expressed genes were identified using DESeq2, with a cutoff of absolute log<sub>2</sub>-transformed fold change  $\geq 1$  and adjusted  $P < 0.01$  between experimental conditions<sup>48</sup>. Mouse genes were mapped to human homologues using gene homologies provided by the Mouse Genome Database project<sup>49</sup>. Principal components analysis was performed with output from DESeq2<sup>48</sup>. For fluorescent competition assays, FACS-purified *Kras*<sup>G12D/+</sup> *Trp53*<sup>-/-</sup> lung adenocarcinoma cells stably transduced with either pMIG-empty or pMIG-p53-R277T-E284D were mixed at about 60:40 with untransduced parental cells and cultured in vitro for 10 days. The percentage of GFP<sup>+</sup> cells was monitored over time using a Guava easyCyte HT flow cytometer (Millipore).

All mouse experiments were approved by the MSKCC Internal Animal Care and Use Committee. No pre-specified sample size was required, and 5 or 10 mice per condition were used. Mice were maintained under specific-pathogen-free conditions, and food and water were provided ad libitum. Mice (Hsd:athymic nude-*Foxn1*<sup>nu</sup>, abbreviated *Nu/Nu*) were purchased from Envigo (stock no. 069). For experiments involving orthotopic transplantation of *Kras*<sup>G12D/+</sup> *Trp53*<sup>-/-</sup> lung adenocarcinoma cells, 100,000 cells stably transduced with either empty vector (pMIG-empty) or *Trp53*-mutant cDNAs (pMIG-p53-R277T, pMIG-p53-E284D or pMIG-p53-R277T-E284D) were resuspended in 200 µl of PBS and injected into the tail vein of 6–8-week-old *Nu/Nu* female mice. These stable cell populations were generated and FACS-purified as described above, and injected at 120 h after transduction.

## TERT promoter mutation analysis and validation

*TERT* promoter mutations present in five or more patients, accounting for multiple samples per patient, were assessed for co-occurrence and

mutual exclusivity among composite mutations via two-sided Fisher's exact test. A pair of somatic mutations with  $P < 0.01$  were considered co-occurring (or mutually exclusive) if their log-odds ratio was greater (or less) than zero. To predict the affinity for GABPA to bind *TERT* promoter mutant alleles, 31-bp DNA sequences for wild-type or mutant *TERT* centred on each of 205G>A (that is, chromosome 5, 1295205G>A), 228G>A and 250G>A mutations were extracted and generated by editing the appropriate base. The position frequency matrix for GABPA binding profiles in humans was acquired from JASPAR2018<sup>50</sup> (Matrix identifier MA0062.1), and scores quantifying the predicted affinity of GABPA for each *TERT* promoter sequence were calculated using TFBSTools<sup>51</sup>. Only binding site motifs overlapping the relevant locus in each of the wild-type and mutant sequence were retained.  $P$  values quantifying the likelihood of a GABPA binding site in each sequence to arise by chance were calculated using TFMPvalue<sup>52</sup>.

To assess the effect of *TERT* promoter composite mutations on *TERT* expression, A375, SK-Mel2 and SK-Mel30 melanoma cell lines were obtained (kindly provided by laboratories of N. Rosen and T. Mergenhouw). pGL4.0-TERT wild type, G228A and G250A plasmids were provided by the J. Costello laboratory (Addgene plasmids no. 84924, 84926 and 84925)<sup>21</sup>. pGL4.0-TERT G205A, G205A/G228A and G205A/G250A plasmids were generated using Q5 Site-Directed mutagenesis kit (NEB, EO554S). All plasmids were verified using Sanger sequencing. Thereafter,  $1 \times 10^4$  cells from A375, SK-Mel2, and SK-Mel30 were seeded into each well of 96-well plates. Cells were transiently transfected with pGL4.0-empty vector (Promega), TERT wild type or mutant plasmids (180 ng per well) along with pGL4.74[hRluc/TK] vector (18 ng per well, Promega) as an internal control using Lipofectamine 3000 (Thermo Fisher). Dual luciferase activity measurement was performed 48 h after transfection using the Dual-Luciferase Reporter Assay System (Promega) following the manufacturer's instructions. The firefly luciferase activity of individual wells was normalized relative to Renilla luciferase activity. Experiments were performed in biological quadruplicates or pentaplicates. To quantify the effect of a specific *TERT* variant, we compared individual genotypes (for example, *TERT*<sup>G205A</sup> to wild type) using linear models of luciferase expression, in which we controlled for the baseline telomerase expression of each cell line—that is, luc<sup>-</sup> variant + cell line + constant, in which variant is a binary term that encodes the presence or absence of a genotype (relative to the chosen reference), and cell line is a factor introduced to control for the contribution of the baseline expression of each cell line. All cell lines used for either the *TERT* or *TP53* functional validation experiments were authenticated by short-tandem-repeat analysis and confirmed negative for mycoplasma.

## Statistical analyses and figures

All statistical analyses were performed using the R statistical programming environment (version 3.5.0). Figures were generated using either base R or the ggplot2 library. Error bars indicate the 95% binomial confidence intervals calculated using the Pearson–Klopper method, unless otherwise noted. Confidence intervals for the down-sampling analysis were calculated using the loess.sd function from the msir library.  $P$  values for the difference in proportions were calculated using Fisher's exact test or two-sample Z-tests, unless otherwise noted.  $P$  values were corrected for multiple comparisons using the Benjamini–Hochberg method and reported as  $Q$  values when applicable.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

All mutational data from the prospective sequencing cohort are available at [http://download.ncbi.nlm.nih.gov/composite\\_mutations\\_maf.txt.gz](http://download.ncbi.nlm.nih.gov/composite_mutations_maf.txt.gz). Mutational data from The Cancer Genome Atlas were acquired

from <https://gdc.cancer.gov/about-data/publications/pancanatlas>. RNA sequencing data have been deposited in the Gene Expression Omnibus with accession number GSE136295. All other genomic and clinical data accompany the Article, and are available in the Extended Data and Supplementary Information. All other materials are available upon request from the corresponding authors.

## Code availability

Source code for these analyses is available at <https://github.com/taylor-lab/composite-mutations>.

25. Hyman, D. M. et al. Precision medicine at Memorial Sloan Kettering Cancer Center: clinical next-generation sequencing enabling next-generation targeted therapy trials. *Drug Discov. Today* **20**, 1422–1428 (2015).
26. Cheng, D. T. et al. Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. *J. Mol. Diagn.* **17**, 251–264 (2015).
27. Zehir, A. et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat. Med.* **23**, 703–713 (2017).
28. Chakravarty, D. et al. OncoKB: a precision oncology knowledge base. *JCO Precis. Oncol.* **1**, 1–16 (2017).
29. Campbell, B. B. et al. Comprehensive analysis of hypermutation in human cancer. *Cell* **171**, 1042–1056.e10 (2017).
30. Niu, B. et al. MSIsensor: microsatellite instability detection using paired tumor-normal sequence data. *Bioinformatics* **30**, 1015–1016 (2014).
31. Middha, S. et al. Reliable pan-cancer microsatellite instability assessment by using targeted next-generation sequencing data. *JCO Precis. Oncol.* **1**, 1–17 (2017).
32. Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Campbell, P. J. & Stratton, M. R. Deciphering signatures of mutational processes operative in human cancer. *Cell Rep.* **3**, 246–259 (2013).
33. Dixon, P. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* **14**, 927–930 (2003).
34. Smedley, D. et al. The BioMart community portal: an innovative alternative to large, centralized data repositories. *Nucleic Acids Res.* **43**, W589–W598 (2015).
35. Forbes, S. A. et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* **43**, D805–D811 (2015).
36. Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
37. Alexandrov, L. et al. The repertoire of mutational signatures in human cancer. *Nature* **578**, 94–101 (2020).
38. Pich, O. et al. Somatic and germline mutation periodicity follow the orientation of the DNA minor groove around nucleosomes. *Cell* **175**, 1074–1087.e18 (2018).
39. Sabarinathan, R., Mularoni, L., Deu-Pons, J., Gonzalez-Perez, A. & López-Bigas, N. Nucleotide excision repair is impaired by binding of transcription factors to DNA. *Nature* **532**, 264–267 (2016).
40. Buisson, R. et al. Passenger hotspot mutations in cancer driven by APOBEC3A and mesoscale genomic features. *Science* **364**, eaaw2872 (2019).
41. Hess, J. M. et al. Passenger hotspot mutations in cancer. *Cancer Cell* **36**, 288–301.e14 (2019).
42. Needleman, S. B. & Wunsch, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**, 443–453 (1970).
43. McGranahan, N. et al. Clonal status of actionable driver events and the timing of mutational processes in cancer evolution. *Sci. Transl. Med.* **7**, 283ra54 (2015).
44. Dimitrova, N. et al. Stromal expression of miR-143/145 promotes neoangiogenesis in lung cancer development. *Cancer Discov.* **6**, 188–201 (2016).
45. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
46. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
47. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
48. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
49. Bult, C. J., Blake, J. A., Smith, C. L., Kadin, J. A. & Richardson, J. E. Mouse genome database (MGD) 2019. *Nucleic Acids Res.* **47**, D801–D806 (2019).
50. Khan, A. et al. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.* **46**, D260–D266 (2018).
51. Tan, G. & Lenhard, B. TFBSTools: an R/bioconductor package for transcription factor binding site analysis. *Bioinformatics* **32**, 1555–1556 (2016).
52. Touzet, H. & Varré, J.-S. Efficient and accurate P-value computation for position weight matrices. *Algorithms Mol. Biol.* **2**, 15 (2007).
53. Supek, F. & Lehner, B. Clustered mutation signatures reveal that error-prone DNA repair targets mutations to active genes. *Cell* **170**, 534–547.e23 (2017).
54. Nik-Zainal, S. et al. Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993 (2012).

**Acknowledgements** We thank the members of the E.R. and B.S.T. laboratories for discussion and support. This work was supported by National Institutes of Health awards P30 CA008748, P01 CA087497 (S.W.L.), U54 OD020355 (S.W.L. and B.S.T.), R01 CA207244 (B.S.T.), R01 CA204749 (B.S.T.), R01 CA245069 (B.S.T.); Brown Performance Group ICI Fund (N.V. and E.R.), Society of MSK (N.V. and E.R.), American Cancer Society, Anna Fuller Fund and the Josie Robertson Foundation (B.S.T.). F.J.S.-R. is an HHMI Hanna Gray Fellow supported in part by an MSKCC Translational Research Oncology Training Fellowship (T32-CA160001). S.W.L. is an investigator of the Howard Hughes Medical Institute.

**Author contributions** A.N.G., E.R. and B.S.T. conceived the study. C.M.B., E.B., P.J., A.V.P., A.L.R., N.D.F., C.B., N.S., E.R. and B.S.T. assisted with genomic data collection and analytical methodology development. F.J.S.-R., Y.C., N.V., M.S. and S.W.L. designed and performed the experiments. Y.J.H. and T.B. assisted with RNA sequencing. A.N.G., E.R. and B.S.T. wrote the manuscript with input from all authors.

**Competing interests** N.V. reports advisory board activities for Novartis and consulting activities for Petra Pharmaceuticals. M.S. has received research funding from Puma Biotechnology, Daiichi-Sankyo, Immunomedics, Targimmune and Menarini Ricerche; is a cofounder of Medendi.org, and is on the advisory boards of the Bioscience Institute and Menarini Ricerche. S.W.L. is a founder and scientific advisory board member of Oric Pharmaceuticals, Mirimus, Inc. and Blueprint Medicines; and is on the scientific advisory boards of Constellation Pharmaceuticals, Petra Pharmaceuticals and PMV Pharmaceuticals. B.S.T. reports receiving honoraria and research funding from Genentech and Illumina, and advisory board activities for Boehringer Ingelheim and Loxo Oncology, a wholly owned subsidiary of Eli Lilly, Inc. All stated activities were outside of the work described here. The other authors declare no competing interests.

### Additional information

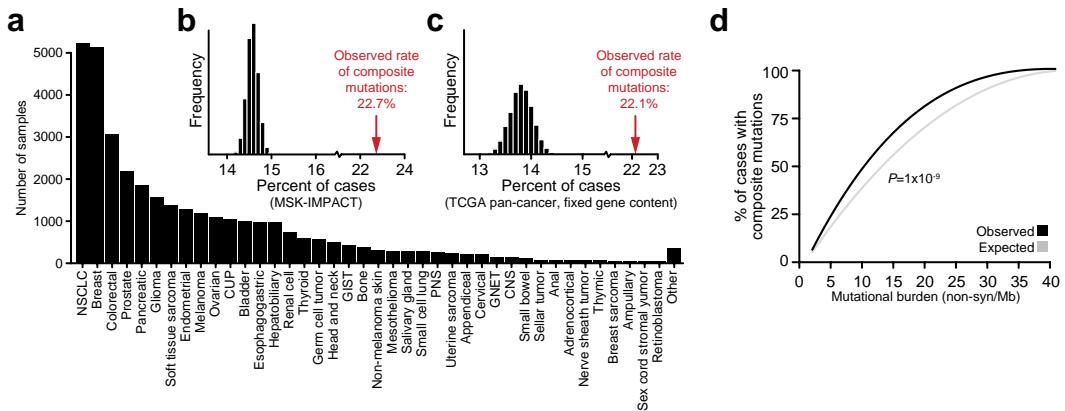
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2315-8>.

**Correspondence and requests for materials** should be addressed to E.R. or B.S.T.

**Peer review information** *Nature* thanks Moritz Gerstung, Mark Lackner and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

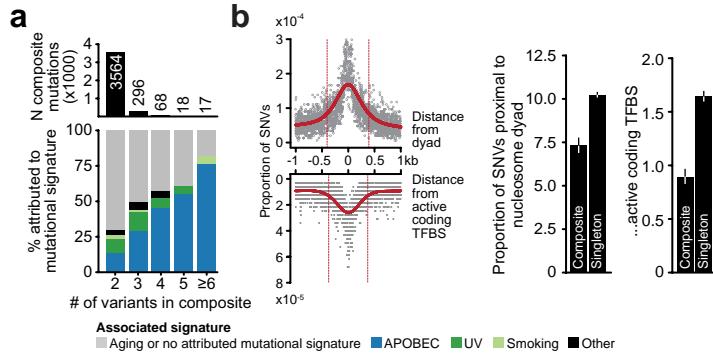
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

# Article



**Extended Data Fig. 1 | Study cohort and rates of composite mutations.** **a**, Distribution of cancer types in the study cohort. **b**, The rate of composite mutations (22.7% of all tumours) compared to a simulated background rate (black,  $P=10^{-5}$  from one-sided permutation test for enrichment with 100,000 random permutation-based simulations (no permutation exceeded observed value)). **c**, The observed rate of composite mutations in the primary untreated cancers of the TCGA cohort ( $n=10,908$  solid tumours) when

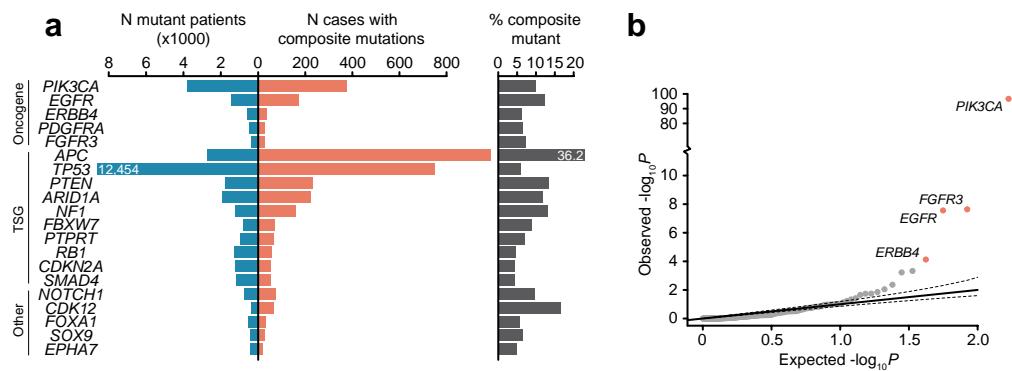
controlling for gene content for consistency with the targeted sequencing panel of the prospective cohort studied here. The null distribution from sampling (Methods) is shown in black. **d**, The observed and expected rate of composite mutations in tumours of the indicated tumour mutational burden (as in Fig. 1b,  $n=30,505$  biologically independent tumour samples with tumour mutational burden  $\leq 40$ ,  $P=1 \times 10^{-9}$  from two-sided Wilcoxon signed-rank test).



**Extended Data Fig. 2 | Sources of local hypermutation.** **a**, The number of composite mutations comprising two or more constituent variants (top) and the distribution of likely causative mutational signatures among them (bottom). Composite mutants comprising greater than three mutations were increasingly produced by APOBEC-associated mutagenesis, indicative of localized hypermutation<sup>53,54</sup>, but accounted for a minority of events cohort-wide. **b**, Left, the somatic mutational data in the study cohort reflect the elevated mutation rates previously observed at both the positions closest to

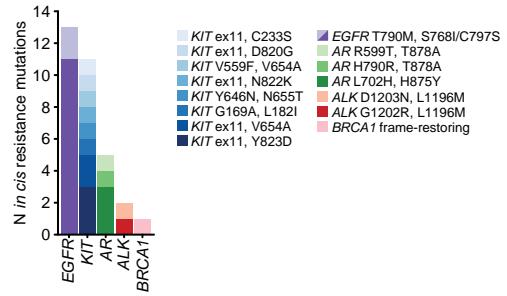
the nucleosome dyad as well as DNA bound to active transcription-factor binding sites<sup>38,39</sup>. However, mutations arising in composite events were proportionally less often proximal to such sites (defined here as within the full width at half maximum of the peak of mutation rate (red)) than were singleton mutations (right,  $P=10^{-27}$  and  $10^{-47}$ , respectively; two-sided two-sample Z-test,  $n=323,883$  single-nucleotide substitutions arising in 471 biologically distinct melanoma samples).

# Article

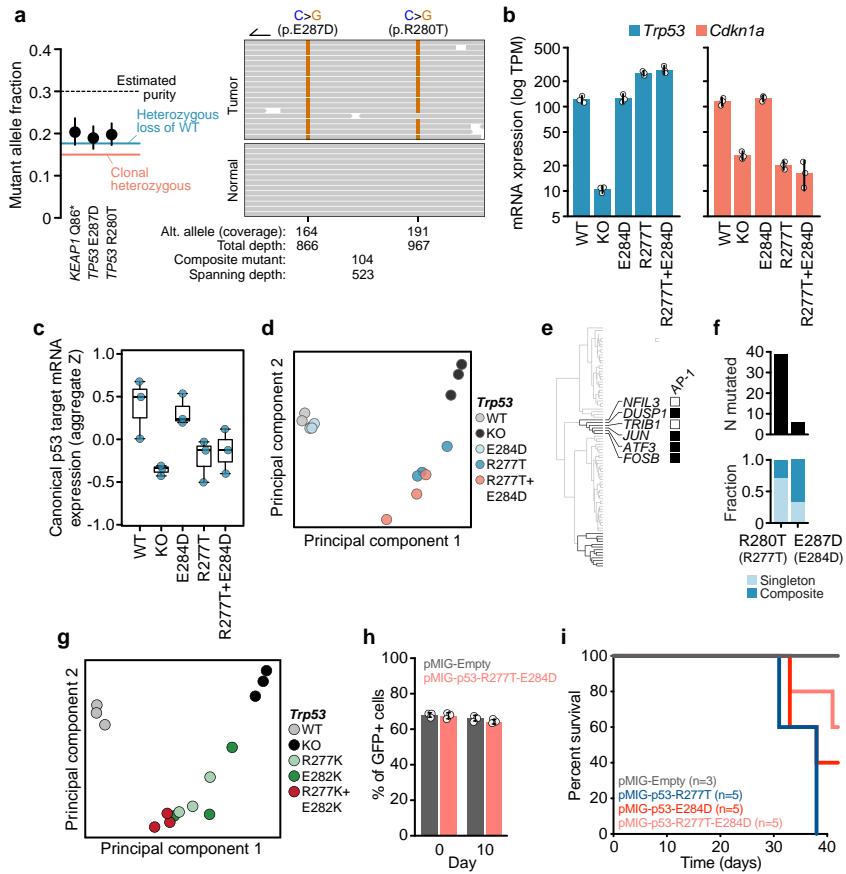


**Extended Data Fig. 3 | Number and distribution of composite events across genes.** **a**, The number and percentage of cases in the study cohort containing composite mutations in the indicated genes (right) juxtaposed to their overall mutation rate (left). Genes with a significant enrichment of composite mutations are shown ( $Q < 0.01$ , FDR-adjusted  $P$  values from one-sided binomial test for enrichment,  $n = 26,997$  as in Fig. 2b), limited to the top 10 genes by significance in each category of gene function, unless fewer. **b**, The significance of enrichment for composite mutations ( $n$  and statistical tests as described in **a** and Fig. 2b) limited to 168 oncogenes.

test for enrichment,  $n = 26,997$  as in Fig. 2b), limited to the top 10 genes by significance in each category of gene function, unless fewer. **b**, The significance of enrichment for composite mutations ( $n$  and statistical tests as described in **a** and Fig. 2b) limited to 168 oncogenes.

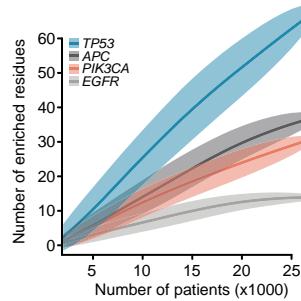


**Extended Data Fig. 4 | cis composite secondary-resistance mutations.** The *cis* composite mutations classified as arising in post-treatment specimens due to acquired resistance to one of several molecularly targeted therapies in the study cohort.



**Extended Data Fig. 5 | Phenotypic characterization of TP53 composite mutants.** **a**, *TP53<sup>R280T/E287D</sup>* mutant lung adenocarcinoma. Left, mutant allele fractions of clonal *TP53* mutations consistent with loss of wild-type *TP53* (error bars, 95% binomial confidence intervals). Expected mutant allele fractions of different copy number states are shown as horizontal lines. Mutant *KEAP1* in the same tumour (with LOH) is shown for reference. Right, spanning reads indicating *cis* mutations. **b**, Right and left, *Trp53* and *Cdkn1a* mRNA expression in *Kras<sup>G12D</sup>/Trp53<sup>mut</sup>* mouse lung cancer cells expressing distinct *Trp53* genotypes. Bars, average of three replicates, error bars are 95% confidence intervals. **c**, The aggregate Z-score per replicate for the mRNA expression of canonical p53-target genes ( $n = 3$  replicates per allele; box centre is median, edges are 25% and 75% quartiles, whiskers are minimum and maximum of the most extreme values). **d**, Principal component analysis of the transcriptomes of *Trp53* genotypes ( $n = 3$  replicates shown per condition). **e**, Dendrogram as in Fig. 3f, indicating the genes of interest (effectors of the AP-1 transcription

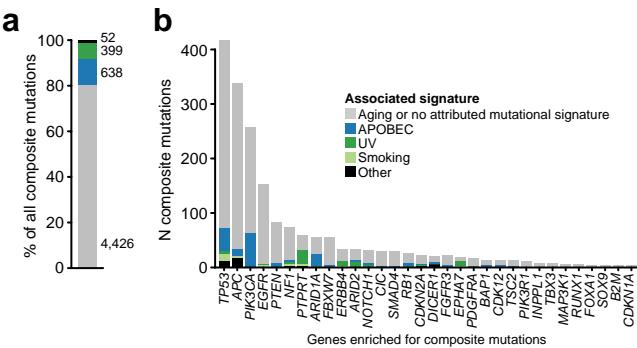
factor network (PID\_AP1\_PATHWAY;  $Q = 1.4 \times 10^{-7}$  based on computed overlap (using mSigDB) with  $n = 5,501$  gene sets from the curated C2 collection)). **f**, The prevalence of *TP53<sup>R280T</sup>* and *TP53<sup>E287D</sup>* mutations (top), and the fraction arising as composite mutants (bottom). The corresponding mouse alleles are given in parentheses. **g**, Principal component analysis of the transcriptomes of the *Trp53<sup>R277K/E282K</sup>* composite mutation genotypes (as in **d**).  $n = 3$  replicates per allele. **h**, The percentage of GFP<sup>+</sup> FACS-purified *Kras<sup>G12D</sup>/Trp53<sup>-/-</sup>* lung adenocarcinoma cells stably transduced with pMIG-empty or pMIG-p53-R277T-E284D, and cultured in vitro for 10 days in a 60:40 mixture with untransduced parental cells. Bar indicates mean, error bars are s.d.,  $n = 3$  independent infections. **i**, Overall survival of immunocompromised mice bearing lung tumours of the indicated *Trp53* genotypes generated by tail vein injection of stably transduced and FACS-purified *Kras<sup>G12D/+</sup>/Trp53<sup>-/-</sup>* lung adenocarcinoma cells ( $n = 100,000$  cells).



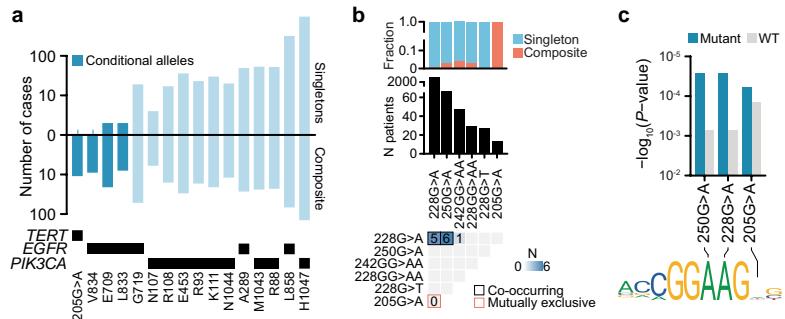
**Extended Data Fig. 6 | Saturation analysis of genes for composite mutation detection.**

Down-sampling indicates the number of residues identified as enriched for arising in composite mutations in each of four genes ( $Q < 0.1$ , FDR-adjusted one-sided Fisher's exact tests as in Fig. 4a;  $n = 1,000\text{--}26,997$  patients per down-sample) as a function of the number of tumours sequenced (LOESS fit is shown with 95% confidence interval). Four genes that accounted for the greatest proportion of all enriched residues detected are shown (Fig. 4a). *EGFR* appears to reach saturation for discovery of residues enriched for arising in composite, whereas the other genes have not yet reached saturation for discovery at the current cohort size.

# Article



**Extended Data Fig. 7 | Mutational signature attribution among composite mutations.** **a**, The fraction of all composite mutations identified here in which one or both individual mutations could be unambiguously attributed to an established mutational signature. The majority of composite variants could not be directly attributed to APOBEC, ultraviolet, smoking or other known mutational signatures. **b**, The fraction of composite mutations per gene in which one or both variants could be attributed to an established mutational signature.



**Extended Data Fig. 8 | Conditional mutant alleles.** **a**, The number of affected cases containing each of the indicated somatic mutations in *TERT*, *EGFR* or *PIK3CA* as either individual mutations (top) or as part of composite mutants (bottom). Conditional mutations were defined as those statistically enriched for arising as part of composite mutations, but seldom as individual hotspot mutations in cancer (predominantly accompanied by a second somatic mutation). **b**, The incidence of *TERT* promoter mutations and the fraction

arising as composite mutations (orange). Bottom, the co-occurrence and mutual exclusivity of composite mutations in the *TERT* promoter (The *P* values for  $n=5$  and 6 co-occurring mutations are 0.002 and  $3 \times 10^{-7}$ , respectively, and for 0 mutually exclusive mutations is  $1 \times 10^{-25}$ ; two-sided Fisher's exact test,  $n=29,507$  patients). **c**, Transcription factor GABPA binding affinity for mutant and wild-type *TERT* promoter sequences at the 228G>A, 250G>A and the conditional 205G>A allele.

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

R Statistical Computing environment (v3.5.0)

Data analysis

FACETS (v0.5.6); maf2maf (v1.6.17); MSIensor (v0.2); OncoKB (v1.0.4); R (v3.5.0); R packages: binom (v1.1-1), BiomaRt (v2.36.1), car (v3.0-3), cowplot (v1.0.0), data.table (v1.12.2), DESeq2 (v1.22.2), MASS (v7.3-51.4), Rcpp (v1.0.2), TFBSTools (v1.20.0), TFMPvalue (v0.0.8), vegan (v2.5-6); custom phasing software (<https://github.com/taylor-lab/MutationPhaser>); mutation signature decomposition (<https://github.com/mskcc/mutation-signatures>); featureCounts (v1.6.3); Trimmomatic (v0.36); STAR (2.5.3a); Human genome reference (GRCh37); Mouse genome reference (GRCm38.p5, mm10); Source code for these analyses is available at <http://github.com/taylor-lab/composite-mutations>.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All mutational data from the prospective sequencing cohort is available through the cBioPortal for Cancer Genomics: [http://download.cbioportal.org/composite\\_mutations\\_maf.txt.gz](http://download.cbioportal.org/composite_mutations_maf.txt.gz). Mutational data from The Cancer Genome Atlas was acquired from <https://gdc.cancer.gov/about-data/publications/pancanatlas>. RNA sequencing data were deposited in the GEO with accession number GSE136295. All other genomic and clinical data accompanies the manuscript and is available as Extended Data and Supplementary Information.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Clinical sequencing data was comprised of 34,650 tumor and matched normal specimens from 31,359 patients prospectively characterized as part of their active care at Memorial Sloan Kettering Cancer Center (MSKCC) between Jan. 2014 and Apr. 2019. Sequencing data from 10,908 primary untreated cancers of The Cancer Genome Atlas cohort were included for comparative frequency analyses (data acquired from <a href="https://gdc.cancer.gov/about-data/publications/pancanatlas">https://gdc.cancer.gov/about-data/publications/pancanatlas</a> ).
Data exclusions	No exclusion criteria other than including solid tumors were specified for the study population
Replication	Experimental replication was performed as described in the Methods section, which included 3 to 5 replicates per condition, and all attempts at replication were successful.
Randomization	Data were randomized for permutation-based statistical testing as described in the Methods section. No other randomized allocation among groups was performed, and all further allocation was based on stated variables and conditions.
Blinding	Blinding was not applicable for this study/analytical design.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	Antibodies
<input type="checkbox"/>	Eukaryotic cell lines
<input checked="" type="checkbox"/>	Palaeontology
<input type="checkbox"/>	Animals and other organisms
<input type="checkbox"/>	Human research participants
<input checked="" type="checkbox"/>	Clinical data

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	MRI-based neuroimaging

## Eukaryotic cell lines

Policy information about [cell lines](#)

### Cell line source(s)

Cell lines utilized here included: A375, Sk-Mel2, Sk-Mel30 (kindly provided by the N. Rosen and T. Merghoub laboratories at MSK), HEK293T (obtained from ATCC, CRL-3216), and murine KP lung adenocarcinoma cells (Kras G12D+/+, Trp53-/-; provided by the T. Jacks laboratory, MIT).

### Authentication

All cell lines have been authenticated by short tandem repeat analysis.

### Mycoplasma contamination

All cell lines were confirmed tested negative for mycoplasma contamination

### Commonly misidentified lines (See [ICLAC](#) register)

No commonly mis-identified cell lines were utilized.

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

### Laboratory animals

Mice used in experiments were Hsd:Athymic Nude-Foxn1nu strain purchased from Envigo (stock #069), 6-8 weeks old, female.

### Wild animals

No wild animals were used in this study.

Field-collected samples	No field-collected samples were used in this study.
Ethics oversight	All mouse experiments were approved by the Memorial Sloan-Kettering Cancer Center (MSKCC) Internal Animal Care and Use Committee

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	Age at time of sequencing: median 61.7 years Male/female: 46.4%/53.6% Additional details in Supplementary Table 1.
Recruitment	Passive recruitment were for patients who underwent prospective sequencing as part of their active clinical care at Memorial Sloan Kettering Cancer Center (MSKCC) from January 2014 to April 2019. All such patients whose tumor sequencing was performed with a matched normal sample were included and biases include only those related to the demographic composition of the catchment area for cancer patients at MSKCC.
Ethics oversight	MSKCC Institutional Review Board

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	NCT01775072
Study protocol	Details available at ClinicalTrials.gov #NCT01775072 or upon request.
Data collection	Locale of data collection: Memorial Sloan Kettering Cancer Center and affiliate sites. Dates of recruitment for prospectively characterized patients utilized here were from January 2014 to April 2019.
Outcomes	Primary and secondary outcome measures not assessed as part of the present study.

## Article

# Anti-PfGARP activates programmed cell death of parasites and reduces severe malaria

<https://doi.org/10.1038/s41586-020-2220-1>

Received: 22 February 2019

Accepted: 20 February 2020

Published online: 22 April 2020

 Check for updates

Dipak K. Raj<sup>1,2</sup>, Alok Das Mohapatra<sup>1,2</sup>, Anup Jnawali<sup>1,2</sup>, Jenna Zuromski<sup>1,2</sup>, Ambrish Jha<sup>1</sup>, Gerald Cham-Kpu<sup>1,2</sup>, Brett Sherman<sup>1</sup>, Rachel M. Rudlaff<sup>3,4</sup>, Christina E. Nixon<sup>1,2</sup>, Nicholas Hilton<sup>1,19</sup>, Andrew V. Oleinikov<sup>5</sup>, Olga Chesnokov<sup>5</sup>, Jordan Merritt<sup>5</sup>, Sunthorn Pond-Tor<sup>1,2</sup>, Lauren Burns<sup>1,2</sup>, Grant Jolly<sup>2</sup>, Choukri Ben Mamoun<sup>6,7</sup>, Edward Kabyemela<sup>8,9,10</sup>, Atis Muehlenbachs<sup>11</sup>, Lynn Lambert<sup>12</sup>, Sachy Orr-Gonzalez<sup>12</sup>, Nina F. Gnädig<sup>13</sup>, David A. Fidock<sup>13,14</sup>, Sangshin Park<sup>1,15</sup>, Jeffrey D. Dvorin<sup>3,4</sup>, Norbert Pardi<sup>16</sup>, Drew Weissman<sup>16</sup>, Barbara L. Mui<sup>17</sup>, Ying K. Tam<sup>17</sup>, Jennifer F. Friedman<sup>118</sup>, Michal Fried<sup>12,20</sup>, Patrick E. Duffy<sup>12,20</sup> & Jonathan D. Kurtis<sup>1,2,20</sup> 

Malaria caused by *Plasmodium falciparum* remains the leading single-agent cause of mortality in children<sup>1</sup>, yet the promise of an effective vaccine has not been fulfilled. Here, using our previously described differential screening method to analyse the proteome of blood-stage *P. falciparum* parasites<sup>2</sup>, we identify *P. falciparum* glutamic-acid-rich protein (PfGARP) as a parasite antigen that is recognized by antibodies in the plasma of children who are relatively resistant—but not those who are susceptible—to malaria caused by *P. falciparum*. PfGARP is a parasite antigen of 80 kDa that is expressed on the exofacial surface of erythrocytes infected by early-to-late-trophozoite-stage parasites. We demonstrate that antibodies against PfGARP kill trophozoite-infected erythrocytes in culture by inducing programmed cell death in the parasites, and that vaccinating non-human primates with PfGARP partially protects against a challenge with *P. falciparum*. Furthermore, our longitudinal cohort studies showed that, compared to individuals who had naturally occurring anti-PfGARP antibodies, Tanzanian children without anti-PfGARP antibodies had a 2.5-fold-higher risk of severe malaria and Kenyan adolescents and adults without these antibodies had a twofold-higher parasite density. By killing trophozoite-infected erythrocytes, PfGARP could synergize with other vaccines that target parasite invasion of hepatocytes or the invasion of and egress from erythrocytes.

To identify novel vaccine candidates for malaria caused by *P. falciparum*, we pooled plasma collected at two years of age from the most-resistant individuals and the most-susceptible individuals who participated in our Tanzanian birth cohort<sup>3</sup> (Supplementary Table 1) and performed differential biopanning experiments on a *P. falciparum* 3D7 strain blood-stage cDNA library constructed in bacteriophage (hereafter phage) T7. After differentially biopanning  $1.0 \times 10^8$  recombinant phages and sequencing  $n=100$  differentially recognized clones, we identified 11 parasite genes the protein products of which were uniquely recognized by antibodies in plasma from resistant—but not susceptible—individuals (Supplementary Table 2).

On the basis of its *in silico* properties, its high degree of enrichment (44 out of 100 differentially biopanned clones) and its representation by

clones derived from three overlapping but distinct cDNAs, we focused our attention on PF3D7\_0113000.

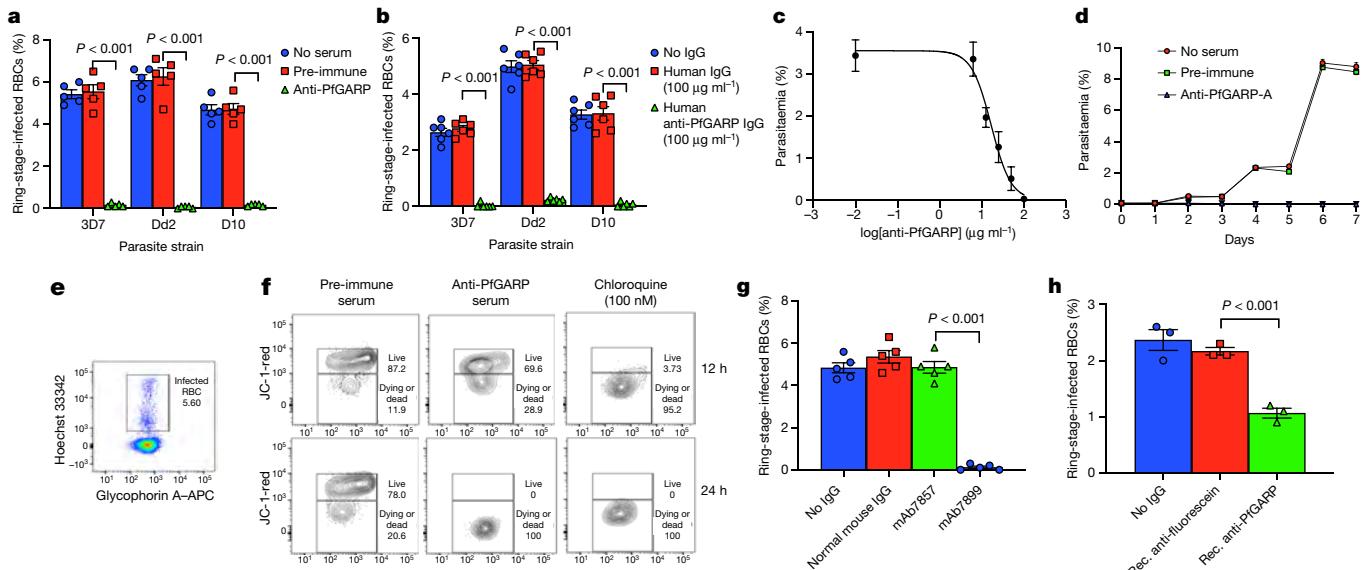
An *in silico* analysis using PlasmoDB<sup>4</sup> ([www.plasmodb.org](http://www.plasmodb.org)) and OrthoMCL-DB<sup>5</sup> ([www.orthomcl.org](http://www.orthomcl.org)) predicted that PF3D7\_0113000 contains a 2,236-bp gene (*PfGARP*) that encodes an 80-kDa acidic protein with a functional *Plasmodium* export element (PEXEL) motif (Supplementary Information, Extended Data Fig. 10), and that has one intron near its 5' end. *PfGARP* has syntenic orthologues in *Plasmodium praefalciparum*, *Plasmodium gaboni* and *Plasmodium reichenowi*, but no orthologues in any other species of *Plasmodium* or in any other organism that has been examined so far. *PfGARP* has no homology with proteins of known function and contains several complex repeat

<sup>1</sup>Center for International Health Research, Rhode Island Hospital, Brown University Medical School, Providence, RI, USA. <sup>2</sup>Department of Pathology and Laboratory Medicine, Brown University Medical School, Providence, RI, USA. <sup>3</sup>Division of Infectious Diseases, Boston Children's Hospital, Boston, MA, USA. <sup>4</sup>Department of Pediatrics, Harvard Medical School, Boston, MA, USA.

<sup>5</sup>Charles E. Schmidt College of Medicine, Florida Atlantic University, Boca Raton, FL, USA. <sup>6</sup>Department of Internal Medicine, Yale University, New Haven, CT, USA. <sup>7</sup>Department of Microbial Pathogenesis, Yale University, New Haven, CT, USA. <sup>8</sup>Mother Offspring Malaria Studies (MOMS) Project, Seattle Biomedical Research Institute, Seattle, WA, USA. <sup>9</sup>Muheza Designated District Hospital, Muheza, Tanzania. <sup>10</sup>Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania. <sup>11</sup>Infectious Disease Pathology Branch, Centers for Disease Control and Prevention, Atlanta, GA, USA. <sup>12</sup>Laboratory of Malaria Immunology and Vaccinology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD, USA. <sup>13</sup>Department of Microbiology and Immunology, Columbia University Irving Medical Center, New York, NY, USA. <sup>14</sup>Division of Infectious Diseases, Department of Medicine, Columbia University Irving Medical Center, New York, NY, USA. <sup>15</sup>Graduate School of Urban Public Health, University of Seoul, Seoul, South Korea. <sup>16</sup>Department of Medicine, University of Pennsylvania, Philadelphia, PA, USA.

<sup>17</sup>Acuitas Therapeutics, Vancouver, British Columbia, Canada. <sup>18</sup>Department of Pediatrics, Rhode Island Hospital, Brown University Medical School, Providence, RI, USA. <sup>19</sup>Present address: Department of Microbiology and Molecular Genetics, University of California, Davis, Davis, CA, USA. <sup>20</sup>These authors contributed equally: Michal Fried, Patrick E. Duffy, Jonathan D. Kurtis.

<sup>✉</sup>e-mail: Jonathan\_Kurtis@Brown.edu



**Fig. 1 | Antibodies to PfGARP inhibit parasite growth.** **a, b**, Mouse anti-PfGARP (1:10 dilution) (green; **a**) or human anti-rPfGARP IgG (100 µg ml⁻¹) purified from pooled serum collected from individuals who had been exposed to malaria (green; **b**) inhibits the growth of three different strains of *P. falciparum* (3D7, Dd2 and D10) by 94–99%. Negative controls were no serum (blue) and pre-immune mouse serum (red) (**a**); or control medium with no IgG (blue) and human IgG purified from human serum obtained from malaria-naïve individuals (red) (**b**). Data are mean ± s.e.m. of five biologically independent replicates. **c**, The IC<sub>50</sub> of anti-PfGARP purified from sera from vaccinated mice is approximately 16.8 µg ml⁻¹. Data are mean ± s.d. of three biologically independent replicates. **d**, Anti-PfGARP kills *P. falciparum* parasites in long-term cultures. Data are mean ± s.e.m. of three biologically independent replicates. **e, f**, Anti-PfGARP disrupts mitochondrial membrane potential. RBCs infected with ring-stage 3D7 *P. falciparum* were incubated with pre-immune or anti-PfGARP serum (1:10 dilution) or with chloroquine

(100 nM). **e**, Gating strategy for **f**. **f**, Contour plots showing infected RBCs with live parasites in the upper gate and dying or dead parasites in the lower gate (defined by JC-1 staining). Data are representative of three independent experiments. **g, h**, Monoclonal anti-PfGARP inhibits parasite growth. **g**, RBCs infected with ring-stage 3D7 *P. falciparum* were cultured with medium alone, normal mouse IgG (1 mg ml⁻¹) or anti-PfGARP monoclonal antibodies (mAb7857 or mAb7899; 1 mg ml⁻¹). **h**, Recombinant monoclonal anti-PfGARP produced from the sequences of the heavy- and light-chain variable regions of mAb7899 inhibits parasite growth. RBCs infected with ring-stage 3D7 *P. falciparum* were cultured with medium alone, recombinant (rec.) anti-fluorescein (250 µg ml⁻¹) or recombinant anti-PfGARP monoclonal antibody (250 µg ml⁻¹). Data are mean ± s.e.m. of three biologically independent replicates. All *P* values were calculated by two-sided non-parametric Mann–Whitney *U*-test. Data in **a–h** are representative of five independent experiments. Data in **b–h** are representative of three independent experiments.

regions and extensive regions of low amino-acid complexity; 50% of the protein is composed of three amino acids (Lys, Glu and Asp).

Expression of *PfGARP* is highly restricted to the early-trophozoite stage of the parasite life cycle<sup>6</sup> and the gene shows minimal sequence variation in the immunorelevant region—the region encoded by the largest clone identified in resistant sera from our differential screens (nucleotides 1,222–2,022). A previous deep-sequencing analysis of 227 field samples identified only one non-synonymous single-nucleotide polymorphism (SNP) in the immunorelevant region<sup>7</sup>, and an expanded (as yet unpublished) analysis of 3,248 field samples has provisionally identified 15 non-synonymous SNPs in this region (<https://www.malariaigen.net/apps/pf/4.0/#doc=DocAboutData.htm>).

### Anti-PfGARP-A kills parasites in vitro

We expressed and purified the polypeptide encoded by the differentially recognized, immunorelevant region of PfGARP from the referent *P. falciparum* 3D7 strain (nucleotides 1,228–2,022; amino acids 410–673) in *Escherichia coli* and designated this recombinant protein rPfGARP-A (Extended Data Fig. 1a). In addition, we cloned this immunorelevant region into a eukaryotic expression plasmid (VR2001). To generate anti-PfGARP-A antiserum, we immunized mice either with the recombinant protein (rPfGARP-A) in TiterMax adjuvant, or with the eukaryotic expression plasmid. In a western blot analysis, both the anti-PfGARP-A antiserum that was generated by immunization with the recombinant protein and the antiserum generated by immunization with DNA recognized a protein of around 100 kDa in trophozoite-infected red blood cells (RBCs) (Extended Data Fig. 1b, k). This higher apparent molecular weight is consistent with the acidic composition of PfGARP<sup>8</sup>.

We performed growth-inhibition assays (GIAs) using anti-PfGARP-A antiserum that was prepared by immunizing mice with either DNA or recombinant protein. Anti-PfGARP-A inhibited parasite growth by 94–99% compared to controls in three parasite strains and four freshly isolated parasite lines (all *P* < 0.001; Fig. 1a, Extended Data Fig. 1d). Anti-PfGARP-A-treated parasites displayed a dysmorphic, pyknotic appearance on Giemsa-stained blood smears, consistent with ‘crisis’ forms that are associated with dying or dead parasites<sup>9,10</sup> (Extended Data Fig. 1e–h).

We purified human polyclonal anti-PfGARP-A antibodies from plasma pooled from adults living in our Tanzanian field site (Extended Data Fig. 1i, j), and found that these human anti-PfGARP-A antibodies significantly inhibited parasite growth by 94–99% compared with controls in three parasite strains (all *P* < 0.001, Fig. 1b). In addition, we purified mouse polyclonal anti-PfGARP-A antibodies from plasma pooled from rPfGARP-A-immunized mice (Extended Data Fig. 1k, l), and calculated that the half-maximum inhibitory concentration (IC<sub>50</sub>) for parasite growth was 16.8 µg of anti-PfGARP-A per ml of culture medium (Fig. 1c).

In long-term cultures, anti-PfGARP-treated parasites never expanded in number (Fig. 1d), and by day two they had decreased in size and appeared pyknotic. By day four, the parasites had become small, pyknotic dots that were difficult to visualize, and by day six, they were no longer identifiable as parasite-infected RBCs. These data indicate that anti-PfGARP antibodies kill parasite-infected RBCs even in the absence of complement or cellular effector functions.

We performed GIA assays and quantified the viability of parasites by flow cytometry using JC-1, a mitochondrial membrane potential probe. Ring-stage parasites that were treated with anti-PfGARP antibodies showed a marked loss of mitochondrial membrane potential

(a characteristic of programmed cell death) within 12 h, and almost all parasites had lost their mitochondrial function within 24 h (Fig. 1e, f).

We produced a series of monoclonal antibodies in mice that were immunized with rPfGARP-A. Of the sixteen monoclonal antibodies that reacted with PfGARP-A in an enzyme-linked immunosorbent assay (ELISA), only mAb7899 killed parasites in culture (Fig. 1g). We sequenced and expressed the heavy-chain and light-chain variable regions of mAb7899. The resulting recombinant monoclonal antibody had a dissociation constant ( $K_d$ ) of 2.9 nM (95% confidence interval (CI) [1.3, 5.9]; Extended Data Fig. 2a), inhibited parasite growth by 51% at a concentration of 250  $\mu\text{g ml}^{-1}$  in GIA assays (Fig. 1h) and recognized amino acids 443–459 (VKNVIEDEDKDGVEIIN) of full-length PfGARP (Extended Data Fig. 2b). Consistent with the loss of mitochondrial function, the levels of lactate in the medium of cultures treated with recombinant mAb7899 were significantly higher than those in control cultures (Extended Data Fig. 2c). A monovalent antigen-binding fragment (Fab) of mAb7899 inhibited parasite growth by 76–87% across three strains of *P. falciparum* (Extended Data Fig. 2d). These data confirm that anti-PfGARP-mediated killing occurs in the absence of complement, cellular effector functions or antigen cross-linking.

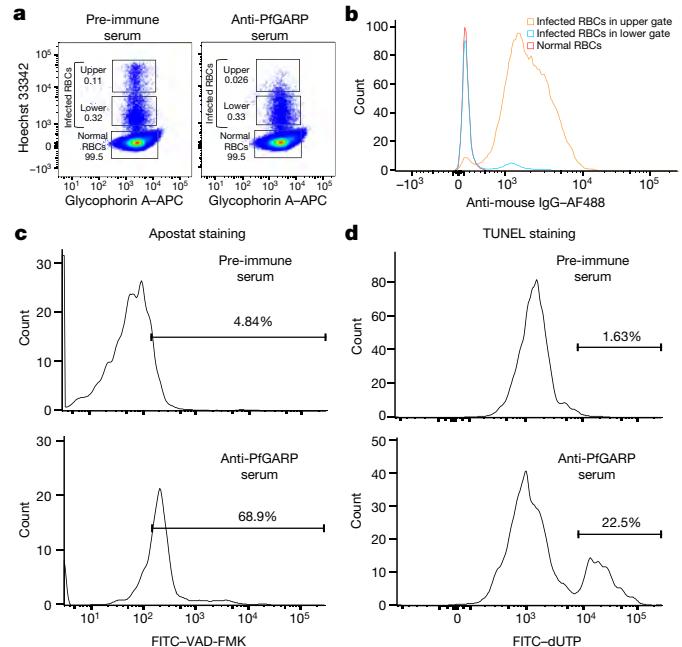
We used the TetR system<sup>11</sup> to construct parasites with a conditional knockdown of PfGARP (3D7-PfGARP KD) (Extended Data Fig. 3a, Supplementary Table 3). The 3D7-PfGARP KD parasites, when grown in the absence of the inducer anhydrotetracycline, did not exhibit an overt growth phenotype in vitro, despite showing a reduction of up to 90% in the levels of PfGARP protein (Extended Data Fig. 3b–e). In GIA assays, the killing efficacy of anti-PfGARP antibodies was only mildly reduced in 3D7-PfGARP KD parasites cultured without anhydrotetracycline compared to those cultured with it ( $P = 0.014$ ; Extended Data Fig. 3f). Anti-PfGARP-mediated killing of 3D7-PfGARP KD parasites cultured without anhydrotetracycline may be due to the residual expression of PfGARP even in the absence of the inducer. To test this hypothesis, we constructed parasites in which the *PfGARP* gene was deleted (3D7-PfGARP KO) (Extended Data Fig. 4a–d), and evaluated them in growth and GIA assays. The 3D7-PfGARP KO parasites did not exhibit an overt growth phenotype in vitro (Extended Data Fig. 4e). As expected, the killing efficacy of anti-PfGARP antibodies was completely abrogated in the 3D7-PfGARP KO compared to the wild-type parasites (Extended Data Fig. 4f).

## Immunolocalization of PfGARP

To assess the immunolocalization of PfGARP, we used immunofluorescence confocal microscopy and immunogold transmission electron microscopy in 3D7 *P. falciparum* parasites probed with anti-PfGARP. We also evaluated 3D7-PfGARP KD parasites, in which PfGARP is linked to a C-terminal V5 tag, probed with anti-V5 antibodies. In all experiments, PfGARP was localized to the exofacial surface of the RBC membrane in early-to-late-trophozoite-infected RBCs (Extended Data Figs. 5–7).

## Anti-PfGARP disrupts food vacuole integrity

In transmission electron microscopy studies, the food vacuole in anti-PfGARP treated parasites was markedly diminished in size compared to control-treated parasites and appeared absent or condensed tightly around hemozoin crystals—an effect that has not been reported for any previous antimalarial antibodies (Extended Data Fig. 8a, b). We explored the possibility that PfGARP also localized to the membrane of the food vacuole by colocalization studies with the *P. falciparum* chloroquine resistance transporter (PfCRT). In the majority of infected RBCs, PfCRT and PfGARP did not colocalize (Extended Data Fig. 7d). We assessed the effect of anti-PfGARP by confocal microscopy using the calcium-binding dye Fluo-4 AM, which preferentially labels the food vacuole<sup>12</sup>. Parasites treated with anti-PfGARP showed a marked loss of integrity of the food vacuole, as evidenced by the redistribution of Fluo-4 AM (and therefore calcium) from the food vacuole<sup>12</sup> to the cytosol (Extended Data Fig. 8c). We further evaluated the effect of anti-PfGARP antibodies on



**Fig. 2 | Antibodies to PfGARP bind to *P. falciparum*-infected RBCs, leading to the activation of caspase-like proteases and the fragmentation of parasite DNA.** **a**, RBCs infected with ring-stage 3D7 *P. falciparum* were cultured for 24–30 h in the presence of 10% mouse pre-immune or anti-PfGARP serum that was generated by vaccination with DNA. Representative dot plots show infected RBCs with a higher DNA content (mature parasites) in the upper gate, infected RBCs with a lower DNA content (immature and dying parasites) in the middle gate and uninfected RBCs in the lowest gate. **b**, Binding of polyclonal anti-PfGARP antibodies to the infected RBCs in the upper two gates and the uninfected RBCs in the lowest gate from **a**. **c, d**, Incubation of infected RBCs with anti-PfGARP results in the activation of caspase-like proteases (assessed by Apostat staining; **c**) and DNA fragmentation (assessed by TUNEL staining; **d**) in parasites that were present in the upper gate. Percentages in **c, d** indicate the proportion of gated events occurring within the range bounded by the horizontal line. All data are representative of three independent experiments.

food vacuole integrity using serial block-face scanning electron microscopy (SBF-SEM), which demonstrated that treatment with anti-PfGARP resulted in a complete loss of food vacuole integrity, with hemozoin dispersed broadly throughout the parasitophorous vacuole (Supplementary Video 1).

## Anti-PfGARP activates programmed cell death

The effect of anti-PfGARP antibodies on parasite morphology, mitochondrial membrane potential and food vacuole integrity (with increased intracellular levels of calcium) suggests parasites are killed through the activation of programmed cell death. We therefore analysed anti-PfGARP-treated parasites for the activation of caspase-like enzymes. We performed GIA assays and probed the parasites with Apostat, a fluorescent dye that labels activated-caspase-like cysteine proteases. Parasites that were treated with anti-PfGARP showed a marked increase in the activation of caspase-like proteases compared to untreated parasites, as assessed by flow cytometry (Fig. 2a–c). We also performed TUNEL (terminal deoxynucleotidyl transferase dUTP nick-end labelling) assays to determine whether treatment with anti-PfGARP resulted in the fragmentation of parasite DNA. Parasites that were treated with anti-PfGARP showed substantial fragmentation of their DNA, as assessed by flow cytometry (Fig. 2d).

## Tanzanian birth cohort study

To investigate the effects of naturally acquired anti-PfGARP antibodies on clinical malaria, we measured the levels of anti-PfGARP IgG antibodies in

# Article

members of our Tanzanian birth cohort using a fluorescent, bead-based assay, and related the results to subsequent malaria outcomes.

We measured the levels of anti-PfGARP IgG antibodies in plasma that was obtained from 246 children at the age of 48 weeks. The average duration of follow-up was 64 weeks per child. Anti-PfGARP antibodies were detectable in 48.8% of these samples and children were followed for a total of 15,737 child weeks of observation.

We used generalized estimating equation (GEE)-based repeated measures models to evaluate the relationship between the levels of anti-PfGARP IgG and the risk of severe malaria. When analysed as a continuous variable, anti-PfGARP IgG levels predicted a significantly decreased risk of severe malaria over the follow-up period ( $P=0.008$ ). When analysed dichotomously, individuals with undetectable anti-PfGARP IgG antibodies ( $n=126$  individuals who contributed 7,327 weeks of follow-up) had a 2.5-fold-higher risk of severe malaria than individuals with detectable anti-PfGARP IgG antibodies ( $n=120$  individuals who contributed 8,410 weeks of follow-up; 95% CI [1.2, 5.5],  $P=0.018$ ) (Fig. 3a). These results remained significant (odds ratio = 2.5, 95% CI [1.1, 5.5],  $P=0.026$ ) even after adjusting for potential confounders (see Supplementary Methods for the modelling approach).

## Kenyan cohort study

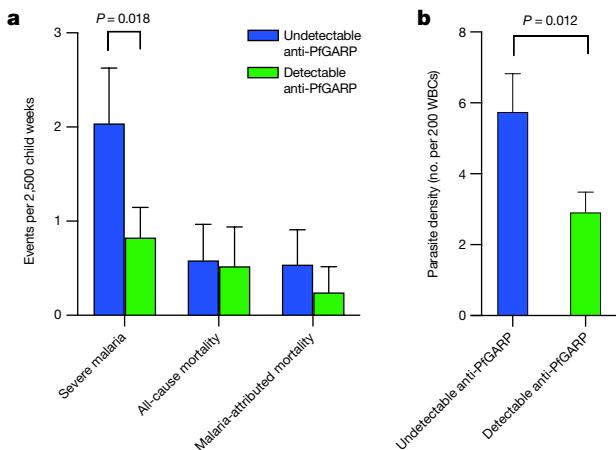
To generalize these results to a completely independent cohort, we measured anti-rPfGARP-A IgG responses in a cohort of Kenyan male individuals who participated in a treatment-reinfection study<sup>13–15</sup>. Volunteers ( $n=135$ ) aged 12–35 years were entered into the study at the beginning of the high-transmission season in April 1997, treated for malaria and followed with weekly blood smears for 18 weeks. Serum was collected two weeks after treatment and stored at  $-80^{\circ}\text{C}$ . In this age group, clinical or severe malaria is very uncommon.

When analysed as a continuous variable in GEE-based repeated measures models, higher anti-PfGARP IgG levels predicted a significantly decreased parasite density over 18 weeks of follow-up ( $P<0.004$ ). When analysed dichotomously, individuals with undetectable anti-rPfGARP-A IgG antibodies ( $n=61$  individuals who contributed 1,018 weeks of follow-up blood smears) had 1.97-fold higher densities of parasites over the 18 weeks of follow-up than individuals who had detectable anti-rPfGARP-A IgG antibodies ( $n=74$  individuals who contributed 1,237 weeks of follow-up blood smears; 95% CI [0.94, 4.23],  $P=0.012$ ) (Fig. 3b). These results remained significant (odds ratio = 1.82, 95% CI [0.9, 3.7],  $P=0.019$ ) after adjusting for potential confounders (see Supplementary Methods).

## Monkey vaccine trials

Because PfGARP is only found in *Plasmodium* species that infect primates, we conducted a vaccine trial in the *P.falciparum/Aotus* model, in which *Aotus* monkeys are infected with *P.falciparum*. Nucleoside-modified mRNA vaccines (purified by fast protein liquid chromatography (FPLC)) against the Zika and influenza viruses have been shown to have potent effects in small and large animals<sup>16–18</sup>. We therefore immunized  $n=5$  monkeys intradermally with 50  $\mu\text{g}$  of nucleoside-modified mRNA encoding PfGARP-A, encapsulated in lipid nanoparticles (PfGARP-A-mRNA LNPs). Control monkeys ( $n=4$ ) were intradermally immunized with 50  $\mu\text{g}$  of LNPs containing poly(C) RNA (poly(C)-RNA LNPs). The monkeys received three doses at three-week intervals, and before each dose, serum was obtained for antibody assays. On day 63, the monkeys were challenged by intravenous injection with  $1 \times 10^4$  RBCs infected with blood-stage *P.falciparum* FVO strain parasites, and blood films were prepared daily. This represents a heterologous challenge with a highly virulent parasite as the sequence for the PfGARP-A-mRNA LNP vaccine was derived from the 3D7 strain.

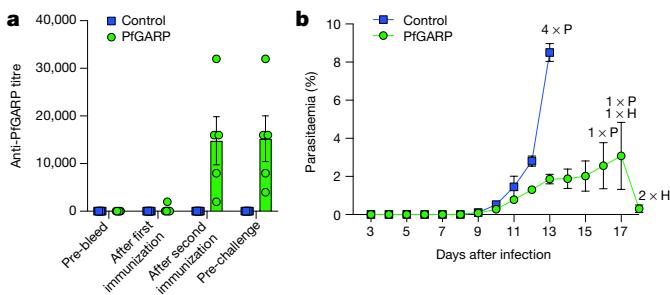
Immunized monkeys generated antibody responses that plateaued after the second injection (Fig. 4a). Control monkeys had significantly



**Fig. 3 | Antibodies to PfGARP predict a reduced risk of severe malaria and parasitaemia.** **a**, Tanzanian children with undetectable anti-PfGARP IgG antibodies at the age of 48 weeks ( $n=126$  individuals who contributed 7,327 weeks of follow-up) had a 2.5-fold-higher risk of developing severe malaria than children who had detectable anti-PfGARP IgG antibodies at this age ( $n=120$  individuals who contributed 8,410 weeks of follow-up; 95% CI [1.2, 5.5],  $P=0.018$ , GEE model with two-sided fixed-effect test). Bars represent the least-square mean number of events per 2,500 child weeks, with s.e.m. adjusted for repeated measures. **b**, Kenyan male individuals with undetectable anti-rPfGARP-A IgG antibodies ( $n=61$  individuals who contributed 1,018 weeks of follow-up blood smears) had 1.97-fold-higher parasite densities over 18 weeks of follow-up than did individuals with detectable IgG anti-rPfGARP-A antibodies ( $n=74$  individuals who contributed 1,237 weeks of follow-up blood smears; 95% CI [0.94, 4.23],  $P=0.012$ , GEE model with two-sided fixed-effect test). Bars depict the least-square mean parasite density (measured as the number of asexual-stage parasites per 200 white blood cells (WBCs)) with s.e.m.

higher parasitaemia on day 12 than monkeys that were immunized with PfGARP-A ( $P<0.009$ ), with 4.6-fold-higher parasitaemia on day 13, the final day with complete follow-up of all monkeys ( $P<0.001$ , Fig. 4b). On day 13, all of the control monkeys met pre-specified criteria for antimalarial treatment (parasitaemia  $> 7.5\%$ ). Of the five PfGARP-A-vaccinated monkeys, one was treated for hyperparasitaemia on day 16 and one for hyperparasitaemia on day 17; one was treated for anaemia—a common complication in the *Aotus/P.falciparum* model<sup>19</sup>—on day 17; and the two remaining monkeys were treated for anaemia on day 18, despite controlling their parasitaemia (Extended Data Fig. 9a). Notably, the three monkeys that controlled their parasitaemia and required treatment for anaemia had higher anti-PfGARP titres (16,000, 16,000 and 32,000) compared to the monkeys that required treatment for their parasitaemia (8,000 and 4,000).

We also conducted a trial in the *P.falciparum/Aotus* model using rPfGARP-A (expressed in *E.coli*) emulsified in Ribi adjuvant as the immunogen. We immunized monkeys subcutaneously with 50  $\mu\text{g}$  of either PfGARP-A in Ribi adjuvant ( $n=4$  monkeys) or Ribi alone ( $n=5$  monkeys). The monkeys received three doses at three-week intervals, and before each dose, samples of serum were obtained for antibody assays. On day 63, the monkeys were challenged by intravenous injection with  $1 \times 10^4$  RBCs infected with blood-stage *P.falciparum* FVO strain parasites, and blood films were produced daily. Immunized monkeys generated antibody responses that increased after the third injection (Extended Data Fig. 9b). Notably, control monkeys had significantly higher parasitaemia on days 7–12 than PfGARP-A-vaccinated monkeys (all  $P<0.05$ ; Extended Data Fig. 9c, d). On day 11—the final day with complete follow-up of all monkeys—control monkeys had 3.5-fold-higher parasitaemia than PfGARP-A-vaccinated monkeys ( $P<0.01$ ). Four control monkeys met pre-specified criteria for drug treatment (parasitaemia greater than 7.5%, haematocrit lower than 25% or evidence of clinical illness) on day 11, and the remaining control monkey met these criteria



on day 12. On day 11, one PfGARP-vaccinated monkey underwent drug treatment despite not meeting the pre-specified criteria.

## Summary and conclusions

Malaria remains a leading cause of childhood mortality and vaccines are urgently needed to attenuate this threat to public health. Using our vaccine discovery platform, we identified PfGARP—an antigen that is localized to the exofacial surface of the RBC membrane in trophozoite-infected RBCs—as a vaccine candidate targeting the blood stage of *P. falciparum*. Anti-PfGARP markedly attenuates the growth of parasites by arresting and killing trophozoite-infected RBCs in the absence of immune effector cells or complement.

Parasites that were treated with anti-PfGARP displayed several canonical features of programmed cell death, including shrunken, pyknotic nuclei; a loss of mitochondrial membrane integrity; the activation of caspase-like proteases; DNA fragmentation; and the release of calcium from intracellular stores. Although *P. falciparum* lacks classic caspases, it does encode three meta-caspases<sup>20</sup>, and activation of *P. falciparum* metacaspase 1 (PfMCA1) functions as an upstream activator of a caspase-like enzyme that leads to programmed cell death<sup>21</sup>. This ability to induce programmed cell death in *Plasmodium* parasites is a novel mode of action for an antimalarial antibody. Because PfGARP is located on the exofacial surface of RBCs, and because antibody engagement of PfGARP leads to the activation of programmed cell death in parasites, we speculate that PfGARP might function in the density-dependent regulation of parasitaemia by sensing either parasite or host factors<sup>22–24</sup>.

In non-human primates, immunization with PfGARP as either an mRNA-based or a recombinant-protein-based immunogen conferred marked protection against parasitaemia compared with controls after a heterologous challenge with *P. falciparum*. This is the first report, to our knowledge, of an mRNA-based vaccine that induces protection against *P. falciparum*.

In longitudinal cohort studies, naturally occurring anti-PfGARP predicted a decreased risk of severe malaria in children and decreased parasitaemia in adolescents and adults. Together, our data support PfGARP as a vaccine candidate against malaria caused by *P. falciparum*.

Furthermore, our recombinant monoclonal anti-PfGARP antibody could serve as a platform for the development of therapeutic and prophylactic antibody-based interventions, and could also be used in high-throughput drug screens that target PfGARP-induced programmed cell death. By killing trophozoite-infected RBCs, immunization with PfGARP could synergize with other vaccines that target the parasite invasion of hepatocytes<sup>25</sup> or the invasion of and egress from erythrocytes<sup>2,26</sup>.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2220-1>.

1. World Health Organization. *World Health Statistics 2018: Monitoring Health for the SDGs* <https://www.who.int/data/gho/publications/world-health-statistics> (WHO, 2018).
2. Raj, D. K. et al. Antibodies to PFSEA-1 block parasite egress from RBCs and protect against malaria infection. *Science* **344**, 871–877 (2014).
3. Mutabingwa, T. K. et al. Maternal malaria and gravidity interact to modify infant susceptibility to malaria. *PLoS Med.* **2**, e407 (2005).
4. Aurrecoechea, C. et al. PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Res.* **37**, D539–D543 (2009).
5. Chen, F., Mackey, A. J., Stoeckert, C. J. Jr & Roos, D. S. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res.* **34**, D363–D368 (2006).
6. López-Barragán, M. J. et al. Directional gene expression and antisense transcripts in sexual and asexual stages of *Plasmodium falciparum*. *BMC Genomics* **12**, 587 (2011).
7. Manske, M. et al. Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing. *Nature* **487**, 375–379 (2012).
8. Tiwari, P., Kaita, P. & Guptasarma, P. Understanding anomalous mobility of proteins on SDS-PAGE with special reference to the highly acidic extracellular domains of human E- and N-cadherins. *Electrophoresis* **40**, 1273–1281 (2019).
9. Ockenhouse, C. F., Schulman, S. & Shear, H. L. Induction of crisis forms in the human malaria parasite *Plasmodium falciparum* by gamma-interferon-activated, monocyte-derived macrophages. *J. Immunol.* **133**, 1601–1608 (1984).
10. Jensen, J. B., Boland, M. T. & Aklood, M. Induction of crisis forms in cultured *Plasmodium falciparum* with human immune serum from Sudan. *Science* **216**, 1230–1233 (1982).
11. Ganesan, S. M., Falla, A., Goldfless, S. J., Nasamu, A. S. & Niles, J. C. Synthetic RNA–protein modules integrated with native translation mechanisms to control gene expression in malaria parasites. *Nat. Commun.* **7**, 10727 (2016).
12. Ch'Ng, J. H., Liew, K., Goh, A. S., Sidhartha, E. & Tan, K. S. Drug-induced permeabilization of parasite's digestive vacuole is a key trigger of programmed cell death in *Plasmodium falciparum*. *Cell Death Dis.* **2**, e216 (2011).
13. Kurtis, J. D., Lanar, D. E., Oppollo, M. D. & Duffy, P. E. Interleukin-10 responses to liver-stage antigen 1 predict human resistance to *Plasmodium falciparum*. *Infect. Immun.* **67**, 3424–3429 (1999).
14. Kurtis, J. D., Matalib, R., Onyango, F. K. & Duffy, P. E. Human resistance to *Plasmodium falciparum* increases during puberty and is predicted by dehydroepiandrosterone sulfate levels. *Infect. Immun.* **69**, 123–128 (2001).
15. Nixon, C. P. et al. Antibodies to rhoptry-associated membrane antigen predict resistance to *Plasmodium falciparum*. *J. Infect. Dis.* **192**, 861–869 (2005).
16. Pardi, N. et al. Nucleoside-modified mRNA vaccines induce potent T follicular helper and germinal center B cell responses. *J. Exp. Med.* **215**, 1571–1588 (2018).
17. Pardi, N. et al. Zika virus protection by a single low-dose nucleoside-modified mRNA vaccination. *Nature* **543**, 248–251 (2017).
18. Pardi, N. et al. Nucleoside-modified mRNA immunization elicits influenza virus hemagglutinin stalk-specific antibodies. *Nat. Commun.* **9**, 3361 (2018).
19. Egan, A. F., Fabucci, M. E., Saul, A., Kaslow, D. C. & Miller, L. H. *Aotus* New World monkeys: model for studying malaria-induced anemia. *Blood* **99**, 3863–3866 (2002).
20. Vandana, D., Dixit, R., Tiwari, R., Katyal, A. & Pandey, K. C. Metacaspases: potential drug target against protozoan parasites. *Front. Pharmacol.* **10**, 790 (2019).
21. Mestlin, B., Beavogui, A. H., Fasel, N. & Picot, S. *Plasmodium falciparum* metacaspase PfMCA-1 triggers a z-VAD-fmk inhibitable protease to promote cell death. *PLoS One* **6**, e23867 (2011).
22. Mutai, B. K. & Waitumbi, J. N. Apoptosis stalks *Plasmodium falciparum* maintained in continuous culture condition. *Malar. J.* **9**, S6 (2010).
23. Engelbrecht, D. & Coetzer, T. L. *Plasmodium falciparum* exhibits markers of regulated cell death at high population density in vitro. *Parasitol. Int.* **65**, 715–727 (2016).
24. Chou, E. S. et al. A high parasite density environment induces transcriptional changes and cell death in *Plasmodium falciparum* blood stages. *FEBS J.* **285**, 848–870 (2018).
25. RTS,S Clinical Trials Partnership. First results of phase 3 trial of RTS,S/AS01 malaria vaccine in African children. *N. Engl. J. Med.* **365**, 1863–1875 (2011).
26. Crosnier, C. et al. Basigin is a receptor essential for erythrocyte invasion by *Plasmodium falciparum*. *Nature* **480**, 534–537 (2011).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

# Article

## Methods

### Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were blinded to allocation during the non-human primate experiments and outcome assessments. Investigators were blinded to clinical status during blood-smear reading for the human cohort studies.

### Ethical approval

The study complied with all relevant ethical regulations for both animal and human studies. Ethical clearance was obtained from the institutional review boards of Seattle Biomedical Research Institute and Rhode Island Hospital; the Medical Research Coordinating Committee of the National Institute for Medical Research, Tanzania; and the Kenyan Medical Research Institute. Ethical clearance for animal studies was obtained from the relevant review boards of Brown University, Rhode Island Hospital and the US National Institutes of Health (NIH).

### Tanzanian birth cohort

**Study population.** Volunteers participated in the Mother Offspring Malaria Studies (MOMS) project, which is based at Muheza Designated District Hospital, in northeastern Tanzania. Mothers presenting at Muheza Designated District Hospital for delivery were enrolled and provided signed, informed consent for their participation, and that of their newborns, in the study. The entomological inoculation rate in our study site exceeds 400 infectious mosquito bites per person per year<sup>27</sup>. Details of the MOMS study design, enrolment methods and exclusion criteria have been published elsewhere<sup>3,28</sup>.

**Inclusion criteria and clinical monitoring.** We monitored 785 children for *P.falciparum* infection from birth up to 3.5 years of age. Children were evaluated at routine 'well-child' visits by a clinician every two weeks from birth to one year of age, and monthly thereafter, including analysis of blood smears. Routine blood samples were collected once every 6 months from the age of 1.5 to 3.5 years. Blood smears and blood samples were also collected any time the child became sick. Sick children who came to hospitals or mobile clinics were examined by a medical officer. Treatment outside the study was minimized by active, weekly surveillance by our mobile clinics.

Clinical malaria was defined as asexual *P.falciparum* parasitaemia by blood smear, coupled with symptoms suggestive of malaria, such as a temperature higher than 37.5 °C, nausea or vomiting, irritability and poor feeding. Prompt treatment was provided to sick children according to the guidelines of the Tanzanian Ministry of Health, and study participants were instructed to obtain all medications including antimalarial drugs through the project staff.

**Sample collection and processing.** Venous blood was collected and stored at 4 °C until processing. After centrifugation, plasma was stored at -80 °C. *P.falciparum* parasitaemia was determined by Giemsa staining of thick blood smears prepared from the capillary or venous blood. Parasite density was expressed as the number of asexual-stage parasites per 200 white blood cells in the thick smear. Sickle cell trait was determined by electrophoresis (Helena Laboratories). Haemograms were obtained on an impedance-based analyser (Abbott Cell Dyn 1200).

**Case definitions.** Mild malaria was defined as a positive blood smear and one or more of the following: (1) anaemia defined by a blood haemoglobin level of less than 8 g dl<sup>-1</sup>; (2) vomiting; (3) diarrhoeal disease or gastroenteritis; (4) lower respiratory infection; or (5) oral temperature higher than 38 °C.

Severe malaria was defined as a positive blood smear and one or more of the following: (1) respiratory distress defined by respiratory rate exceeding 40 breaths per minute for children older than two months

or 50 breaths per minute for children younger than two months; (2) a history of one or more convulsions in the 24 h before or during hospitalization; (3) prostration defined by inability to sit unaided; (4) hypoglycaemia defined by a blood glucose level of less than 2.2 mmol l<sup>-1</sup>; (5) severe anaemia defined by a blood haemoglobin level of less than 6 g dl<sup>-1</sup>; or (6) oral temperature higher than 40 °C.

Malaria-associated mortality was defined as death with a positive blood film obtained during the terminal illness. One child who died of bacterial meningitis but had a positive blood film was adjudicated as a non-malarial death.

### Kenyan cohort

**Study population.** To generalize the protective nature of anti-PfGARP antibodies, we measured the levels of anti-PfGARP antibody in an entirely distinct longitudinal cohort using epidemiological data and blood samples that were collected in 1997 as part of a treatment-reinfection study<sup>14,29</sup>. Volunteers were residents of subsistence-farming villages in which *P.falciparum* is endemic, in western Kenya north of Lake Victoria. The entomological inoculation rate in this area can exceed 300 infectious bites per year<sup>30</sup>. After informed consent was obtained, 144 male individuals aged 12–35 years were entered into the study at the beginning of the high-transmission season in April 1997. Detectable parasitaemia was eradicated in 143 of the 144 participants with quinine sulfate (10 mg kg<sup>-1</sup> twice daily for 3 days) and doxycycline (100 mg twice daily for 7 days). One volunteer remained parasitaemic during the week after treatment and was removed from the analysis, and five volunteers did not have available serum samples, thus our analytic sample size was *n* = 138. Immunological and epidemiological analyses of this cohort have been reported elsewhere<sup>14,29,31,32</sup>.

**Malaria assessment.** Thick and thin blood smears were obtained from each volunteer before treatment and then weekly for 18 weeks after treatment to quantify reinfection. Each smear was interpreted by two microscopists who were blinded to the clinical status of the participants and the mean of the two values was recorded.

**Entomology measurements.** The abundance of female *Anopheles* mosquitoes was measured weekly for 18 weeks in each volunteer's domicile using the daytime resting indoors (DRI) method<sup>33</sup> as previously described<sup>29</sup>.

**Blood collection.** Two weeks after treatment with quinine and doxycycline, volunteers donated 10 ml of blood into heparinized tubes. Within 4 h of collection, samples were centrifuged and plasma was aliquoted and stored at -80 °C for subsequent analysis.

**Clinical assays.** Sickle cell trait was determined by electrophoresis (Helena Laboratories).

### Selection of resistant and susceptible individuals for differential screening assays

Our overall purpose was to identify acquired differences in antibody repertoire that mediate resistance to parasitaemia. In our cohort, parasitaemia did not decline until after the age of two years (see Supplementary Fig. 1 from a previous study<sup>2</sup>). Any differences in parasitaemia between groups that were detectable in the first two years of life were unlikely to be due to differences in acquired antibody repertoire, as the children had made relatively little specific antibody by this age. Therefore, we wanted to ensure that the susceptibility to parasitaemia was similar (and high) between our resistant and susceptible groups in the first two years of life and then diverged after two years of age. The divergence after the age of two years was unlikely to be due to a constitutive (that is, genetic) feature and more likely to be due to a feature that is acquired (that is, antibody repertoire). Thus, we selected children for our two groups (resistant versus susceptible) to have similar levels of

parasitaemia in the first two years of life, and then to have very divergent levels of parasitaemia from the age of two to four years.

From our Tanzanian birth cohort, we excluded individuals with less than 9 of the total  $n=18$  scheduled monthly blood smears collected between the age of 2 and 3.5 years, individuals with less than 200  $\mu\text{l}$  of plasma available from the plasma sample obtained at age 2 ( $\pm 2$  weeks), and individuals who were parasitaemic when the plasma sample from the age of 2 years ( $\pm 2$  weeks) was obtained. We then ordered individuals by rank on the basis of the mean parasite density on all blood films collected between the age of 2 and 3.5 years. This mean parasite density included the scheduled monthly blood smears as well as positive blood smears that were obtained during sick visits. Individuals from the low and high extremes of this distribution were chosen to comprise the 'resistant' ( $n=12$ ) and 'susceptible' ( $n=14$ ) groups. To minimize differential exposure as a possible confounder, resistant individuals were selected from those children who did not sleep under bed nets, whereas susceptible individuals were selected from those children who did sleep under bed nets. Selections were made with matching based on the village of residence and sex. Potential confounders examined included: haemoglobin phenotype, the presence of placental malaria, maternal age, birth season and number of previous pregnancies (Supplementary Table 1). By matching and demonstrating that potential non-immunological variables influencing resistance (for example, haemoglobin S (HbS)) were not differentially distributed between the resistant and susceptible groups, we substantially reduced the chance that these covariates were confounding the relationship between antibody specificities discovered and the outcome of resistance or susceptibility.

#### Differential screening of whole blood-stage proteome

We constructed a *P.falciparum* blood-stage cDNA expression library prepared in a T7Select 10-3b vector (Invitrogen) using RNA prepared from freshly isolated parasites collected in our Tanzanian field site. This vector displays 5–15 copies of the cloned gene on the surface of phage capsids as a fusion with phage 10B protein.

We bound Immulon 4HB (Thermo Fisher Scientific) ELISA wells with 100  $\mu\text{l}$  of a 1:100 dilution of serum that was pooled from malaria-resistant children ( $n=12$ ; see Supplementary Table 1) for 1 h at room temperature. Wells were washed five times with 1× TBST (10 mM Tris HCl, 150 mM NaCl, 0.05% Tween 20, pH 7.4) and blocked with 2% bovine serum albumin (BSA) in 1× TBST for 1 h at room temperature. Wells were probed with 1× $10^8$  phages in 100  $\mu\text{l}$  of 1× TBST and incubated for 1 h at room temperature. Unbound phages were removed and the wells were washed five times with 1× TBST. Bound phages were then eluted in 100  $\mu\text{l}$  of 5 M NaCl. Eluted phages were amplified and titred using BLT5403 bacteria according to the manufacturer's instructions. Amplified eluted phages were used as input phages for three additional rounds of positive selection.

After four rounds of positive selection, the titre of eluted phages was determined and adjusted to 1× $10^5$  per ml in 1× TBST buffer. For negative selection, we bound Immulon 4HB (Thermo Fisher Scientific) ELISA wells with 100  $\mu\text{l}$  of a 1:100 dilution of serum that was pooled from malaria-susceptible children ( $n=14$ ; see Supplementary Table 1) for 1 h at room temperature. An aliquot of 100  $\mu\text{l}$  of the diluted phages (10,000 total phages) was added to the wells, wells were incubated for 1 h at room temperature and the unbound phages were collected. Unbound phages were transferred to an additional well coated with serum pooled from susceptible children, incubated for 1 h at room temperature and collected. This process was repeated a total of five times.

After negative selection, the titre of the eluted phages was determined and 100 individual plaques were isolated. Their cDNA inserts were amplified by PCR using the vectors T7SelectUP (5'-GGAGCTGCGTATTCCAGTC-3') and T7Select Down (5'-AACCCCTCAAGACCCGTTA-3') and the PCR products were sequenced.

#### Expression and purification of PfGARP

We subcloned the open reading frame encoding amino acids 410–673 of PfGARP into pET30 (Novagen) and transformed the resulting plasmid into the expression host *E.coli* BL21(DE3) (Novagen). The pET30 vector encodes a His tag at both the amino and carboxy ends of the recombinant protein, thus facilitating purification by metal chelate chromatography. Transformants were grown in Terrific broth supplemented with 100  $\mu\text{g ml}^{-1}$  kanamycin, at 37 °C in a 10-l fermenter with oxygen sparging (10 l min<sup>-1</sup>) until an optical density at 600 nm (OD<sub>600</sub>) of 8.0 was reached. Isopropyl-β-D-thiogalactopyranoside was added to a final concentration of 1 mmol l<sup>-1</sup>, and the culture was fed continuously with 0.3 g ml<sup>-1</sup> glucose, 0.09 g ml<sup>-1</sup> yeast extract at 50 ml h<sup>-1</sup> for 12 h. Cultures were collected by centrifugation and 750 g of wet cell paste was resuspended in 10 l of 10 mmol l<sup>-1</sup> potassium phosphate, 150 mmol l<sup>-1</sup> NaCl, 10 mmol l<sup>-1</sup> imidazole, 0.5% Tween 20 and 0.5% Triton X-100, pH 8.0 and lysed by high-pressure disruption at 20,000 psi (Microfluidics, Model 110-T). The lysate was clarified by tangential flow microfiltration (filter area 1 m<sup>2</sup>, pore size 1  $\mu\text{m}$ ; Millipore) and 8 l of clarified lysate was recovered. Protein purification was achieved by a 4-step process on BioPilot chromatography equipment (Pharmacia). In brief, clarified lysate was applied to a FineLine Pilot 35 (GE Healthcare) column containing 90 ml of Ni-NTA Superflow Resin (Novagen). The protein of interest was eluted with a stepped gradient containing increasing concentrations of imidazole. Fractions containing the protein of interest were pooled, adjusted to 400 mmol l<sup>-1</sup> ammonium sulfate, 10 mmol l<sup>-1</sup> DTT and further purified, by hydrophobic-interaction chromatography on a Fine Line Pilot 35 (GE Healthcare) column containing 150 ml of Source 15PHE (GE Healthcare). Recombinant proteins were eluted with a linear gradient of elution buffer (10 mmol l<sup>-1</sup> Tris, 1 mmol l<sup>-1</sup> DTT, 1 mmol l<sup>-1</sup> EDTA, pH 8.0). Fractions containing the protein of interest were pooled and further purified by anion-exchange chromatography on a Fine Line Pilot 35 (GE Healthcare) column containing 130 ml of MacroPrep High Q (BioRad). Recombinant proteins were eluted with a linear gradient of elution buffer (10 mmol l<sup>-1</sup> Tris, 1 mol l<sup>-1</sup> NaCl, 1 mmol l<sup>-1</sup> DTT, 1 mmol l<sup>-1</sup> EDTA, pH 8.0). Final purification was achieved by ceramic hydroxyapatite (CHT) chromatography on a FineLine Pilot 35 (GE Healthcare) column containing 70 ml of CHT type 1 (BioRad). Recombinant proteins were eluted with a linear gradient of elution buffer (500 mmol l<sup>-1</sup> potassium phosphate and 1 mmol l<sup>-1</sup> DTT, pH 7.4).

Purified recombinant protein, designated rPfGARP-A, was buffer-exchanged into 10 mmol l<sup>-1</sup> sodium phosphate, 0.05% Tween 20 and 3% sucrose, and concentrated to 500  $\mu\text{g ml}^{-1}$  by tangential flow ultrafiltration (filter area 50 cm<sup>2</sup>, pore size 5 kDa; Pall). rPfGARP was lyophilized at 500  $\mu\text{g}$  per vial and stoppered under nitrogen. Endotoxin levels were less than 2 EU per mg protein as determined by an FDA cleared assay (Lonza). Typical yields were over 50 mg rPfGARP per 750 g wet cell paste.

Notably, rPfGARP-A expresses immunorelevant epitopes, which generate functional polyclonal antibodies that block parasite growth and kill trophozoites (Fig. 1) and recognize native PfGARP by western blot (Extended Data Fig. 1).

#### Parasite strains and culture

*P.falciparum* strains (3D7, Dd2 and D10) were obtained from MR4. Two parasite isolates from adults and two parasite isolates from children were collected from our Tanzanian field site and culture-adapted. The parasites were cultured *in vitro* according to the previously published methods, with minor modifications<sup>34</sup>. In brief, parasites were maintained in RPMI 1640 medium containing 25 mM HEPES, 5% human O+ erythrocytes, 5% Albumax II (Invitrogen), 24 mM sodium bicarbonate and 10  $\mu\text{g ml}^{-1}$  gentamycin at 37 °C with 5% CO<sub>2</sub>, 1% O<sub>2</sub> and 94% N<sub>2</sub>.

# Article

## Anti-PfGARP antisera

Mouse anti-PfGARP antisera were produced by immunization based on DNA, recombinant protein or mRNA. For DNA immunization, we subcloned the open reading frame encoding amino acids 410–673 of PfGARP into VR2001, transformed this into the host *E. coli* NovaBlue (Novagen), and purified the endotoxin-free plasmid (Endofree Giga, Qiagen). BALB/cJ mice were immunized with 100 µg of plasmid (25 µg intramuscular injection in each hind leg and 50 µg intradermal injection at the base of the tail) followed by 50 µg intradermal injections at the base of the tail every two weeks for a total of four doses.

For protein immunization, we emulsified rPfGARP-A in an equal volume of TiterMax adjuvant (CytRx Corporation) and injected 50 µg of rPfGARP-A intraperitoneally at two-week intervals for a total of four doses.

For mRNA-based immunization, BALB/cJ mice were immunized intradermally with 10 µg lipid-encapsulated mRNA (see below) encoding amino acids 410–673 of PfGARP every three weeks for a total of three doses.

## Affinity purification of anti-PfGARP antisera

To purify polyclonal mouse and human anti-PfGARP IgG, we coupled 6 mg of rPfGARP-A to one ml of NHS-activated Sepharose 4 Fast Flow (GE Health Sciences) according to the manufacturer's instructions. For mouse anti-PfGARP IgG, we used plasma pooled from rPfGARP-A immunized mice. For human anti-PfGARP IgG, we used plasma pooled from placental blood collected from women delivering in our Tanzanian birth cohort.

In brief, rPfGARP-A coupled resin was incubated with 600 µl of pooled mouse or human plasma diluted in 6 ml of phosphate-buffered saline (PBS). After extensive washing in PBS with 0.05% Tween 20, bound antibody was eluted in 0.1 M glycine, pH 2.5 and immediately neutralized with 1 M Tris HCl, pH 8. Eluted antibodies were buffer exchanged into PBS by diafiltration in spin columns (Centricon) and sterilized before use in immunoblot and in vitro growth assays.

## Western blot

Parasite pellets were prepared by treatment of parasitized RBCs with 0.15% saponin in PBS, pH 7.4 on ice for 10 min, followed by centrifugation (3,000g, 5 min), resuspension in cold PBS and centrifugation (3,000g, 5 min). Alternatively, RBCs infected with mature-stage parasites were purified on a magnetic column and resuspended in distilled water (1:25 v/v) to lyse the RBCs, followed by centrifugation (3,000g, 5 min) to collect parasite pellets.

Parasite pellets or rPfGARP-A were dissolved in SDS sample-loading buffer with reducing agent (Bio-Rad) and heated to 95 °C for 10 min, and proteins were separated in 4–11% gradient SDS-PAGE gels. Separated proteins were transferred to nitrocellulose membranes, which were blocked in 5% milk PBS (pH 7.4) and 0.05% Tween 20 for 1 h. Membranes were probed with polyclonal anti-PfGARP or pre-immune mouse serum, detected by an anti-mouse IgG antibody conjugated to IRDyes and imaged on an LI-COR (Odyssey Imaging Systems).

## Growth inhibition assays

GIAs were carried out with polyclonal anti-PfGARP or IgG monoclonal and polyclonal antibodies, control anti-fluorescein monoclonal antibodies, control mouse serum or IgG as described<sup>35,36</sup> with minor modifications. In brief, anti-PfGARP antibodies or controls were dialysed overnight in PBS, pH 7.4, heat-inactivated at 56 °C for 30 min and pre-incubated with human RBCs for 1 h before use in GIA assays. GIA assays were carried out using Dd2, 3D7, D10 or four newly adapted isolates of *P. falciparum* collected at our Tanzanian field site. Parasites were synchronized to the ring stage by treatment with 5% sorbitol<sup>37</sup> for three successive replication cycles and cultured to the ring stage. Parasites at 0.3–0.4% parasitaemia and 2% haematocrit were incubated

with antiserum or IgG, in a final volume of 100 µl in microtitre wells. Cultures were performed in triplicate with five replicates (comprising a total of 15 individual wells) prepared for each treatment condition. After 48 h, blood films were prepared from each replicate and stained with Giemsa. A microscopist blinded to the treatment conditions counted the RBCs that were infected with ring-stage parasites, and the results from the three wells were averaged. The relationship between the treatment group and the parasitaemia outcome of the five replicates was analysed by Mann–Whitney *U*-test.

In some GIA assays (Fig. 1d), parasites were plated at 0.08% parasitaemia and were incubated with antiserum, in a final volume of 100 µl in microtitre wells. Cultures were examined by microscopy daily for seven days. The medium (with appropriate antiserum) was changed daily. Cultures were performed in triplicate with five replicates (comprising a total of 15 individual wells) prepared for each treatment condition.

## Production of monoclonal antibodies and epitope mapping

We immunized BALB/cJ mice intraperitoneally three times at two-week intervals with 25 µg of rPfGARP-A emulsified in TiterMax, and boosted with 25 µg of rPfGARP-A intravenously three days before fusion of spleenocytes with P3X63Ag8.653 myeloma cells (ATCC) according to our published method<sup>33</sup>.

Hybridomas were cultured and cloned by limited dilutions. Hybridomas producing anti-PfGARP antibodies were screened and used for large-scale production of monoclonal antibodies.

The heavy- and light-chain variable regions of hybridoma clone 7899 (IgG1k) were sequenced, and recombinant plasmids were constructed and used for the production of recombinant monoclonal antibodies in HEK293 cells according to the manufacturer's protocols (Absolute Antibody). Recombinant monoclonal antibodies expressing the heavy- and light-chain variable regions from clone 7899 on a mouse IgG1 framework were purified by Protein A chromatography. Recombinant monoclonal antibodies expressing the heavy- and light-chain variable regions from clone 7899 on a mouse IgG1Fab framework (monovalent) were purified by Ni-NTA chromatography.

For epitope mapping, a custom 15-mer peptide microarray was designed and printed. The array contained 264 different peptides that spanned the PfGARP-A sequence (amino acids 410–673). The peptides overlapped by a single amino acid and were printed in duplicate, framed by haemagglutinin control peptides. The array was probed with recombinant mAb7899 (red) and anti-HA (green) and imaged on a LI-COR Odyssey according to the manufacturer's protocol (PepperPrint).

## Mitochondrial membrane potential

Mitochondrial membrane potential was assessed with the dye JC-1, a cationic dye that exhibits membrane-potential-dependent accumulation in mitochondria. RBCs infected with ring-stage parasites from *P. falciparum* in vitro cultures were collected at 5–10% parasitaemia and incubated with anti-glycophorin A antibodies (Invitrogen) for 30 min at room temperature followed by Hoechst 33342 dye at 1 µg ml<sup>-1</sup>. JC-1 staining was performed at a final concentration of 2 µM for 30 min at 37 °C in the dark with constant agitation and analysed by flow cytometry.

## Caspase-activation assay

The activation of caspase-like proteases was quantified using a cell-permeable, fluorescein isothiocyanate (FITC)-conjugated pan-caspase inhibitor (FITC-VAD-FMK, Apostat, R&D Systems) which irreversibly binds and labels activated cysteine proteases. Antibody-treated *P. falciparum*-infected RBCs (treated with anti-PfGARP or control) were incubated with FITC-VAD-FMK, at a final concentration of 1%, 1 h before collection. Samples were washed with PBS to remove unbound reagent, and this was followed by staining with anti-glycophorin A antibodies (Invitrogen) and Hoechst 33342 dye (Thermo Fisher Scientific) as above. Samples were analysed by flow cytometry.

### TUNEL assay

Fragmentation of intracellular DNA in the cultured *P.falciparum* parasites was evaluated by TUNEL assay, using the APO-DIRECT Kit (BD Biosciences). Antibody-treated *P.falciparum*-infected RBCs (treated with anti-PfGARP or control) were collected, washed and stored in ice-cold 70% ethanol for at least 18 h before staining with 50 µl of DNA labelling solution, prepared as per the manufacturer's instructions. After incubation for about 1 h with the DNA labelling solution, samples were stained with anti-glycophorin A antibodies (Invitrogen) and Hoechst 33342 dye (Thermo Fisher Scientific) as above and analysed by flow cytometry.

### Lactate assay

The spent medium from 3D7 parasites incubated with recombinant anti-PfGARP monoclonal antibody or control recombinant anti-fluorescein in GIA assays was collected and assayed for lactate levels (lactate assay kit II, Sigma-Aldrich, MAK065) according to the manufacturer's protocol.

### Generation of 3D7-PfGARP KD and 3D7-PfGARP KO parasite lines

For 3D7-PfGARP KD parasites, approximately 100 µg of pRR203 was linearized with EcoRV, purified and co-transfected with 100 µg pRR183 into 3D7 parasites. Parasites were maintained on 500 nM anhydrotetracycline. One day after transfection, drug pressure was applied with 2.5 nM WR99210 and the PfDHODH inhibitor *N*-(3-chloro-4-methylphenyl)-5-methyl-2-(trifluoromethyl)[1,2,4]triazolo[1,5-a]pyrimidin-7-amine (MMV665874 or AD1) at 150 nM. Five days after transfection, AD1 selection was removed. Parasites were cloned by limiting dilution.

3D7-PfGARP KO parasites were made as above, with the exception that pRR248 was used as the homology-directed repair plasmid. pRR248 was linearized with AvrII, NotI and Sapi and transfected with pRR183 into a strain of 3D7 parasites with episomal expression of spCas9. Only WR99210 drug pressure was applied after transfection.

For 3D7-PfGARP KD parasites, integration of the targeting construct was confirmed by PCR with oJDD1027/oJDD4507 (control), oJDD1027/oJDD2933 (integration) and oJDD4279/4280 (locus size). Tet aptamer size was confirmed by amplifying the aptamer with oJDD3560/oJDD44 and digesting the PCR fragment with PspOMI and KpnI. For 3D7-PfGARP KO parasites, integration of the targeting construct was similarly confirmed by PCR.

### Construction of PfGARP homology-directed repair plasmid.

To prevent homology-directed repair occurring between the spCas9-directed cut site and the 3' end of PfGARP, we synthesized caPfGARP, in which the last 438 bp of PfGARP were codon-altered. We amplified the synthesized gene with oJDD4506/oJDD3566, and the 500 bp preceding the codon-altered region from 3D7 genomic DNA with oJDD3563/oJDD4507. We amplified 500 bp from the 3' end of PfGARP for the 3' homology region with oJDD3561/oJDD3562 from 3D7 genomic DNA. We used PCR with splicing by overhang extension (SOE) to create a PfGARP 3'HR-EcoRV-PfGARP 5'HR-caPfGARP fusion. We cloned this fusion into pRR69<sup>38</sup> with NotI and NcoI to generate pRR203.

**Construction of PfGARP-targeting spCas9 vector and generation of knock-down parasite line.** A PfGARP-targeting guide was cloned into the U6 cassette of pBAM203<sup>38</sup> by PCR SOE. The U6 promoter with guide was amplified by oJDD3058/oJDD4088 and gRNA and U6 terminator with guide amplified with oJDD3059/oJDD4089. The cassette was cloned into pBAM203 with EcoRI and AvrII to generate pRR183. The 3D7 parasites were transfected with pRR183 and genome-integrated parasites were selected by serial dilution.

**Construction of PfGARP KO homology-directed repair plasmid.** To knock out PfGARP, we generated a cassette for double crossover that integrated into the 5' UTR of PfGARP and at the 3' end of the sequence coding for PfGARP. We amplified the 5' UTR of PfGARP from genomic DNA with oJDD4077/4979 and cloned it into a plasmid containing a cassette expressing the KAHRP signal sequence fused to nanoluciferase followed by a 2A skip peptide and the hDHFR selection marker. The 5' UTR of PfGARP was cloned upstream of the cassette with NotI/AflIII. The 3' homology region was amplified with oJDD5029/4083 and cloned downstream of the selection cassette with EcoRI/AvrII to generate pRR248.

### Immunofluorescence assays

Blood smears of asynchronous 3D7 strain parasite cultures were prepared, fixed in cold methanol for 15 min and probed with anti-PfGARP prepared by DNA vaccination, rabbit anti-PfMSP4 (obtained from MR4) or anti-rabbit glycophorin A diluted 1:200 in PBS, 5% BSA, pH 7.4. Blood smears were incubated with primary antibodies for 1 h at 25 °C, washed three times in PBS, 0.05% Tween 20 and incubated with goat anti-mouse IgG conjugated with Alexa Fluor 488 (Molecular Probes) and goat anti-rabbit IgG conjugated with Alexa Fluor 594 (Molecular Probes). Blood smears were incubated for 10 min in 1 µg ml<sup>-1</sup> DAPI (Sigma-Aldrich) to label nuclei and coverslipped with ProLong Gold anti-fade reagent (Invitrogen). Blood smears were imaged using a confocal microscope (Leica SP2, Leica Microsystems) equipped with a 100× oil-immersion objective, and sequential Z-sections of the infected RBC were collected. Immunofluorescence assays were also performed using anti-V5 monoclonal antibodies (Thermo-Fisher Scientific) and control IgG as described above.

To assess the colocalization of PfGARP with the membrane of the food vacuole, we probed parasites with both anti-PfGARP and anti-PfCRT, a protein which localizes to the food vacuole membrane<sup>39</sup>.

In some experiments, parasite-infected RBCs were incubated with primary antibodies before fixation and slide preparation to ensure that the methanol used for fixation did not permeabilize the RBC membrane.

### Transmission electron microscopy

To assess the effect of anti-PfGARP on parasite ultrastructure, we performed transmission electron microscopy on parasites after a 24-h incubation with anti-PfGARP or control antibodies. 3D7 strain parasites were grown to high parasitaemia (10%) consisting of predominantly trophozoites. Parasites were incubated with anti-PfGARP prepared by DNA or rPfGARP-A vaccination at 10% serum concentration for 24 h at 37 °C. Pre-immune mouse serum at 10% serum concentration was used as a negative control.

Parasites were washed three times in 1× PBS, and were fixed for 30 min at 4 °C with 2% glutaraldehyde and 1% paraformaldehyde in 0.1 M sodium cacodylate buffer. Samples were dehydrated, embedded in Epon (EMS), sectioned on an ultra-microtome, counterstained for 10 min in 5% aqueous uranyl acetate and examined on a Philips CM10 electron microscope.

For immunoelectron microscopy, we performed live-cell staining followed by fixation. Parasitized RBCs were blocked for 1 h at 25 °C in 1× PBS containing 2% BSA. Samples were incubated with anti-PfGARP prepared by DNA or rPfGARP-A vaccination (diluted 1:200 in PBS) for 3 h at 25 °C. Pre-immune mouse serum was used as a negative control. The samples were washed three times in 1× PBS and probed with gold-conjugated anti-mouse antibodies for 1 h at 25 °C at 1:2 dilution (Invitrogen). Samples were washed three times in 1× PBS, fixed and processed as described above.

### Assessment of food vacuole integrity

We further evaluated the effect of anti-PfGARP antibodies on food vacuole integrity by confocal microscopy using the calcium-binding

## Article

dye Fluo-4 AM (a cell-permeable calcium-sensing dye that fluoresces green when bound with  $\text{Ca}^{2+}$ ), which specifically labels the food vacuole. Ring-stage 3D7 parasites were incubated with culture medium (negative control), 1  $\mu\text{M}$  chloroquine (positive control) or anti-PfGARP prepared by rPfGARP-A immunization for 24 h. Parasites were washed with 1 $\times$  PBS three times and incubated with Fluo-4 AM at a final concentration of 2  $\mu\text{M}$  and DAPI for 30 min at 37 °C. The parasites were washed with 1 $\times$  PBS and observed under a fluorescent microscope. Under normal conditions,  $\text{Ca}^{2+}$  ions sequester inside the food vacuole of the parasite, which results in a punctate staining pattern of Fluo-4 AM. Treatment with chloroquine or anti-PfGARP disrupted the food vacuole leading to dispersion of  $\text{Ca}^{2+}$  ions into the cytosol of the parasite and resulted in a diffuse staining pattern of Fluo-4 AM.

### Serial block-face scanning electron microscopy

To further assess the effect of anti-PfGARP on parasite ultrastructure, we performed SBF-SEM on parasite-infected RBCs that had been treated with anti-PfGARP or pre-immune antiserum. Ring-stage 3D7 strain parasites were grown to 5% parasitaemia and incubated with anti-PfGARP (prepared by DNA vaccination) or pre-immune serum (10% final serum concentration) for 30 h at 37 °C. Parasites were washed three times in 1 $\times$  PBS and fixed for 1 h at 4 °C with 2% glutaraldehyde in 0.1 M sodium cacodylate buffer containing 1 mM  $\text{CaCl}_2$  and 1 mM  $\text{MgCl}_2$ . Samples were washed five times with sodium cacodylate buffer. Samples were incubated at 4 °C for 1 h with 1% osmium tetroxide and 1.5% potassium ferrocyanide in sodium cacodylate buffer followed by five washes with deionized water. Samples were incubated in 1% aqueous thiocarbohydrazide for 20 min followed by five washes with deionized water. Samples were incubated in 1% aqueous osmium tetroxide for 30 min, followed by overnight incubation at 4 °C in 1% uranyl acetate in 70% ethanol followed by five washes with deionized water. Samples were finally incubated in 0.2% aqueous lead citrate for 30 min before being dehydrated, treated with propylene oxide and embedded in Epon resin (EMS). Once cured, samples were cut and mounted on a 6.6-mm specimen mount with conductive silver epoxy (MG Chemicals). Specimens were coated with gold palladium using an Emitech K550 sputter coater before being sectioned and imaged with a Thermo Apreo Volume Scope scanning electron microscope in high-vacuum conditions. Acquired images were processed and the final three-dimensional reconstruction was performed with Amira 2019.2 software.

### Processing of PfGARP

PfGARP encodes a predicted N-terminal signal sequence or transmembrane region (amino acids 1–22) and an appropriately located PEXEL motif (amino acids 48–52). To determine whether parasites process and cleave the PEXEL motif, we probed western blots of parasite extracts and rPfGARP-A using peptide-specific antisera generated against peptides that flank the PEXEL motif (amino acids 31–48 and 504–522).

### Anti-PfGARP antibody assays

Initial, confirmatory antibody assays were performed with rPfGARP-coated ELISA plates according to our published methods. To measure IgG anti-rPfGARP antibody levels in the Kenyan cohort; we developed a bead-based assay according to our published methods<sup>40</sup>. In brief, 100  $\mu\text{g}$  of rPfGARP-A or 100  $\mu\text{g}$  of BSA was conjugated to  $1.25 \times 10^7$  microspheres (Luminex), and conjugated rPfGARP and BSA beads were pooled and lyophilized in single-use aliquots. Reconstituted beads were incubated for 30 min at 37 °C with human plasma samples at 1:80 dilution in Assay Buffer E (ABE; PBS pH 7.4 containing 0.1% BSA, 0.05% Tween 20 and 0.05% sodium azide) in microtitre filter-bottom plates (Millipore). Beads were washed three times in ABE by vacuum filtration and incubated for 30 min at 37 °C with biotinylated anti-human IgG (Pharmingen) diluted 1:1,000 in ABE. Beads were washed three times in ABE by vacuum filtration and incubated for 10 min at 37 °C with phycoerythrin-conjugated streptavidin (Pharmingen) diluted 1:500 in ABE. Beads were washed three times in ABE by vacuum filtration, resuspended in ABE and analysed on a BioPlex 200 multi-analyte analyser. Fluorescence values for AdEx beads were subtracted from rPfGARP beads. The cut-off for detectable anti-PfGARP antibody levels was defined as fluorescence values for rPfGARP beads exceeding the value for AdEx beads.

1:500 in ABE. Beads were washed three times in ABE by vacuum filtration, resuspended in ABE and analysed on a BioPlex 200 multi-analyte analyser. Fluorescence values for BSA beads were subtracted from rPfGARP beads. The cut-off for detectable anti-PfGARP antibody levels was defined as fluorescence values greater than the mean + 2 s.d. fluorescence level of 22 healthy North American adults.

To measure IgG anti-rPfGARP antibody levels in the Tanzanian cohort, we performed bead-based assays using amino acids 23–673 of PfGARP expressed and purified from COS-7 cells as the target antigen. We expressed and purified amino acids 23–673 of PfGARP (excluding initial signal sequence or transmembrane domain) in a eukaryotic expression system (COS-7 cells) according to our published methods<sup>41–43</sup>. In brief, the coding sequence for amino acids 23–673 of PfGARP was PCR-amplified with primers encoding BamH1 and EcoR1 restriction sites, gel-purified, digested with BamH1 and EcoR1, and ligated into BamH1- and EcoR1-digested vector pAdEx<sup>41–43</sup>. The integrity of the PfGARP construct was verified by sequencing on both strands. Transfection of COS-7 cells, expression, extraction and immobilization of recombinant PfGARP protein and control AdEx protein on the surface of BioPlex beads have been described in detail in our previous publications<sup>41–43</sup>. Immobilization of recombinant PfGARP on beads was verified by reactivity with mouse anti-PfGARP antibodies. PfGARP and AdEx-only beads were incubated for 30 min at 37 °C with human plasma samples at 1:80 dilution in ABE (PBS pH 7.4 containing 0.1% BSA, 0.05% Tween 20 and 0.05% sodium azide) in microtitre filter-bottom plates (Millipore). Beads were washed three times in ABE by vacuum filtration and incubated for 30 min at 37 °C with biotinylated anti-human IgG (Pharmingen) diluted 1:1,000 in ABE. Beads were washed three times in ABE by vacuum filtration and incubated for 10 min at 37 °C with phycoerythrin-conjugated streptavidin (Pharmingen) diluted 1:500 in ABE. Beads were washed three times in ABE by vacuum filtration, resuspended in ABE and analysed on a BioPlex 200 multi-analyte analyser. Fluorescence values for AdEx beads were subtracted from rPfGARP beads. The cut-off for detectable anti-PfGARP antibody levels was defined as fluorescence values for rPfGARP beads exceeding the value for AdEx beads.

### Statistical analyses of cohort studies

**Tanzanian birth cohort.** To assess the relationship between anti-PfGARP antibody responses and resistance to clinical malaria outcomes, we developed GEE-based repeated measures models with antibody levels as a continuous variable (SAS v.9.3). These models were used to evaluate the relationship between anti-PfGARP antibody levels (log-transformed) and the risk of malaria outcomes.

Potential confounders and effect modifiers—including haemoglobin phenotype, birthweight and transmission season at birth—were evaluated with the two-sided fixed-effect test and retained in the model if their *P* value was less than 0.1 or their inclusion altered the  $\beta$  coefficient for the anti-PfGARP term by more than 10%. Only haemoglobin phenotype met the pre-specified criteria for inclusion (*P*<0.1). We report both the unadjusted results and the model adjusted for haemoglobin phenotype.

**Kenyan cohort.** To assess the relationship between anti-PfGARP antibody responses and resistance to *P. falciparum* parasitaemia, we developed GEE-based repeated measures models using JMP v.10. We evaluated the relationship between detectable anti-PfGARP IgG antibodies and parasite density measured on 18 post-treatment blood films. We assessed several potential confounders and effect modifiers including age, week of follow-up, exposure to *Anopheles* mosquitoes and haemoglobin phenotype using the two-sided fixed-effect test. Variables, including the week of monitoring, were retained in the model if their *P* value was less than 0.1 or they changed the parameter estimate for the antibody of interest by >10%. Age, exposure and haemoglobin phenotype were retained for face validity, although they did not meet the

pre-specified criteria for inclusion, and PfGARP antibodies remained a significant predictor of parasite density with their exclusion.

### Production of PfGARP-A mRNA

mRNAs were produced as previously described<sup>44</sup> using T7 RNA polymerase (Megascript, Ambion) on a linearized plasmid encoding codon-optimized<sup>45</sup> PfGARP-A. mRNAs were transcribed to contain 101 nucleotide-long poly(A) tails. One-methylpseudouridine (m1Ψ)-5'-triphosphate (TriLink) instead of UTP was used to generate modified nucleoside-containing mRNA. RNAs were capped using the m7G capping kit with 2'-O-methyltransferase (ScriptCap, CellScript) to obtain cap1. mRNA was purified by FPLC (Akta Purifier, GE Healthcare) as described<sup>46</sup>. All mRNAs were analysed by denaturing or native agarose gel electrophoresis and were stored frozen at -20 °C.

### Encapsulation of mRNA in LNPs

Poly(C) RNA (Sigma-Aldrich) and FPLC-purified m1Ψ-containing mRNAs were encapsulated in LNPs using a self-assembly process in which an aqueous solution of mRNA at pH 4.0 is rapidly mixed with a solution of lipids dissolved in ethanol<sup>47</sup>. LNPs used in this study were similar in composition to those described previously<sup>48</sup>, containing an ionizable cationic lipid (proprietary to Acuitas)/phosphatidylcholine/cholesterol/PEG lipid (50:10:38.5:1.5 mol/mol) and were encapsulated at an RNA to total lipid ratio of around 0.05 (w/w). They had a diameter of around 80 nm as measured by dynamic light scattering using a Zetasizer Nano ZS (Malvern Instruments). mRNA LNPs were stored at -80 °C at a concentration of mRNA of around 1 µg µl<sup>-1</sup>.

### Studies of *Aotus* monkey vaccination with mRNA LNPs

We immunized  $n = 5$  monkeys with 50 µg of PfGARP-A-mRNA LNPs, and  $n = 4$  monkeys with 50 µg of poly(C)-RNA LNPs, intradermally on four sites of the shaved back. Monkeys received 3 doses at weeks 0, 3 and 6. Before each immunization, serum samples were obtained for antibody assays. At week 9, monkeys were challenged by intravenous injection with  $1 \times 10^4$  RBCs infected with blood-stage *P.falciparum* FVO strain parasites, followed by daily blood films. Antibody assays were performed with PfGARP-A-coated beads according to our published methods<sup>40</sup> using a biotin-conjugated anti-monkey IgG antibody (Invitrogen) for detection of bound anti-PfGARP. Monkeys were monitored daily by blood samples from day 4 after challenge to quantify parasitaemia and haemoglobin concentration. Monkeys with parasitaemia greater than 7.5%, haematocrit lower than 25% or that exhibited signs of illness (fever, immobility, decreased food intake and so on) were treated with oral mefloquine in accordance with our animal protocol.

### Vaccination of *Aotus* monkeys with rPfGARP-A

We immunized  $n = 4$  monkeys intradermally with 50 µg of recombinant, *E. coli*-produced PfGARP-A (amino acids 410–673) emulsified in Ribi adjuvant, and  $n = 5$  monkeys with Ribi adjuvant alone as control. Monkeys received 3 doses at weeks 0, 3 and 6. Before each immunization, samples of serum were obtained for antibody assays. At week 9, monkeys were challenged by intravenous injection with  $1 \times 10^4$  RBCs infected with blood-stage *P.falciparum* FVO strain parasites, followed by daily blood films. Antibody assays were performed with PfGARP-A coated beads according to our published methods<sup>40</sup> using biotin-conjugated anti-monkey IgG antibody (Invitrogen) for detection of bound anti-PfGARP. Monkeys were monitored daily by blood samples from day 4 post-challenge to quantify parasitaemia and haemoglobin concentration. Monkeys with parasitaemia greater than 7.5%, haematocrit lower than 25% or that exhibited signs of illness (fever, immobility, decreased food intake and so on) were treated with oral mefloquine in accordance with our animal protocol.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

### Data availability

Source Data for Extended Data Figs. 5 and 9 are provided with the paper. The DNA sequence for PfGARP is available in PlasmoDB ([www.plasmodb.org](http://www.plasmodb.org)) under the gene ID PF3D7\_0113000. All other relevant data are available within the manuscript and its Supplementary Information.

27. Ellman, R., Maxwell, C., Finch, R. & Shayo, D. Malaria and anaemia at different altitudes in the Muheza district of Tanzania: childhood morbidity in relation to level of exposure to infection. *Ann. Trop. Med. Parasitol.* **92**, 741–753 (1998).
28. Kabyemela, E. R., Fried, M., Kurtis, J. D., Mutabingwa, T. K. & Duffy, P. E. Decreased susceptibility to *Plasmodium falciparum* infection in pregnant women with iron deficiency. *J. Infect. Dis.* **198**, 163–166 (2008).
29. Kurtis, J. D., Lanar, D. E., Opollo, M. & Duffy, P. E. Interleukin-10 responses to liver-stage antigen 1 predict human resistance to *Plasmodium falciparum*. *Infect. Immun.* **67**, 3424–3429 (1999).
30. Beier, J. C. et al. *Plasmodium falciparum* incidence relative to entomologic inoculation rates at a site proposed for testing malaria vaccines in western Kenya. *Am. J. Trop. Med. Hyg.* **50**, 529–536 (1994).
31. Friedman, J. F. et al. Malaria is related to decreased nutritional status among male adolescents and adults in the setting of intense perennial transmission. *J. Infect. Dis.* **188**, 449–457 (2003).
32. Gourley, I. S., Kurtis, J. D., Kamoun, M., Amon, J. J. & Duffy, P. E. Profound bias in interferon-γ and interleukin-6 allele frequencies in western Kenya, where severe malarial anaemia is common in children. *J. Infect. Dis.* **186**, 1007–1012 (2002).
33. Gunasekaran, K., Jambulingam, P., Sadanandam, C., Sahu, S. S. & Das, P. K. Reliability of light trap sampling for *Anopheles fluviatilis*, a vector of malaria. *Acta Trop.* **58**, 1–11 (1994).
34. Trager, W. & Jensen, J. B. Human malaria parasites in continuous culture. *Science* **193**, 673–675 (1976).
35. Bejon, P. et al. Effect of the pre-erythrocytic candidate malaria vaccine RTS,S/AS01E on blood stage immunity in young children. *J. Infect. Dis.* **204**, 9–18 (2011).
36. Malkin, E. M. et al. Phase 1 clinical trial of apical membrane antigen 1: an asexual blood-stage vaccine for *Plasmodium falciparum* malaria. *Infect. Immun.* **73**, 3677–3685 (2005).
37. Lambros, C. & Vanderberg, J. P. Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. *J. Parasitol.* **65**, 418–420 (1979).
38. Rudlaff, R. M., Kraemer, S., Streva, V. A. & Dvorin, J. D. An essential contractile ring protein controls cell division in *Plasmodium falciparum*. *Nat. Commun.* **10**, 2181 (2019).
39. Pulcini, S. et al. Mutations in the *Plasmodium falciparum* chloroquine resistance transporter, PfCRT, enlarge the parasite's food vacuole and alter drug sensitivities. *Sci. Rep.* **5**, 14552 (2015).
40. Cham, G. K. et al. A semi-automated multiplex high-throughput assay for measuring IgG antibodies against *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1) domains in small volumes of plasma. *Malar. J.* **7**, 108 (2008).
41. Oleinikov, A. V. et al. High throughput functional assays of the variant antigen PfEMP1 reveal a single domain in the 3D7 *Plasmodium falciparum* genome that binds ICAM1 with high affinity and is targeted by naturally acquired neutralizing antibodies. *PLoS Pathog.* **5**, e1000386 (2009).
42. Oleinikov, A. V. et al. Effects of sex, parity, and sequence variation on seroreactivity to candidate pregnancy malaria vaccine antigens. *J. Infect. Dis.* **196**, 155–164 (2007).
43. Oleinikov, A. V. et al. A plasma survey using 38 PfEMP1 domains reveals frequent recognition of the *Plasmodium falciparum* antigen VAR2CSA among young Tanzanian children. *PLoS One* **7**, e31011 (2012).
44. Pardi, N., Muramatsu, H., Weissman, D. & Karikó, K. In vitro transcription of long RNA containing modified nucleosides. *Methods Mol. Biol.* **969**, 29–42 (2013).
45. Thess, A. et al. Sequence-engineered mRNA without chemical nucleoside modifications enables an effective protein therapy in large animals. *Mol. Ther.* **23**, 1456–1464 (2015).
46. Weissman, D., Pardi, N., Muramatsu, H. & Karikó, K. HPLC purification of in vitro transcribed long RNA. *Methods Mol. Biol.* **969**, 43–54 (2013).
47. Maier, M. A. et al. Biodegradable lipids enabling rapidly eliminated lipid nanoparticles for systemic delivery of RNAi therapeutics. *Mol. Ther.* **21**, 1570–1578 (2013).
48. Jayaraman, M. et al. Maximizing the potency of siRNA lipid nanoparticles for hepatic gene silencing in vivo. *Angew. Chem.* **51**, 8529–8533 (2012).

**Acknowledgements** We thank MOMS project staff for their efforts in collecting clinical data, processing samples and interpreting malarial blood smears, and we thank the study participants and their families. This work was supported by grants from the US NIH (R01-AI076353, R01-AI127699 and R01-AI110699) and an internal Rhode Island Hospital Research Pilot Award grant to J.D.K.; grants from the US NIH (R01-AI52059) and the Bill & Melinda Gates Foundation (grant no. 1364) to P.E.D.; the Intramural Research Program of the NIH National Institute of Allergy and Infectious Diseases (NIAID), NIH grant R01-AI092120 and Florida Atlantic University start-up funds to A.V.O.; NIH grant R37-AI50234 to D.A.F.; and NIH grants R01-AI145941 and R01-AI102907 to J.D.D. We also acknowledge research core services provided by the Rhode Island Hospital imaging core (G. Hovanessian), the Leduc Bioimaging Facility (G. Williams) and core services supported by the COBRE Center for Cancer Research Development (P20GM103421). The Thermo Apreo Volume Scope scanning electron microscope that was used for the serial block-face imaging was purchased with a high-end instrumentation grant from the Office of the Director at the NIH (S10 OD023461). L.L., S. O.-G., M.F. and P.E.D. are supported by the Intramural Research Program of the NIAID.

**Author contributions** J.D.K., D.K.R., M.F. and P.E.D. conceived and supervised the study. J.D.K., P.E.D., M.F., J.F.F., S.P., C.B.M., D.A.F., D.K.R. and A.D.M. analysed the data and/or drafted the

## Article

text. D.K.R., A.D.M., A. Jnawali, J.Z., A. Jha, G.C.-K., B.S., C.E.N., N.H., S.P.-T. and L.B. contributed to parasite killing assays. D.K.R., J.Z., A.D.M., G.J., D.A.F., A.M. and N.F.G. contributed to imaging studies. P.E.D., M.F., J.D.K. and E.K. contributed to field-based data collection. P.E.D., M.F., S.P., J.F.F., A.V.O., O.C., J.D.K. and J.M. contributed to epidemiological analyses. R.M.R. and J.D.D. contributed to genetic modification of parasites. N.P., D.W., B.L.M. and Y.K.T. contributed to mRNA-based vaccine design and production. D.K.R., L.L., S.O.-G., M.F., P.E.D. and J.D.K. contributed to non-human primate studies.

**Competing interests** The work presented in this manuscript has been submitted in partial support of patent no. US10,213,502 B2 (filed 26 May 2017) on the use of PfGARP as a vaccine,

on which the authors J.D.K., D.K.R., J.F.F., M.F. and P.E.D. are named inventors. The remaining authors declare no competing interests.

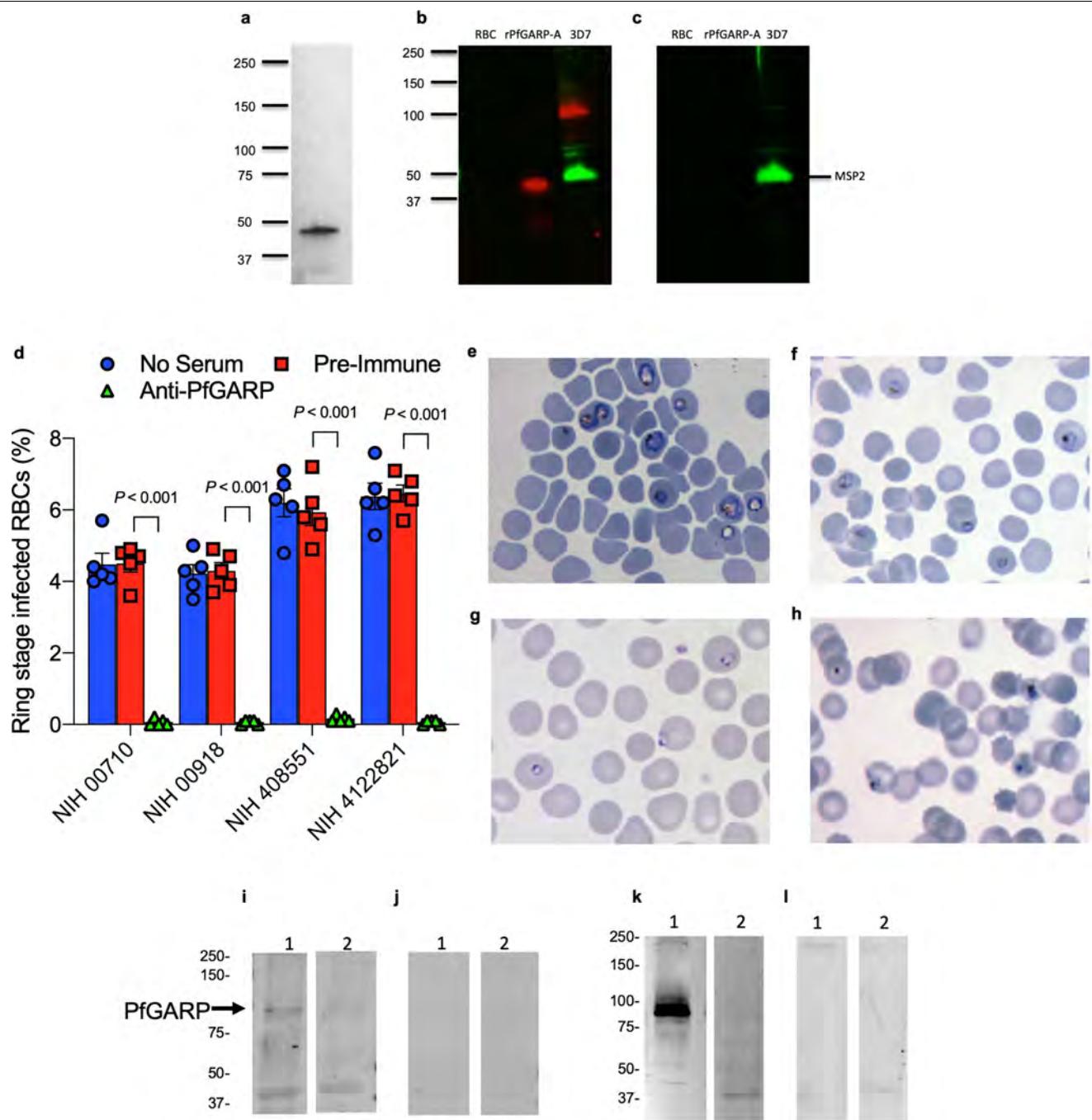
### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2220-1>.

**Correspondence and requests for materials** should be addressed to J.D.K.

**Peer review information** *Nature* thanks Peter Preiser and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

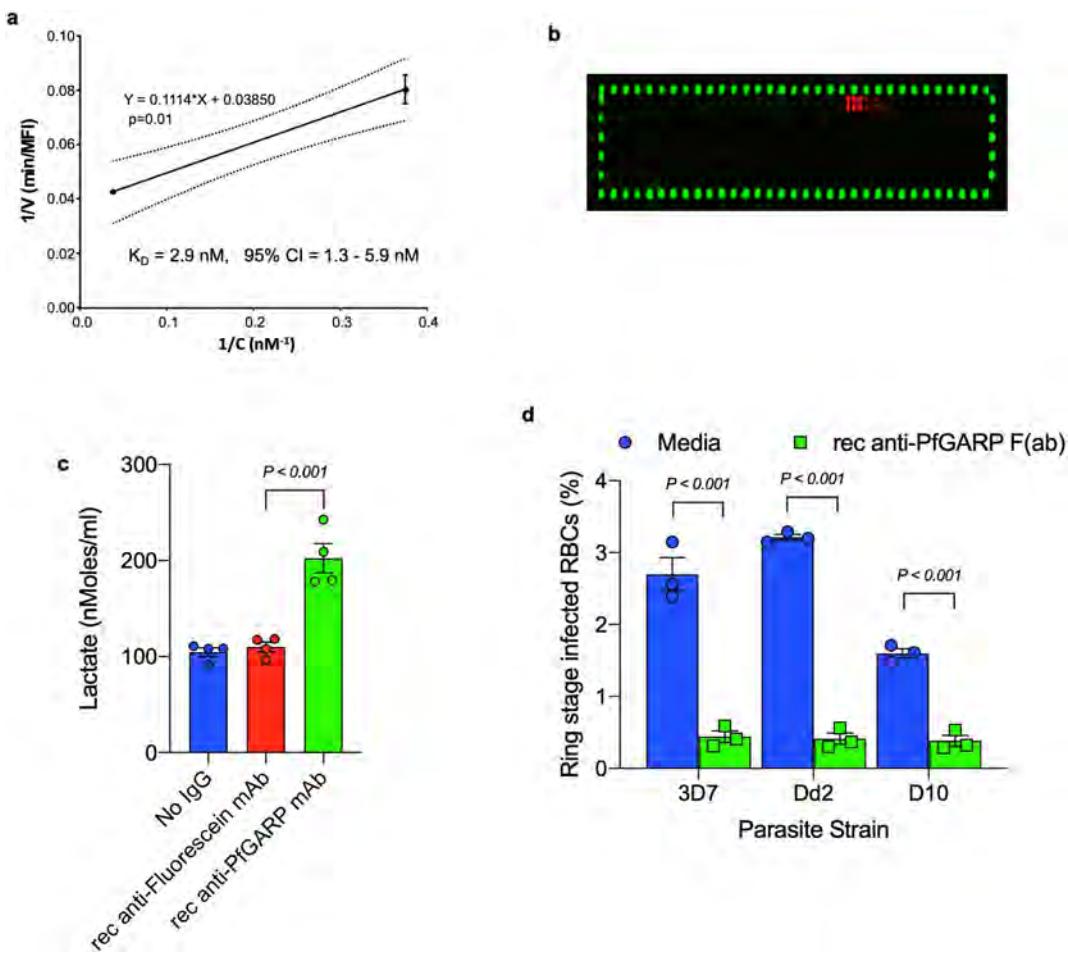
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**Extended Data Fig. 1** | See next page for caption.

# Article

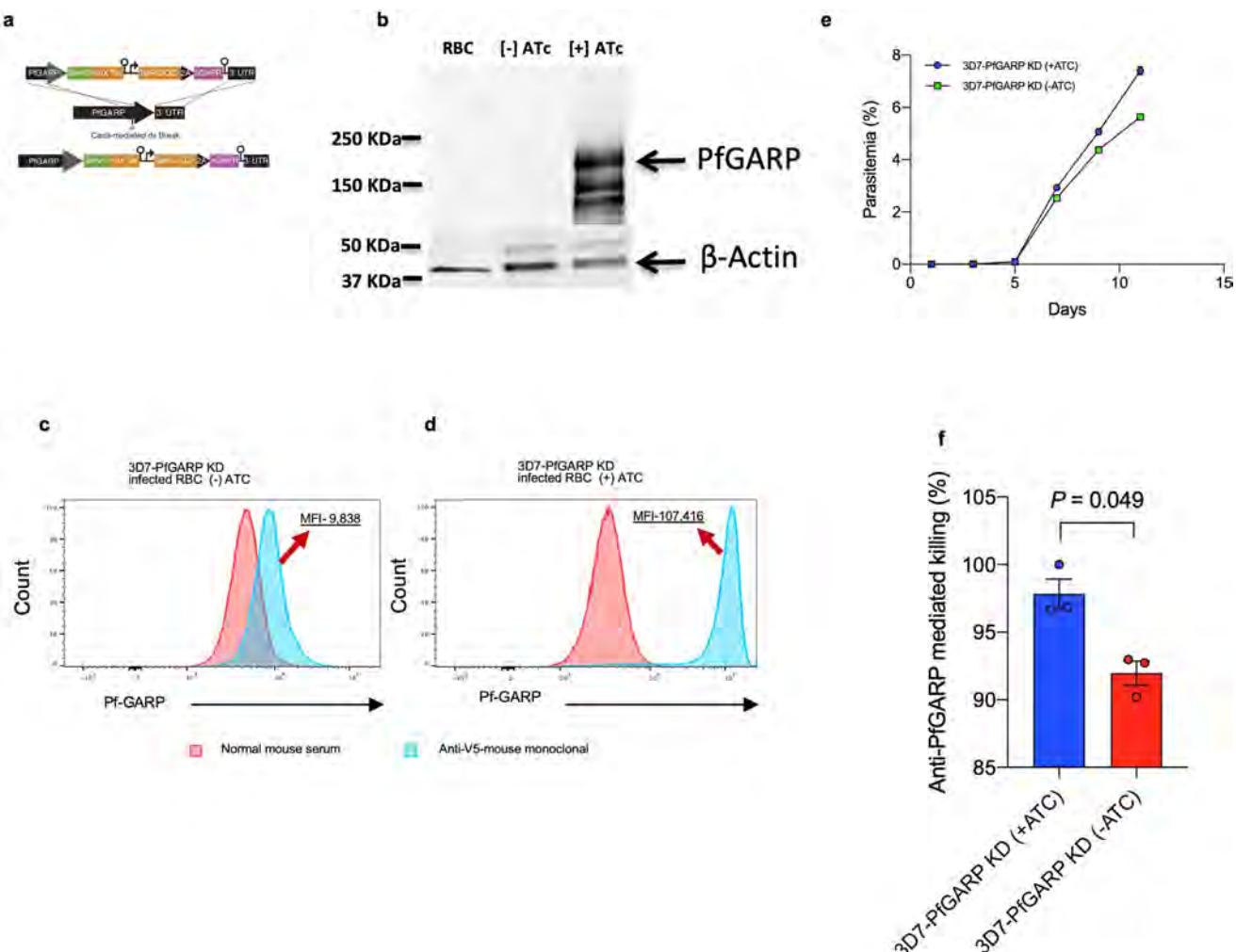
**Extended Data Fig. 1 | Expression of PfGARP, killing of parasites by anti-PfGARP and purification of anti-PfGARP antibodies.** **a**, SDS-PAGE gel of purified rPfGARP-A (250 ng). **b, c**, We analysed rPfGARP-A, as well as extracts prepared from uninfected RBCs and 3D7 trophozoite-infected RBCs by western blot. **b**, Blots were probed with mouse polyclonal anti-PfGARP-A serum (generated by plasmid immunization) and with anti-MSP2 (rabbit polyclonal serum), which were detected with anti-mouse IgG (red) and anti-rabbit IgG (green). **c**, Blots were probed with pre-immune mouse serum and anti-MSP2 (rabbit polyclonal serum), detected with anti-mouse IgG (red) and anti-rabbit IgG (green). Data in **a–c** are representative of five independent experiments. **d**, GIAs performed on parasites collected from our field site in Muheza, Tanzania, after short-term adaptation to the culture. Assays were performed using polyclonal anti-PfGARP-A antibodies generated by immunizing mice with recombinant protein. Ring-stage malaria parasites from adults (NIH 00710 and NIH 00918) or children (NIH 408551 and NIH 412282) were cultured in the presence of anti-PfGARP mouse serum at a 1:10 dilution. Negative controls were no antiserum (blue) and pre-immune mouse serum (red). Parasites were cultured for 48 h at 37 °C and ring-stage and early-trophozoite-stage parasites were counted by microscopy. Data are mean ± s.e.m. of five biologically independent replicates. *P* values were calculated by two-sided non-parametric Mann-Whitney *U*-test. Data are representative of five independent experiments. **e–h**, 3D7 *P. falciparum* parasites were synchronized to the ring stage and plated at 5% parasitaemia in the presence of pre-immune (**e, g**) and anti-rPfGARP-A (**f, h**) mouse serum at a 1:10 dilution. Parasites were cultured for 24 h (**e, f**) or 48 h (**g, h**), stained with Giemsa and photographed by light microscopy. Images in **e–h** are representative of three independent experiments. **i–l**, Purification of anti-PfGARP antibodies from human (**i, j**) and mouse (**k, l**) serum. Serum was affinity-purified using PfGARP-A coupled to sepharose beads. The specificity of the purified anti-PfGARP antibodies was determined by western blot on extracts prepared from unsynchronized 3D7 parasite-infected RBCs. In all panels: lane 1, infected RBCs extracted in RIPA buffer; lane 2, uninfected RBCs extracted in RIPA buffer. Blots in **i, j** were probed with anti-PfGARP purified from serum that was pooled from adults living in a holoendemic area of Tanzania (2 µg ml<sup>-1</sup>) (**i**), or with immunoglobulin from malaria-naïve adults (2 µg ml<sup>-1</sup>) (**j**). Blots in **k, l** were probed with anti-PfGARP purified from serum that was prepared from PfGARP-A-immunized mice (10 µg ml<sup>-1</sup>) (**k**), or immunoglobulin from malaria-naïve mice (10 µg ml<sup>-1</sup>). Blots in **i–l** are representative of two biologically independent experiments.



**Extended Data Fig. 2 | Characterization of recombinant monoclonal anti-PfGARP.** **a**, The kinetics of binding between the recombinant monoclonal antibody mAb7899 and PfGARP-A were measured at two antibody concentrations. For each concentration, two biologically independent replicates were performed. The experiment was performed twice and the data were pooled. The dashed lines represent the 95% CI. Error bars represent s.d. The formula is a linear regression.  $C$ , concentration of biotinylated recombinant mAb7899; MFI, median fluorescence intensity;  $V$ , initial velocity of binding. **b**, Epitope mapping of recombinant mAb7899. We printed a custom 15-mer-peptide microarray containing 264 different peptides that spanned the PfGARP-A sequence (amino acids 410–673). The peptides overlapped by a single amino acid and were printed in duplicate, framed by haemagglutinin (HA) control peptides. The array was probed with recombinant mAb7899 (red)

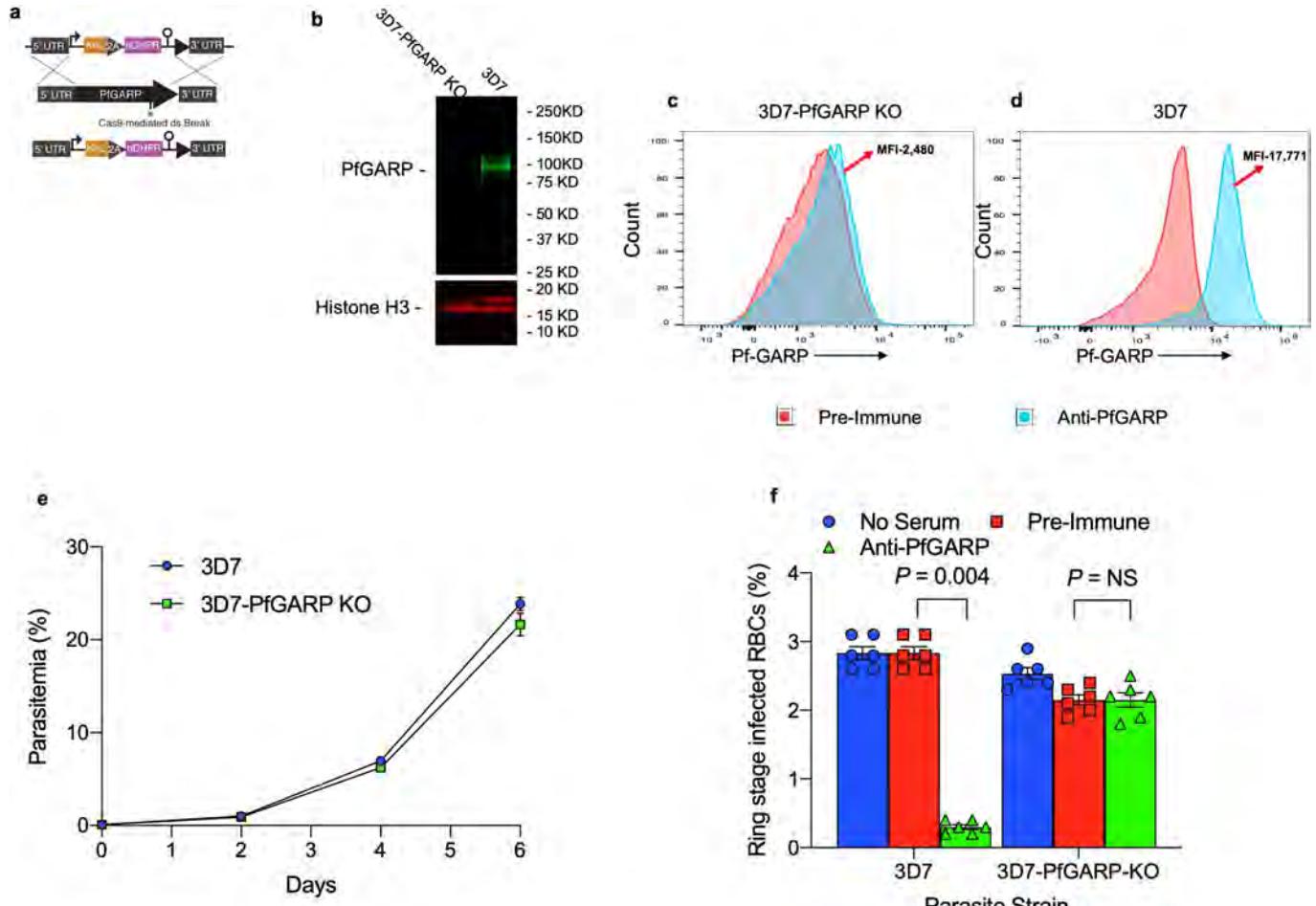
and anti-HA (green) and imaged on an LI-COR Odyssey. **c**, RBCs infected with ring-stage 3D7 parasites were cultured in the presence of medium alone, recombinant anti-PfGARP monoclonal antibody or recombinant anti-fluorescein. Parasites were cultured for 48 h at 37 °C and the levels of lactate were measured in the culture supernatant. Data are mean  $\pm$  s.e.m. of four biologically independent replicates. The  $P$  value was calculated by non-parametric two-sided Mann–Whitney  $U$ -test. **d**, RBCs infected with ring-stage 3D7, D10 or Dd2 parasites were cultured in the presence of medium alone or recombinant anti-PfGARP Fab antibody (1 mg ml<sup>-1</sup>) for 48 h at 37 °C, and ring-stage or early-trophozoite-stage parasites were counted by microscopy. Data are mean  $\pm$  s.e.m. of three biologically independent replicates.  $P$  values were calculated by non-parametric two-sided Mann–Whitney  $U$ -test. Data are representative of two independent experiments.

# Article



**Extended Data Fig. 3 | Construction and characterization of the 3D7-PfGARP KD parasite line.** **a**, Targeting strategy for creating 3D7-PfGARP KD parasites. **b**, Immunoblot analysis of 3D7-PfGARP KD. Parasites were sorbitol-synchronized at the ring stage, incubated with or without anhydrotetracycline (ATC) for 20 h and blots were probed with anti-V5. The expected molecular weight of the PfGARP V5-tagged protein is 124 kDa; however, its apparent mobility is 165 kDa owing to its acidic composition. Lane 1, uninfected RBCs; lane 2, RBCs infected with 3D7-PfGARP KD parasites and cultured without anhydrotetracycline; lane 3, RBCs infected with 3D7-PfGARP KD parasites and cultured with anhydrotetracycline. **c, d**, Expression of PfGARP on the surface of human RBCs infected with 3D7-PfGARP KD parasites. Ring-stage 3D7-PfGARP KD parasites were cultured to the trophozoite stage in the absence (**c**) or presence (**d**) of anhydrotetracycline. Expression of PfGARP

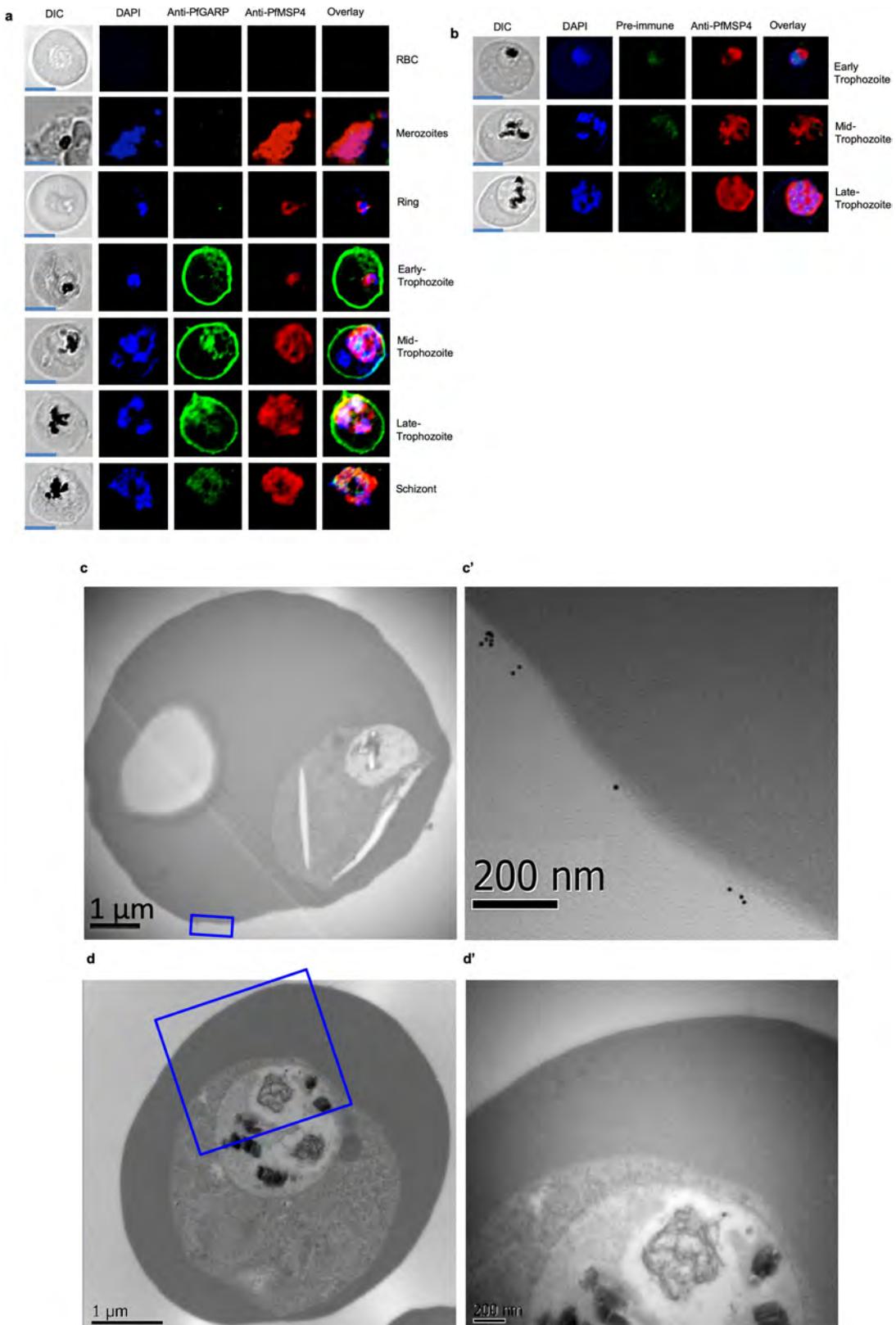
on the surface of fixed but not permeabilized infected human RBCs was determined by flow cytometry, using monoclonal anti-V5 as the primary and anti-mouse IgG-Alexa Fluor 488 as the secondary antibody. Infected RBCs were gated and identified as described in Fig. 1e. **e**, Growth curves for 3D7-PfGARP KD parasites. Ring-stage parasites were cultured with or without anhydrotetracycline. Parasitaemia was measured by microscopy. Data are mean  $\pm$  s.e.m. of three biologically independent replicates. **f**, GIAs using 10% anti-rPfGARP-A serum or pre-immune serum on 3D7-PfGARP KD parasites cultured with or without anhydrotetracycline. Data are mean  $\pm$  s.e.m. of three biologically independent replicates. The *P* value was calculated by non-parametric two-sided Mann–Whitney *U*-test. Data are representative of three independent experiments (**b, d, e, f**) or five independent experiments (**c**).



**Extended Data Fig. 4 | Construction and characterization of the 3D7-PfGARP KO parasite line.** **a**, Targeting strategy for creating 3D7-PfGARP KO parasites. **b**, Immunoblot analysis of 3D7-PfGARP KO. Trophozoite-stage 3D7 wild-type or 3D7-PfGARP KO parasites were probed with anti-PfGARP, and with anti-histone H3 as a loading control. Lane 1, RBCs infected with 3D7-PfGARP KO parasites; lane 2, RBCs infected with 3D7 wild-type parasites. **c**, **d**, Expression of PfGARP on the surface of human RBCs (fixed but not permeabilized) infected with 3D7-PfGARP KO parasites. Ring-stage stage 3D7 wild-type or 3D7-PfGARP KO parasites were cultured to the trophozoite stage. Expression of PfGARP was determined by flow cytometry, using anti-PfGARP as the primary and anti-mouse IgG–Alexa Fluor 488 as the secondary antibody. Infected RBCs were gated and identified as described in Fig. 1e. **e**, Growth

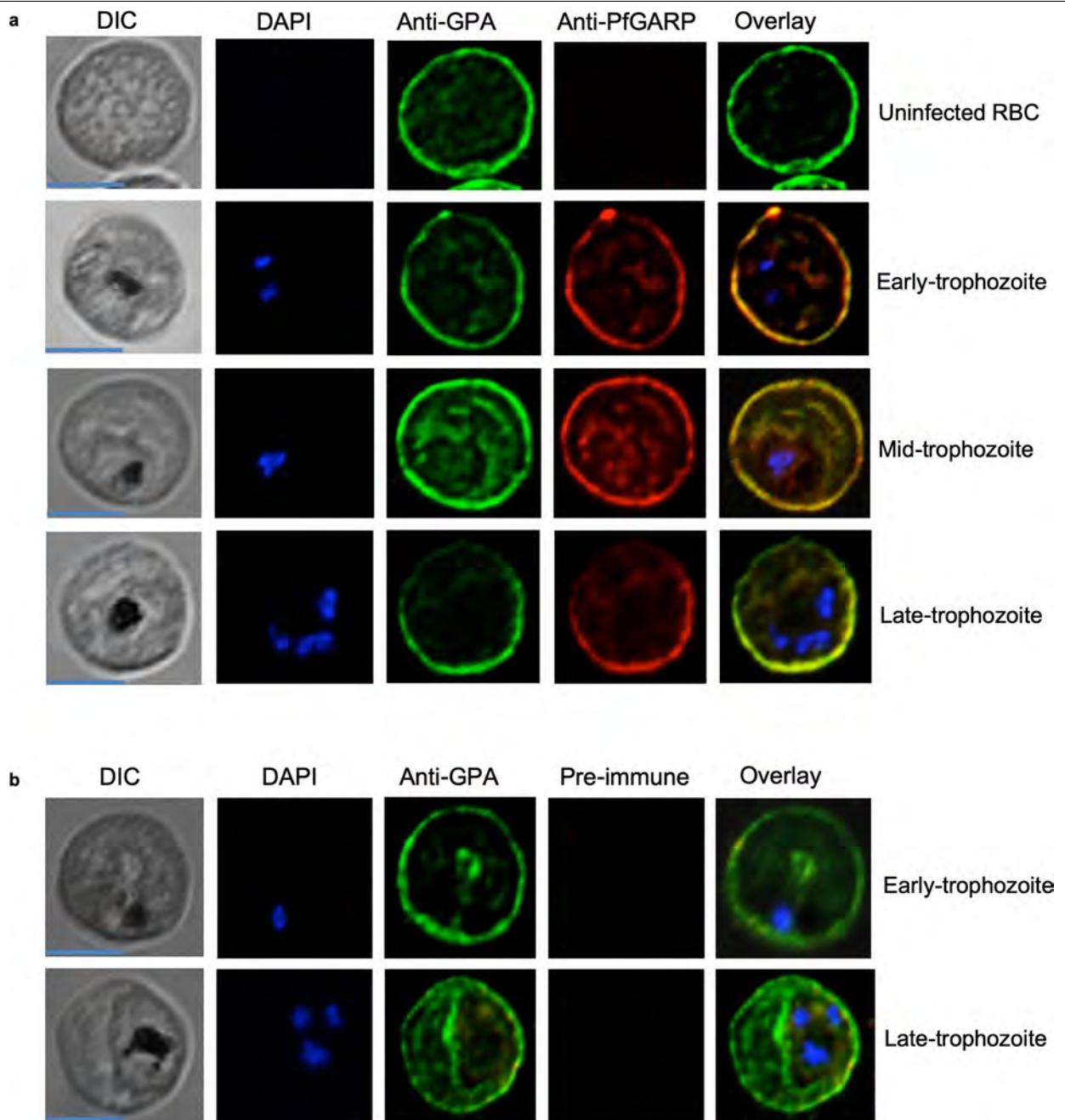
curves for 3D7-PfGARP KO parasites. Ring-stage 3D7 wild-type or 3D7-PfGARP KO parasites were plated at 0.5% parasitaemia and cultured for 6 days. Parasitaemia was measured by microscopy. Data are mean  $\pm$  s.e.m. of three biologically independent replicates. **f**, Ring-stage 3D7 wild-type or 3D7-PfGARP KO parasites with targeted deletion of PfGARP were cultured at a 1:10 dilution in the presence of anti-PfGARP-A mouse serum that was generated by immunizing mice with PfGARP-A mRNA LNPs. Negative controls were no antiserum (blue) and pre-immune mouse serum (red). Parasites were cultured for 48 h at 37 °C and ring-stage and early-trophozoite-stage parasites were counted by microscopy. Data are mean  $\pm$  s.e.m. of six biologically independent replicates. *P* values were calculated by non-parametric two-tailed Mann–Whitney *U*-test. Data are representative of two independent experiments.

# Article



**Extended Data Fig. 5 | Immunolocalization of PfGARP.** **a**, Uninfected and infected RBCs were probed with mouse anti-PfGARP prepared by DNA vaccination (green) and with rabbit anti-PfMSP4 (red) and counterstained with DAPI to label parasite nuclei. PfGARP is detected on the membranes of RBCs infected with early-, mid- and late-trophozoite-stage parasites and does not colocalize with PfMSP4 (which localizes to the parasite membrane). DIC, differential interference contrast microscopy. Scale bars, 5 μm. **b**, Trophozoite-infected RBCs do not label when probed with pre-immune

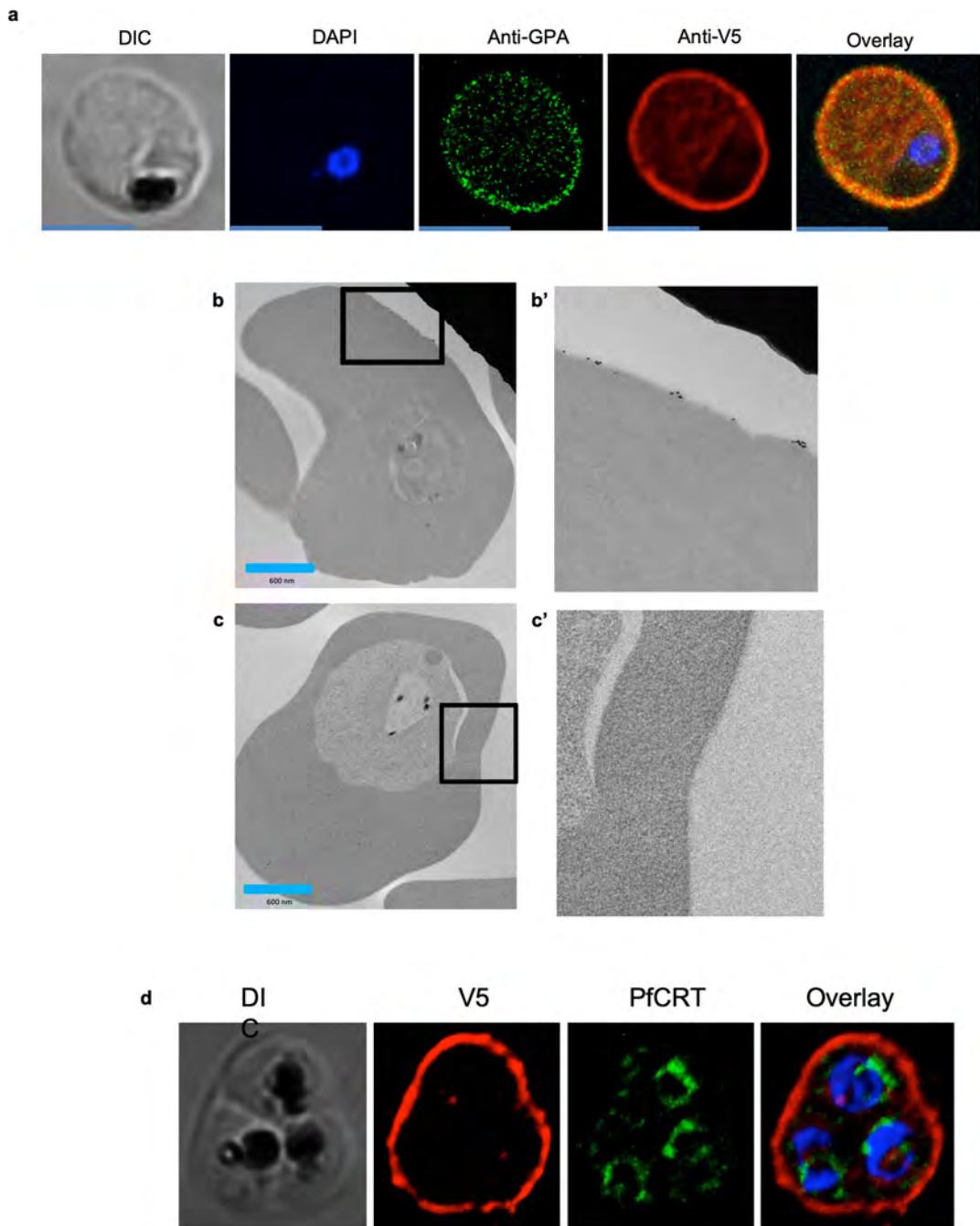
mouse serum. Scale bars, 5 μm. **c, d**, Non-permeabilized, unfixed trophozoite-infected RBCs were incubated with polyclonal anti-rPfGARP (**c**) or control mouse serum (**d**), probed with anti-mouse IgG labelled with 10-nm gold particles, fixed, embedded and visualized by transmission electron microscopy. PfGARP localized to the outer leaflet of trophozoite-infected RBCs. Right, higher-magnification views of the boxed areas on the left. Images are representative of five (**a, b**) or two (**c, d**) biologically independent experiments.



**Extended Data Fig. 6 | PfGARP colocalizes with glycophorin A to the exofacial surface of trophozoite-infected RBCs.** **a**, Uninfected and infected RBCs were probed with rabbit anti-glycophorin A (anti-GYPA; green) and mouse anti-PfGARP prepared by DNA vaccination (red) and counterstained with DAPI to label parasite nuclei. PfGARP is detected only in

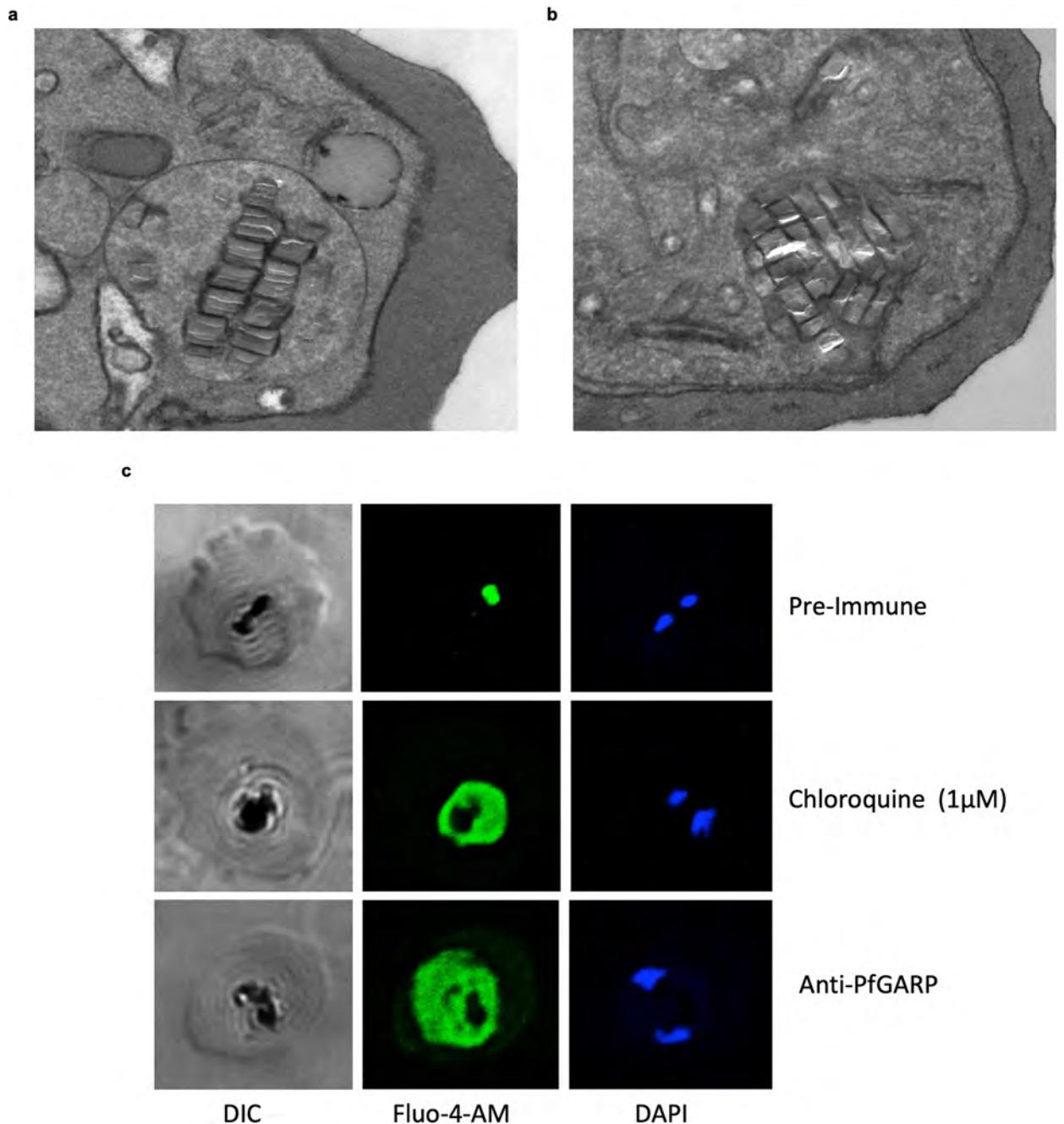
trophozoite-infected RBCs and colocalizes with human glycophorin A on the RBC membrane. Scale bars, 5  $\mu$ m. **b**, Neither early- nor late-trophozoite-stage-infected RBCs label when probed with pre-immune mouse serum. Scale bars, 5  $\mu$ m. Images are representative of three independent experiments (**a, b**).

# Article



**Extended Data Fig. 7 | Localization of VS-tagged PfGARP.** **a**, 3D7-PfGARP KD parasites, in which PfGARP is tagged with the V5 epitope, were grown in the presence of anhydrotetracycline to induce the expression of PfGARP, probed with rabbit anti-glycophorin A (green) and mouse anti-V5 (red) antibodies and counterstained with DAPI to label parasite nuclei. V5-tagged PfGARP colocalizes with glycophorin A to the exofacial surface of trophozoite-infected, non-permeabilized RBCs. Scale bars, 5  $\mu$ m. **b**, 3D7-PfGARP KD parasites were grown in the presence of anhydrotetracycline to induce the expression of PfGARP. RBCs infected with trophozoite-stage 3D7-PfGARP KD parasites were fixed and incubated with anti-V5 mouse IgG (**b**) or buffer (**c**) and probed with

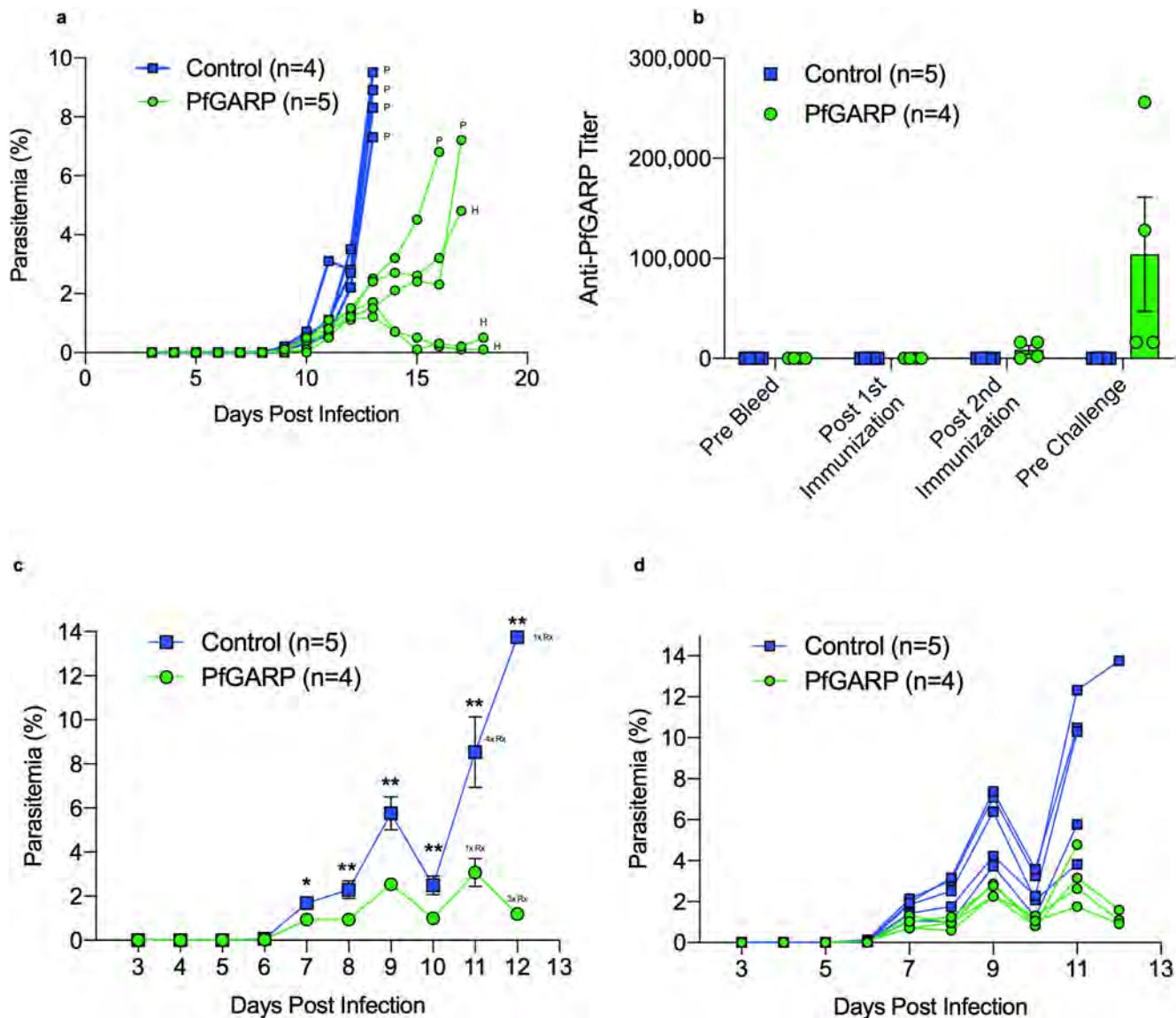
anti-mouse IgG labelled with 10-nm gold particles. V5-tagged PfGARP localized to the outer leaflet of 3D7-PfGARP KD-infected, non-permeabilized RBCs. Right, higher-magnification views of the boxed areas on the left. **d**, 3D7-PfGARP KD parasites were grown in the presence of anhydrotetracycline to induce the expression of PfGARP, probed with anti-PfCRT (green) and anti-V5 (red) antibodies and counterstained with DAPI to label parasite nuclei. PfGARP does not colocalize with PfCRT to the food vacuole in the majority of trophozoite-infected RBCs. Images are representative of three (**a**), two (**b**, **c**) or five (**d**) biologically independent experiments.



**Extended Data Fig. 8 | Anti-PfGARP antibodies disrupt the integrity of the food vacuole.** 3D7 parasites were synchronized to the ring stage and plated at 5% parasitaemia in the presence of pre-immune (**a**) or anti-rPfGARP-A (**b**) mouse serum at a 1:10 dilution. Parasites were cultured for 24 h and processed for transmission electron microscopy. Data are representative of all

trophozoites observed in three biologically independent experiments (**a, b**). **c**, Ring-stage 3D7 parasites were treated with pre-immune or anti-PfGARP serum (1:10) or 1 µM chloroquine for 24 h, followed by staining with DAPI and Fluo-4 AM. Data are representative of two biologically independent experiments.

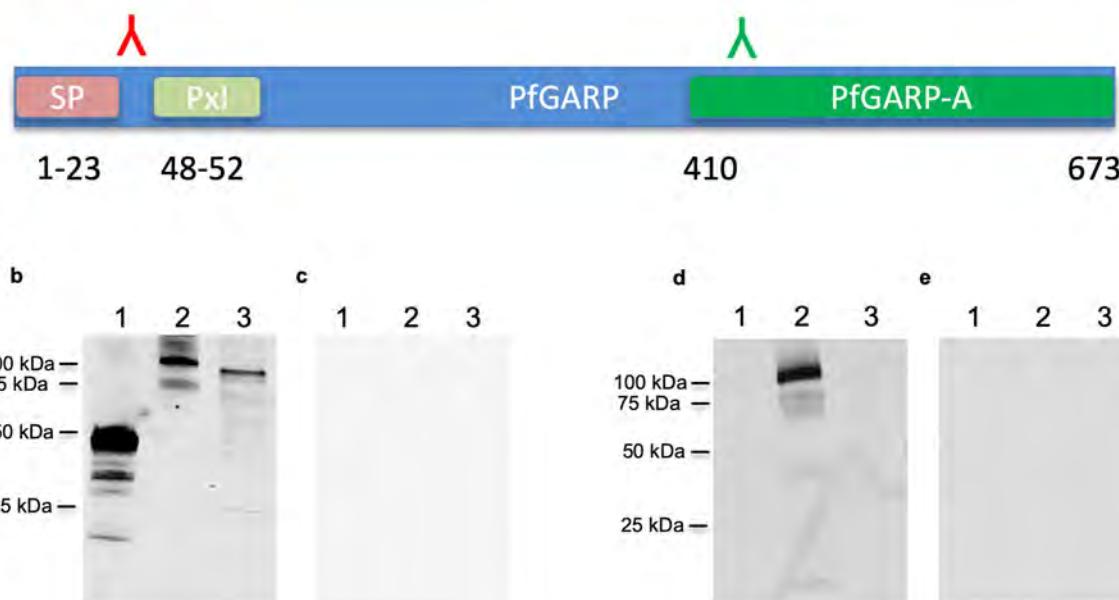
# Article



**Extended Data Fig. 9 | Vaccination with PfGARP protects monkeys from *P. falciparum* challenge.** **a**, Individual parasitaemia data from the monkey trial presented in Fig. 4. *Aotus* monkeys vaccinated with PfGARP-A-mRNA LNP ( $n = 5$ ) and control monkeys vaccinated with poly(C)-RNA LNP ( $n = 4$ ) were challenged intravenously with  $1 \times 10^4$  *P. falciparum*-infected RBCs and parasitaemia was followed daily. H indicates treatment for low haemoglobin; P indicates treatment for high parasitaemia. **b**, Animals were injected subcutaneously either with 50 µg of rPfGARP-A emulsified in 100 µl Ribi ( $n = 4$  monkeys) or with Ribi alone (negative control;  $n = 5$  monkeys) at weeks 0, 3 and 6, and PfGARP-A-specific IgG titres were determined. Data are mean  $\pm$  s.e.m. **c**, Vaccinated *Aotus* monkeys were challenged intravenously with  $10^4$

*P. falciparum* FVO strain-infected RBCs on day 63 and parasitaemia was followed daily. Data are mean  $\pm$  s.e.m. Control monkeys had significantly higher parasitaemia on days 7–12 than monkeys immunized with PfGARP. On day 11, the final day with complete follow-up of all monkeys, control monkeys had 3.5-fold-higher parasitaemia than PfGARP-vaccinated monkeys. Four control monkeys met pre-specified criteria for drug treatment on day 11 and the remaining control monkey met these criteria on day 12. On day 11, one PfGARP-vaccinated monkey underwent drug treatment despite not meeting the pre-specified criteria. \* $P < 0.05$ , \*\* $P < 0.01$  in two-sided *t*-tests without adjustment for multiple comparisons. **d**, Parasitaemia data from the individual monkeys in **c**.

a



**Extended Data Fig. 10 | The PEXEL motif is processed and cleaved in mature PfGARP.** **a**, Schematic depicting the binding sites for the peptide-specific antibodies. **b–e**, Immunoblot of rPfGARP-A (lane 1), recombinant full-length PfGARP (lane 2) and an extract of trophozoite-infected RBCs (lane 3) probed with antibodies raised against amino acids 504–522 of PfGARP (**b**) or pre-immune serum (**c**), or with antibodies raised against amino acids 31–48 of

PfGARP (**d**) or pre-immune serum (**e**). Only antibodies raised against amino acids 504–522 recognized native PfGARP in trophozoite-infected RBCs, whereas antibodies raised against amino acids 31–48 only recognized the full-length recombinant PfGARP, confirming that the PEXEL motif is cleaved during the processing of native PfGARP. PxI, PEXEL motif; SP, signal peptide. Data are representative of three biologically independent experiments.

Corresponding author(s): Jonathan Kurtis

Last updated by author(s): Feb 4, 2020

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

## Software and code

Policy information about [availability of computer code](#)

Data collection

Filemaker Pro v10

Data analysis

SAS v9, JMP v14 and Amira v2019.2

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](http://nature.com/documents/nr-reporting-summary-flat.pdf)

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Both cohort studies were based on open enrolment of participants living in the malaria endemic villages under study.
Data exclusions	no data was excluded
Replication	two independent cohorts were assessed. For in vitro experiments, multiple independent replicates were performed and all were successful.
Randomization	no formal randomization was performed
Blinding	slide readers (enumeration of parasitemia) were blind to treatment status for human blood films, in vitro assays, and the monkey vaccine experiments.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines	<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms		
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants		
<input type="checkbox"/>	<input checked="" type="checkbox"/> Clinical data		

## Antibodies

Antibodies used	Antibodies Catalog no. Clone Lot no. Company F(ab')2-Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 A-11070 853487 Invitrogen Molecular Probes Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 594 A-11032 1985396 Invitrogen Molecular Probes Goat anti-Rat IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 A-A11006 1423045 Invitrogen Molecular Probes F(ab')2-Goat anti-Mouse IgG, IgM (H+L) Secondary Antibody, Alexa Fluor 488 A10684 891190 Invitrogen Molecular Probes Goat anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor 488-5 nm colloidal gold A31565 764359 Invitrogen Molecular Probes IRDye 800CW Goat anti-Rabbit IgG (H + L) Secondary Antibody 926-32211 C30829-02 ODYSSEY imaging system 926-68070 C30723-04 ODYSSEY imaging system FITC anti-human CD235a (Glycophorin A) Antibody 349103 HI264 B138023 BioLegend Inc. Biotin Mouse anti-human IgG 555784 5288794 BD Biosciences APC anti-human CD235a (Glycophorin A), 349113 Clone-HI 264 B262669 Biolegend Inc. V5-Tag Monoclonal antibody R960-25 1937181 Thermo Fisher Scientific
Validation	manufacturer validated each antibody for the intended application

## Eukaryotic cell lines

Policy information about <a href="#">cell lines</a>	
Cell line source(s)	NIH repository (MR4)
Authentication	none
Mycoplasma contamination	not tested

Commonly misidentified lines  
(See [ICLAC](#) register)

none used

## Animals and other organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

Mice  
BALB/cJ both male and female mice (The Jackson Laboratory). Ages 20-26 weeks.

Monkey  
Aotus nancymaae Both male and female (MD Anderson Cancer Center). Adults, age unknown.

Wild animals

none

Field-collected samples

none

Ethics oversight

Animal studies approved by IACUC of Brown University, Rhode Island Hospital and the NIH

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics

Tanzanian Birth cohort (age 0 to 4 years)  
Kenyan Cohort, males only, ages 12-35  
Further details can be found in the MS as well as ref #19 (Kenyan cohort) and ref #2 (Tanzanian cohort)

Recruitment

Kenyan cohort: Open recruitment in the designated malaria endemic villages. Details provided in ref#19  
Tanzanian Cohort: Open recruitment at delivery facilities. Details provided in Ref #2.

Ethics oversight

Ethical clearance was obtained from the IRBs of SBRI and Rhode Island Hospital, the Medical Research Coordinating Committee of the National Institute for Medical Research, Tanzania, and the Kenyan Medical Research Institute.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration

This is not a clinical trial.

Study protocol

*Note where the full trial protocol can be accessed OR if not available, explain why.*

Data collection

*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*

Outcomes

*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

## Flow Cytometry

### Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

Sample preparation

After treatment in culture, the infected RBCs were harvested and washed two times with PBS. After thorough washing, RBCs were appropriately stained and suspended in PBS before acquisition and analysis by using flow cytometer.

Instrument

BD LSR II Flow Cytometer was used for data collection

Software	BD FACS Diva 6.1.1 was used for data collection Flowjo vX.0.7 was used for data analysis
Cell population abundance	No sorting was done in the study; Infected RBCs were identified and gated using standard markers
Gating strategy	Single cells were gated using SSC-H and SSC-A; Normal and Infected RBCs were identified and gated by using Hoechst 33342 and APC conjugated anti-human Glycophorin A antibody. Infected RBCs were identified as cells which were positive for both Glycophorin A and Hoechst 33342. Normal RBCs were identified as cells which were positive for Glycophorin A but negative for Hoechst 33342. Among the infected RBCs, the cells, which were positive for JC-1 red high, represent live parasites and cells, which were JC-1, red low, represent dead parasites. Gating strategy provided in the main Figure 3 H and I.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

## Article

# Mouse models of neutropenia reveal progenitor-stage-specific defects

<https://doi.org/10.1038/s41586-020-2227-7>

Received: 24 September 2018

Accepted: 24 February 2020

Published online: 22 April 2020

 Check for updates

David E. Muench<sup>1,2</sup>, Andre Olsson<sup>1</sup>, Kyle Ferchen<sup>1,3</sup>, Giang Pham<sup>4</sup>, Rachel A. Serafin<sup>1</sup>, Somchai Chutipongtanate<sup>3,5</sup>, Pankaj Dwivedi<sup>3</sup>, Baobao Song<sup>1,6</sup>, Stuart Hay<sup>7</sup>, Kashish Chetal<sup>7</sup>, Lisa R. Trump-Durbin<sup>8</sup>, Jayati Mookerjee-Basu<sup>9</sup>, Kejian Zhang<sup>10,11</sup>, Jennifer C. Yu<sup>12,13</sup>, Carolyn Lutzko<sup>8,11</sup>, Kasiani C. Myers<sup>11,14</sup>, Kristopher L. Nazor<sup>15</sup>, Kenneth D. Greis<sup>3</sup>, Dietmar J. Kappes<sup>9</sup>, Sing Sing Way<sup>4,11</sup>, Nathan Salomonis<sup>7,11</sup> & H. Leighton Grimes<sup>1,8,11</sup>

Advances in genetics and sequencing have identified a plethora of disease-associated and disease-causing genetic alterations. To determine causality between genetics and disease, accurate models for molecular dissection are required; however, the rapid expansion of transcriptional populations identified through single-cell analyses presents a major challenge for accurate comparisons between mutant and wild-type cells. Here we generate mouse models of human severe congenital neutropenia (SCN) using patient-derived mutations in the GFI1 transcription factor. To determine the effects of SCN mutations, we generated single-cell references for granulopoietic genomic states with linked epitopes<sup>1</sup>, aligned mutant cells to their wild-type equivalents and identified differentially expressed genes and epigenetic loci. We find that GFI1-target genes are altered sequentially, as cells go through successive states of differentiation. These insights facilitated the genetic rescue of granulocytic specification but not post-commitment defects in innate immune effector function, and underscore the importance of evaluating the effects of mutations and therapy within each relevant cell state.

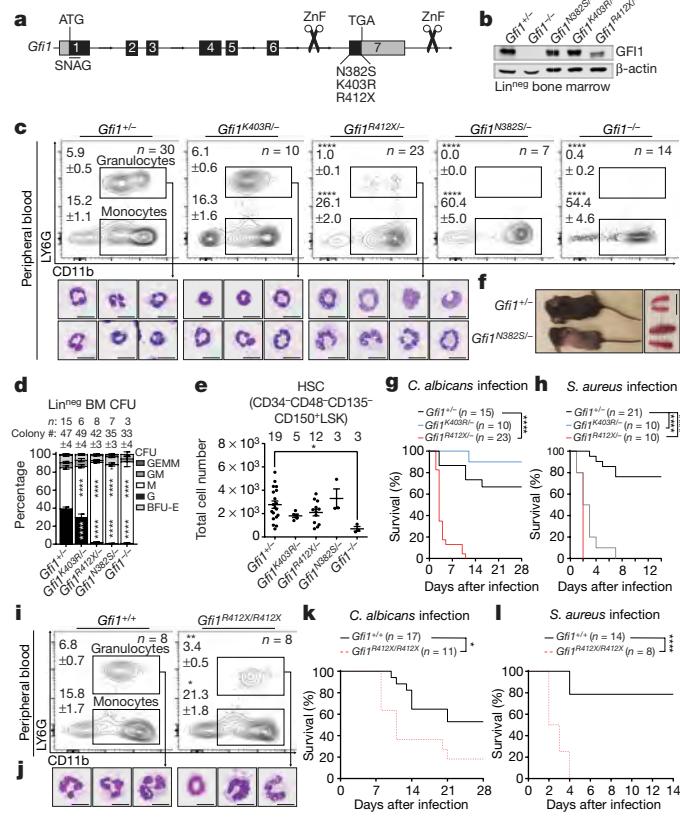
In patients with SCN, inherited and de novo mutations lead to a profound block in neutrophil granulopoiesis. We introduced mutations in the GFI1 transcription factor derived from patients with SCN into the mouse *Gfi1* allele, which resulted in steady-state mouse dysgranulopoiesis and broad in vivo susceptibility to neutrophil-dependent pathogens. Notably, SCN mutations separate neutropenia from *Gfi1*<sup>-/-</sup> defects in haematopoietic stem-cell and lymphocyte numbers. To delineate the mechanism, we mapped the successive genomic states encompassing neutrophil granulocyte specification (ability to form neutrophils) and commitment (cell-fate restriction)<sup>2</sup>, including a new transitional intermediate (AltAnalyze ICGS Viewer; <http://www.altanalyze.org/ICGS/Neutrophil/>), then confirmed the trajectory. Using comparative genomics<sup>3,4</sup>, we assigned GFI1-mutant cells to wild-type cell states to determine differential single-cell gene expression and chromatin. Few GFI1-target genes are deregulated across granulopoiesis, but are instead deregulated in successive granulopoietic cell states; associated with failure of GFI1-bound chromatin to close after specification. However, genetically rescuing specification did not resolve post-commitment defects in immune defence in mice or in G-CSF-rescued neutrophils from human patients with SCN. These experiments underscore the

dominance of cell state in integrating the effects of mutations and therapy, and illustrate a workflow that can be broadly applied to molecularly investigate translationally relevant models of disease.

## Patient mutations induce neutropenia

Clinical resequencing of 225 patients with neutropenia identified previously reported<sup>5,6</sup> and new *Gfi1* sequence polymorphisms of unknown importance (Supplementary Information), which we screened for defects in repressing an *Irf8-eGFP* reporter allele<sup>7</sup> (Extended Data Fig. 1a, b). Next, we edited endogenous mouse *Gfi1* alleles to introduce N382S, K403R or R412X mutations (Fig. 1a, Extended Data Fig. 1c–h, Supplementary Information). Mutant GFI1 proteins accumulate normally, but the R412X mutation results in a smaller, less abundant protein (Fig. 1b, Extended Data Fig. 1i). On postnatal day 1, only *Gfi1*<sup>N382S/+</sup> mice exhibit a blunted response to neonatal cues driven by gut bacteria<sup>8</sup>, comparable to the exacerbated clinical presentation of SCN in neonatal humans (Extended Data Fig. 1j, k). Human *Gfi1* is predicted to be imprinted<sup>9</sup>. RNAscope analysis with *Gfi1* intronic probes (Extended Data Fig. 1l) revealed a pattern that was consistent with monoallelic expression<sup>10</sup>.

<sup>1</sup>Division of Immunobiology and Center for Systems Immunology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. <sup>2</sup>Molecular and Developmental Biology Graduate Program, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. <sup>3</sup>Department of Cancer Biology, University of Cincinnati, Cincinnati, OH, USA. <sup>4</sup>Division of Infectious Diseases, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. <sup>5</sup>Pediatric Translational Research Unit, Department of Pediatrics, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand. <sup>6</sup>Immunology Graduate Program, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. <sup>7</sup>Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. <sup>8</sup>Division of Experimental Hematology and Cancer Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. <sup>9</sup>Fox Chase Cancer Center, Philadelphia, PA, USA. <sup>10</sup>Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. <sup>11</sup>Department of Pediatrics, University of Cincinnati, Cincinnati, OH, USA. <sup>12</sup>Division of Pediatric Hematology/Oncology, Rady Children's Hospital San Diego, San Diego, CA, USA. <sup>13</sup>Department of Pediatrics, University of California, San Diego, La Jolla, CA, USA. <sup>14</sup>Division of Bone Marrow Transplantation and Immune Deficiency, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. <sup>15</sup>BioLegend, Inc., San Diego, CA, USA. <sup>✉</sup>e-mail: nathan.salomonis@cchmc.org; lee.grimes@cchmc.org



**Fig. 1 | GFI1 mutations derived from patients with neutropenia induce steady-state dysgranulopoiesis and broad immune defects in mice.** **a.** Schematic of the *Gfi1* locus; locations of mutations introduced by genome editing. ZnF, zinc-finger nuclelease. **b.** Immunoblot of adult mouse Lin<sup>neg</sup> bone marrow. **c.** FACS plots with percentages of total adult peripheral blood, and cytospins of FACS-sorted CD11b<sup>+</sup> LY6G<sup>+</sup> cells. **d.** Colony-forming unit (CFU) assays performed on Lin<sup>neg</sup> bone marrow (BM) cells. BFU-E, burst-forming-unit erythroid; G, granulocyte; GEMM, granulocyte, erythrocyte, monocyte, megakaryocyte; GM, granulocyte-monocyte; M, monocyte. **e.** FACS analysis of adult mouse bone marrow. HSC, haematopoietic stem cell; LSK, Lin<sup>neg</sup> SCA-1<sup>+</sup> KIT<sup>+</sup>. **f.** Representative images of adult mice and spleens. **g, h.** Survival of mice infected with a median lethal dose (LD<sub>50</sub>) of *C. albicans* (g) or 5 × 10<sup>7</sup> CFUs of *S. aureus* (h). Numbers of mice are in parentheses. **i, j.** FACS plots (i) and cytospins (j) from R412X homozygotes. **k, l.** Survival of mice infected with an LD<sub>50</sub> dose of *C. albicans* (k) or 5 × 10<sup>7</sup> CFU of *S. aureus* (l). Mouse numbers are in parentheses. Data mean ± s.e.m. from independent biological replicates in c–e and i. Data in b represent three independent experiments. Images in f are representative of two littermate mice for each genotype. \*P < 0.05, \*\*P < 0.01, \*\*\*P < 0.0001, two-tailed t-test (c–e, i) or two-sided Mantel-Cox test (g, h, k, l). Scale bars, 10 μm (c, j) and 1 cm (f).

Human induced pluripotent stem (iPS) cells with one *GFI1*<sup>R412X</sup> allele only express the mutant *GFI1* allele (Extended Data Fig. 1m), and are blocked in granulopoiesis (Extended Data Fig. 1n, o). Therefore, to mimic predicted human gene regulation, we genetically silenced one *Gfi1* allele<sup>11</sup> and found *Gfi1*<sup>R412X</sup><sup>–/–</sup>, *Gfi1*<sup>N382S</sup><sup>–/–</sup> and *Gfi1*<sup>–/–</sup> neonatal mice markedly neutropenic (Extended Data Fig. 1j, k). Peripheral blood analyses revealed a range of neutropenia (Fig. 1c, Extended Data Fig. 2a), and *Gfi1*<sup>R412X</sup><sup>–/–</sup> and *Gfi1*<sup>N382S</sup><sup>–/–</sup> mice exhibited monocytosis as a result of cell-intrinsic lineage skewing (Fig. 1c, d, Extended Data Fig. 2a–c), supported by a lack of mature neutrophils in the bone marrow (Extended Data Fig. 2d–f). *Gfi1*<sup>R412X</sup><sup>–/–</sup> neutrophils exhibited abnormal or immature morphologies (Fig. 1c), which suggests a *GFI1* function beyond neutrophil specification. Notably, *Gfi1*<sup>–/–</sup>, *Gfi1*<sup>N382S</sup><sup>–/–</sup> and *Gfi1*<sup>R412X</sup><sup>–/–</sup> mice display steady-state SCA-1<sup>+</sup> emergency granulocyte-monocyte-progenitor (eGMP) gate cells<sup>12,13</sup> (Extended Data Fig. 3a). However, as with serum from patients with SCN<sup>14</sup>, extensive profiling of *Gfi1*<sup>R412X</sup><sup>–/–</sup> serum revealed increased

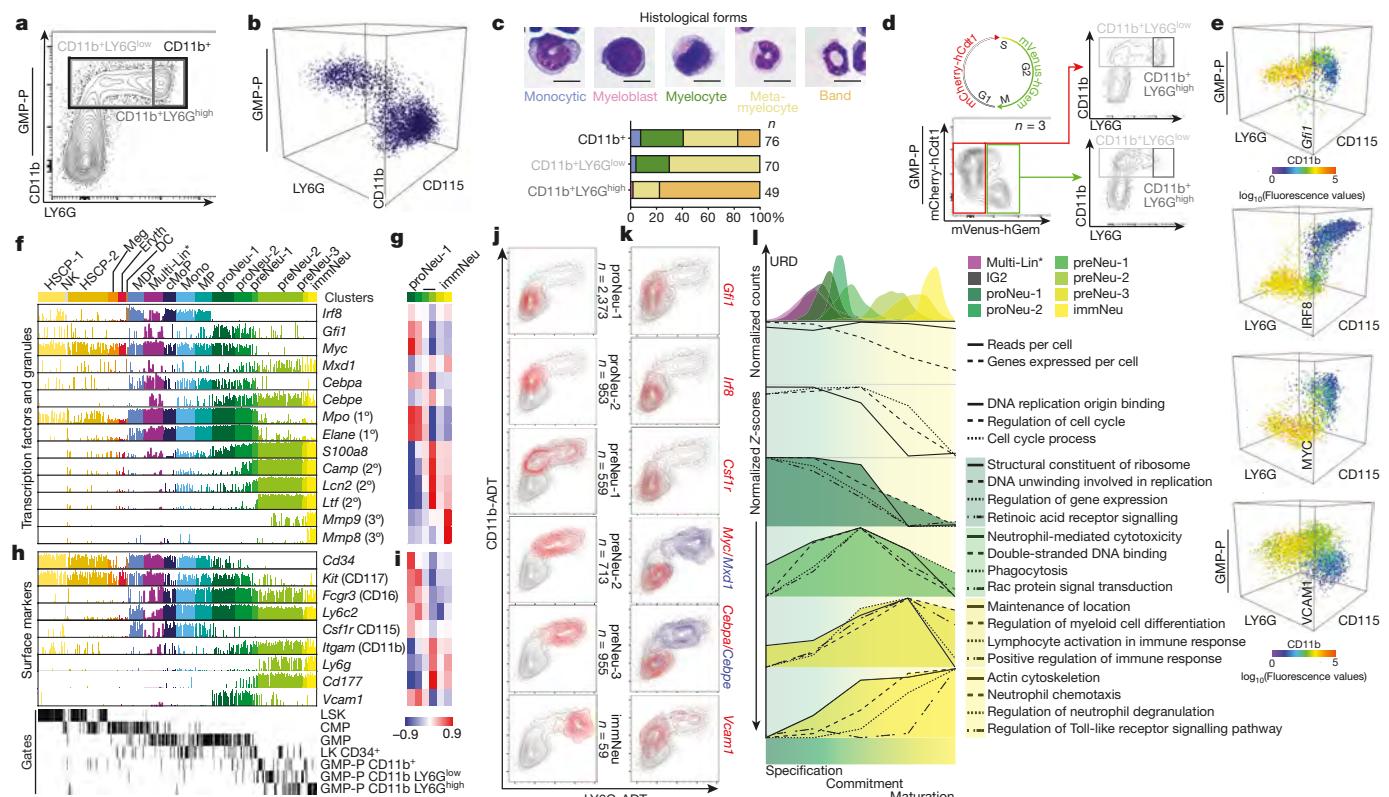
levels of G-CSF, but not eGMP-associated inflammatory cytokines (Extended Data Fig. 3b, c, Supplementary Information).

Unlike *Gfi1*<sup>–/–</sup> mice, the numbers of phenotypic haematopoietic stem cells were normal in *GFI1*-mutant mice (Fig. 1e), but the fitness of haematopoietic stem cells from *Gfi1*<sup>N382S</sup><sup>–/–</sup> mice was defective after transplantation. By contrast, haematopoietic stem cells from *Gfi1*<sup>R412X</sup><sup>–/–</sup> mice rescue lethal irradiation, but display a modest defect in chimerism during transplant (Extended Data Fig. 3d, e). Similar to *Gfi1*<sup>–/–</sup> mice<sup>11,15–17</sup>, *Gfi1*<sup>N382S</sup><sup>–/–</sup> mutants displayed several defects, including runting, lymphopenia and splenomegaly (Fig. 1f, Extended Data Fig. 2b), and rarely survive to adulthood. However, neither *Gfi1*<sup>K403R</sup><sup>–/–</sup> nor *Gfi1*<sup>R412X</sup><sup>–/–</sup> mutants recapitulate these *Gfi1*<sup>–/–</sup> defects (Extended Data Figs. 2b, 3d, e), so we focused on *Gfi1*<sup>K403R</sup><sup>–/–</sup> and *Gfi1*<sup>R412X</sup><sup>–/–</sup> mice to isolate *GFI1* control of granulopoiesis.

*Gfi1*<sup>R412X</sup><sup>–/–</sup> mice were susceptible to *Candida albicans*, with reduced survival and increased pathogen burden (Fig. 1g, Extended Data Fig. 3f). By contrast, *Gfi1*<sup>R412X</sup><sup>–/–</sup> and *Gfi1*<sup>K403R</sup><sup>–/–</sup> mice showed increased susceptibility to *Staphylococcus aureus* (Fig. 1h, Extended Data Fig. 3g). Impaired host defence in *Gfi1*<sup>R412X</sup><sup>–/–</sup> mice probably does not reflect reduced levels of *GFI1* protein, because homozygous mutant alleles in *Gfi1*<sup>R412X/R412X</sup> mice partially restored both *GFI1* protein levels (Extended Data Fig. 3h) and neutrophil numbers (Fig. 1i) albeit with abnormal morphology (Fig. 1j), but did not rescue susceptibility to *C. albicans* and *S. aureus* (Fig. 1k, l). Thus, *GFI1* control of granulocyte production is functionally dissociated from innate host defence against invasive microbial infection.

## Neutrophil specification and commitment

First, we established the genomic states encountered during neutrophil granulopoiesis using a novel flow gate (granulocyte-monocyte-progenitor and precursor ‘GMP-P’ gate) (Extended Data Fig. 4a), with reciprocal expression of the granulocytic marker LY6G and the monocytic marker CD11b (encoded by *Csf1r*) based on CD11b expression levels (Fig. 2a, b, Extended Data Fig. 4a). Histological forms of flow-cytometry-sorted cells, together with fluorescence-activated cell sorting (FACS) plots of GMP-P gate cells from Fucci2 transgenic mice<sup>18</sup>, as well as mice bearing *Gfi1*-locus, or IRF8 and MYC fluorescent-protein reporter alleles, confirmed cells undergoing terminal granulopoiesis (Fig. 2c–e, Extended Data Fig. 4a). Notably, *Gfi1* expression is highest in CD11b<sup>int</sup>LY6G<sup>neg</sup> cells (Fig. 2e). These analyses suggest that CD11b<sup>int</sup>LY6G<sup>neg</sup> cells are specifying, whereas CD11b<sup>high</sup>LY6G<sup>neg-to-high</sup> cells are committing to neutrophil granulopoiesis. We captured these cells (Extended Data Fig. 4b, c) and incorporated them into previous single-cell RNA sequencing (scRNA-seq) datasets<sup>19,20</sup> (Extended Data Fig. 4d, Supplementary Information). An unsupervised discovery workflow (ICGS)<sup>19</sup> (Extended Data Fig. 4c) revealed a notable shift in gene expression between early specifying (‘proNeu’), late committed (‘preNeu’) and immature neutrophil (‘immNeu’) granulocytic clusters, as well as a transitional population (‘preNeu-1’) that expressed both programs. The names proNeu, preNeu and immNeu were adapted from ref.<sup>21</sup>. To validate the cluster order, we identified known reciprocal patterns of transcription factor genes (*Irf8/Gfi1*, *Myc/Mxd1* and *Cebpa/Cebpe*), as well as primary (*Elane* and *Mpo*), secondary (*Camp*, *Lcn2* and *Ltf*), and tertiary (*Mmp8* and *Mmp9*) granule protein-encoding genes (Fig. 2f), the expression of which was considerably correlated with individual clusters (Fig. 2g). We delineated surface markers that are correlated with ICGS clusters (Fig. 2h, i, Extended Data Fig. 4e–g) to find those that are extinguished (*Cd34*, *Csf1r* and *Kit*) or induced (*Itgam*, *Ly6g* and *Cd177*). Notably, *Vcam1* is uniquely expressed in granulocyte-specifying cells, as confirmed by flow cytometry analysis plus 3D visualization of GMP-P gate cells for CD106 expression (Fig. 2e), which is similar to the *Gfi1*-locus reporter (Fig. 2e). Expression of cell cycle genes terminated in the last cluster (Extended Data Fig. 4h–l), along with an overall decrease in expressed genes at the preNeu-1 stage that was



**Fig. 2 | Cell states traversed during commitment to terminal granulopoiesis include a rare transitional state (bridging specification to commitment).**

**a**, FACS plot with annotated gates used to sort cells for scRNA-seq (Extended Data Fig. 4a). **b**, Three-dimensional FACS plot of GMP-P gate (Extended Data Fig. 4a). **c**, Quantified histological forms of FACS-sorted cells (from gates in **a**). **d**, Schematic of fluorescent cell cycle reporter and representative FACS plots of GMP-P gate from Fucci2 transgenic mice. **e**, Three-dimensional FACS plots of bone marrow isolated from adult mice bearing *Gfi1*-locus, IRF8 or MYC fluorescent-protein reporters, or VCAMI (CD106) expression. Events are pseudocoloured for CD11b expression. **f**, Bar chart displaying incidence and amplitude for selected genes (Fluidigm C1) (Extended Data Fig. 4c) ( $n = 509$  cells). ICGS clusters are annotated (top), and FACS gates are annotated (bottom). **g**, Heat map of correlation between gene expression and each

displayed cluster (MarkerFinder). **h**, Bar chart displaying selected genes (cells as in **f**). **i**, Heat map of correlation between gene expression and each displayed cluster (MarkerFinder). **j**, **k**, Plots of ADT unique molecular identifier (UMI) counts (10x Genomics technology) (Extended Data Fig. 5b, e), in which grey indicates all captured cells. In **j**, red identifies cells within ICGS clusters (right). In **k**, red or blue indicates the top 1% of cells expressing the indicated gene(s). **l**, URD pseudotime analysis (of data in Extended Data Fig. 5b), and transcriptional features, or normalized Z-scores of enriched biological processes (of data in Extended Data Fig. 4c). Data in **a**, **b**, **d**, **e** are representative of three independent biological replicates. Data in **c** represent cumulative total cell numbers from one experiment using cells pooled from three male mice. Data in **g** and **i** display Pearson correlation values. CMP, common myeloid progenitor; LK, CD34<sup>+</sup>, Lin<sup>neg</sup> KIT<sup>SCA-1<sup>neg</sup></sup> CD34<sup>+</sup>.

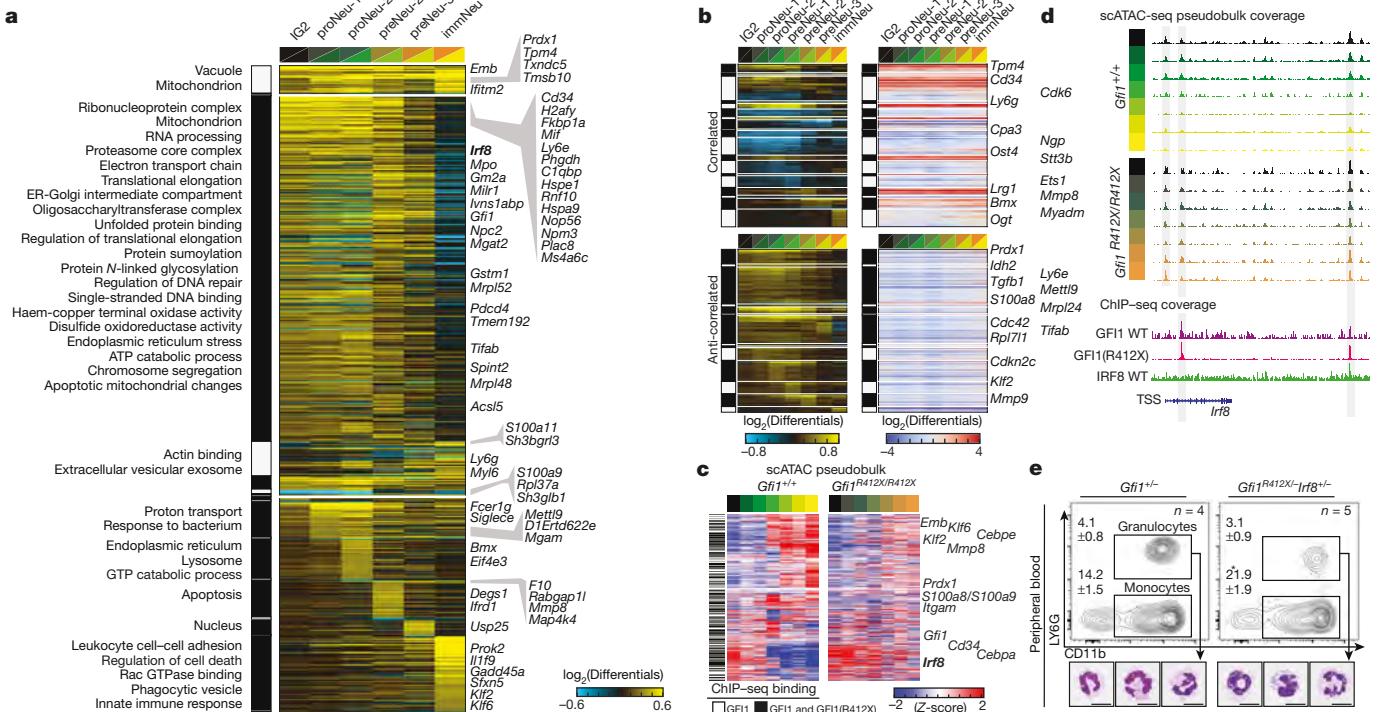
independent of the number of reads per cell (Extended Data Fig. 4m), which suggests that global transcriptional silencing is concomitant with granulocytic commitment.

To directly link GMP-P flow gates (Fig. 2a, Extended Data Fig. 4a) to cell clusters (Extended Data Fig. 4c), we performed cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq)<sup>1</sup> (Fig. 2j, Extended Data Fig. 5a). We supervised analysis of the CITE-seq data using centroids of ICGS clusters<sup>19</sup> (Extended Data Fig. 4c) to classify the newly captured cells (Extended Data Fig. 5b–e), which replicated mixed-lineage and specifying cell states (including multi-lineage primed<sup>19</sup>, *Irf8*-expressing GMP subpopulation 2<sup>19</sup> and preNeu-1 transitional populations) (Extended Data Fig. 5b, d) with no apparent sex-based bias in any cluster (Extended Data Fig. 5b, bottom). The expression of myeloid-associated, cell cycle and surface-marker genes in the CITE-seq data displayed similar trends (compare Fig. 2f, g and Extended Data Fig. 4h, i to Extended Data Fig. 5f–i). Transformed counts of antibody-dependent tags (ADTs) revealed patterns that closely resembled FACS data (compare Extended Data Fig. 5a with Extended Data Fig. 5e). Notably, surface expression of LY6G is discordant with the underlying processes, creating overlap between clusters of neutrophil progenitors and immature neutrophils. Plotting ICGS clusters or individual genes (significantly correlated with clusters) illustrates a similar progression through the CD11b versus LY6G plots (Fig. 2j, k,

Extended Data Fig. 5j, k). Plotting the top 1% of cells expressing *Gfi1*, *Irf8*, *Csf1r*, *Myc*, *Mxd1*, *Cebpa*, *Cebpe* and *Vcam1* (Fig. 2k) confirmed the trajectory, which was further evidenced using pseudotemporal ordering with URD<sup>22</sup> (Fig. 2l, top). Clusters were confirmed in a published dataset<sup>23</sup> (Extended Data Fig. 5l), and reconciled to reported neutrophil progenitor subsets (NMP<sup>24</sup>, NeuP<sup>25</sup>, preNeu<sup>26</sup> and NeP<sup>23</sup>) using 13 of the surface markers used to identify them. Notably, NMP, NeuP, preNeu and NeP are not pure populations and span several ICGS clusters (Extended Data Fig. 5m, n). Next, we identified key pathways and biological processes that are dynamically regulated during terminal granulopoiesis (Fig. 2l, Extended Data Fig. 4l). Specification and commitment are accompanied by distinct biological processes inducing withdrawal from the cell cycle, and a general collapse of gene expression concurrent with the sequential induction of expression of innate immune effector genes (Fig. 2l).

## Mutation affects cell states differently

Given references for granulopoietic genomic states, we could now determine the transcriptional deregulation underlying the *Gfi1*<sup>R412X</sup>– defects in neutrophil production and function. We extended CITE-seq and data complementary Fluidigm captures to granulocyte–monocyte progenitors (GMPs) from *Gfi1*<sup>+/−</sup> and *Gfi1*<sup>R412X</sup><sup>−/−</sup> mice<sup>19</sup> (Extended Data



**Fig. 3 | A single mutation in a lineage-determining transcription factor differentially affects target gene expression in distinct cell states.** **a**, Heat map of differential expression (cellHarmony). scRNA-seq libraries from *Gfi1*<sup>R412X/+</sup> and *Gfi1*<sup>−/−</sup> mice were analysed (10x Genomics gene expression modified-GMP gate and LY6<sup>high</sup> GMP-P gate), then restricted to GFI1 ChIP-seq targets (841 out of 1,462 total), and genes commonly deregulated across 10x Genomics and Fluidigm platforms are displayed (right) (Extended Data Fig. 6k). Each row represents a single gene, each column denotes the fold difference, and enriched Gene Ontology terms (left). **b**, Differential cicero gene activity scores correlated

(top) or anticorrelated (bottom) with scRNA-seq differentially expressed genes (from **a**). **c**, Heat maps of scATAC-seq pseudo-bulk data from GFI1- and GFI1(R412X)-bound loci (black bars on left, and Extended Data Fig. 8e). ICGS clusters (top). Each row is a gene locus (right). **d**, Top, schematic of the *Irf8* locus displaying scATAC-seq pseudobulk accessibility with regions bound by GFI1 shaded (top). Bottom, ChIP-seq reads. ICGS clusters (left). WT, wild type. **e**, FACS plots and representative cytopins of FACS-sorted CD11b<sup>+</sup> LY6G<sup>+</sup> cells from adult mouse peripheral blood. In **e**, data are mean ± s.e.m. from independent biological replicates. \**P* < 0.05, two-tailed *t*-test. Scale bars, 10 μm.

Figs. 4b, 6a–k, 7a–d and <http://www.altaanalyze.org/ICGS/Neutrophil/>). Next, we applied a computational workflow (cellHarmony<sup>3</sup>) (Extended Data Fig. 6b) to assign GFI1-mutant cells to ICGS2 (ref.<sup>27</sup>)-extended clusters to determine differential gene expression. Most *Gfi1*<sup>R412X/+</sup> GMPs are monocytic progenitors (51% by CITE-seq, 55% by Fluidigm), with an approximately threefold increase compared with *Gfi1*<sup>+/+</sup> or *Gfi1*<sup>−/−</sup> cells. Few cells were assigned beyond the proNeu-1 cluster (6.5-fold decrease) (Extended Data Figs. 6c, d, 7b, c). However, capturing rare *Gfi1*<sup>R412X/+</sup> CD11b<sup>high</sup> GMP-Ps (Extended Data Fig. 6a) revealed cells within each of the granulocytic clusters (Extended Data Figs. 6c–e, 7b, c) exhibiting transcriptional profiles resembling that of wild-type GMP-Ps (Extended Data Fig. 4c), including known myeloid, cell-cycle and surface-marker genes (Extended Data Fig. 6f). Notably, of GFI1 targets defined by chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq)<sup>19</sup> (57% of differentially expressed genes), only a minority are disrupted across all clusters (see *Gfi1*, *Mpo*, *Irf8* and *Mmp8* expression in Extended Data Fig. 6f). Instead, using either 10x Genomics or Fluidigm technology, most deregulated expression occurs during specification or in specific clusters (Fig. 3a, Extended Data Fig. 6k).

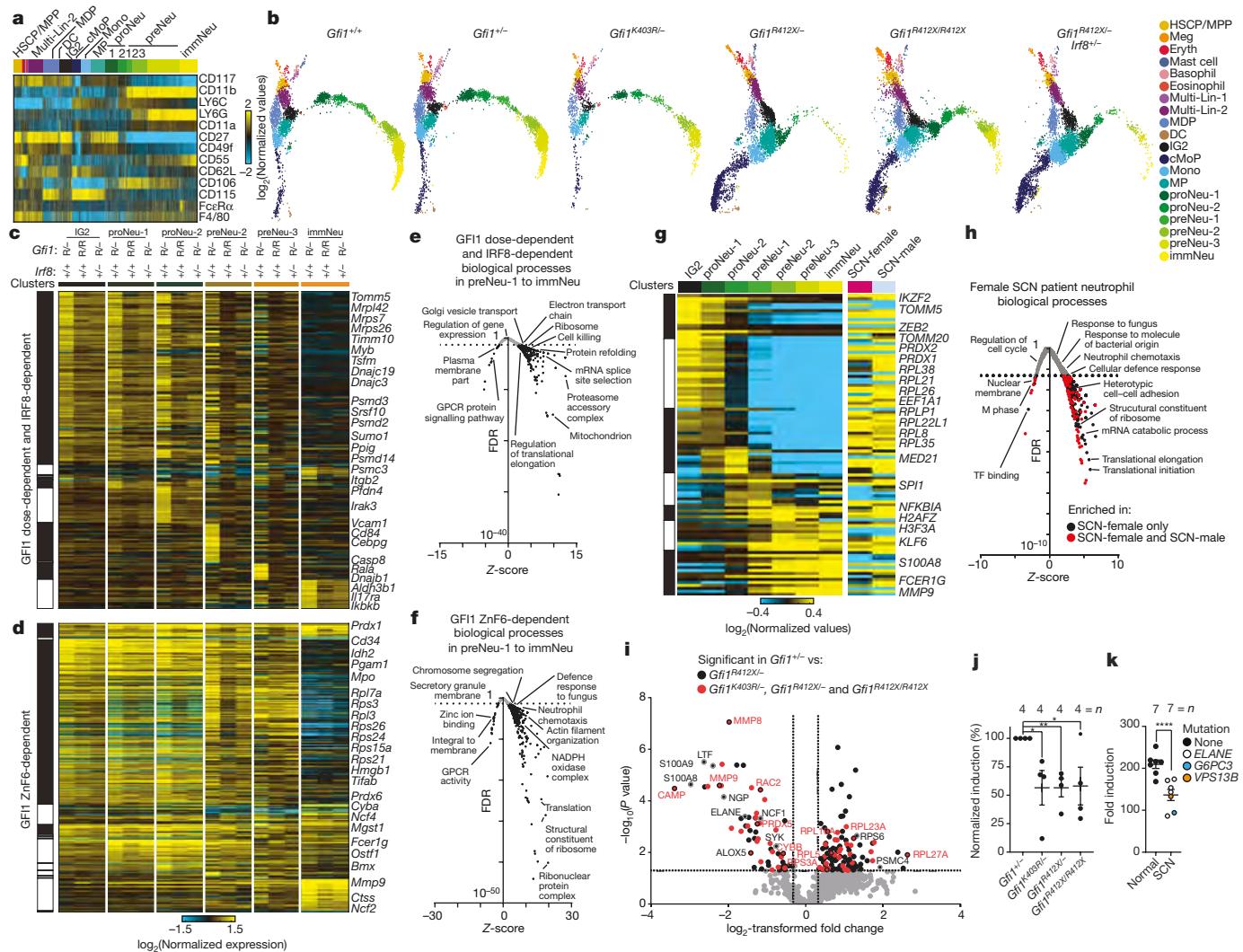
To understand the underlying gene regulatory mechanisms, we performed single-cell assay for transposase-accessible chromatin using sequencing (scATAC-seq) on modified-GMP-gate cells from *Gfi1*<sup>+/+</sup> and *Gfi1*<sup>R412X/R412X</sup> mice (Extended Data Fig. 7e–j). The scATAC-seq clusters were aligned to CITE-seq cell populations (Extended Data Fig. 7e, f) using label transfer in Seurat3<sup>4</sup> (consistent with SNAP-ATAC<sup>28</sup>) (Extended Data Fig. 7i), and cluster-specific ATAC peaks were associated with lineage-specific genes and sequential neutrophil programming (Extended Data Fig. 7k, Supplementary Information). We computed

differentially accessible loci (cicero gene-activity scores), identified regions that correlated or anti-correlated with differential gene expression (Fig. 3b), and inferred dynamic changes in transcription factor activity (Extended Data Fig. 7l). Accessible DNA in *Gfi1*<sup>R412X/R412X</sup> clusters were differentially enriched for GFI1 and IRF8 motifs, which suggests that mutant GFI1(R412X) is unable to silence them. Notably, zinc-finger 6 is dispensable for DNA binding in vitro<sup>29</sup>, which we confirmed for GFI1(R412X) (Extended Data Fig. 8a–e). Moreover, *Gfi1*<sup>R412X/R412X</sup> ChIP-seq on sorted GMPs revealed most loci normally bound by GFI1 (Extended Data Fig. 8f). However, although GFI1-target loci demonstrate dynamic chromatin accessibility in wild-type cells (Fig. 3c, left), these patterns are disrupted in *Gfi1*<sup>R412X/R412X</sup> cells, including *Irf8* (Fig. 3c, right, 3d, Extended Data Fig. 8g–l). Thus, occupancy of GFI1(R412X) is not associated with normal dynamic regulation of chromatin accessibility.

We noted GFI1(R412X) deregulation of both *Irf8* expression and GFI1-bound chromatin accessibility (Fig. 3a, c, d, Extended Data Fig. 6f, k), so we limited *Irf8* alleles. Similar to cells from mice with a homozygous mutation (*Gfi1*<sup>R412X/R412X</sup>) in which IRF8-GFP expression was repressed and granulopoietic specification was partially rescued (Fig. 1i, Extended Data Fig. 9a–f), *Gfi1*<sup>R412X−/−</sup>*Irf8*<sup>+/−</sup> mice display a partial rescue of neutrophil production (Fig. 3e, Extended Data Fig. 9a–d). However, neutrophils from *Gfi1*<sup>R412X/R412X</sup> and *Gfi1*<sup>R412X−/−</sup>*Irf8*<sup>+/−</sup> still exhibit abnormal morphologies (Figs. 1c, j, 3e).

## Rescuing neutropenia but not function

To determine the genes underlying genetic rescue, we extended Fluidigm captures to genetically rescued GMP-gate cells (Extended



**Fig. 4 | Rescuing specification does not repair innate immune effector functions that are programmed after commitment.** **a**, Heat map of median-normalized  $\log_2$ -transformed values of ADTs from CITE-seq data of sorted wild-type modified-GMP gate cells. HSCP, haematopoietic stem-cell progenitors; cMoP, common monocyte progenitor; DC, dendritic cell; IG2, bi-potential monocytic-granuloctyic intermediates; MDP, monocyte dendritic cell progenitor; Mono, monocytic; Multi-Lin-2, multi-lineage progenitor 2; MP, monocyte-committed progenitor; proNeu-1 (or proNeu-2), neutrophil progenitor; preNeu-1 (or preNeu-2, preNeu-3), neutrophil precursor; immNeu, immature neutrophil. **b**, Joint UMAP projection of CITE-seq data from the indicated genotypes. Each dot represents one cell, and cluster identity is colour-coded (cellHarmony) to highlight genetic repair ( $Gfi1^{R412X/R412X}$  and  $Gfi1^{R412X-} \text{Ir}f8^{+/-}$ ) (Extended Data Fig. 7b). Eryth, erythrocytic; immNeu, immature neutrophil; Meg, megakaryocytic. **c, d**, Heat map of gene expression that is genetically repaired (c) or not genetically repaired (d) in each cluster

Data Figs. 4b, 9g–l) and CITE-seq of modified-GMP-gate cells to all mice (Fig. 4a–f, Extended Data Figs. 5c, 7b, d). Joint embedding of CITE-seq datasets into a single uniform manifold approximation and projection (UMAP) reveals similar embedding of cell populations in  $Gfi1^{+/+}$ ,  $Gfi1^{-/-}$  and  $Gfi1^{K403R^{-}}$  mice. All GFI1(R412X) mutants display a markedly altered trajectory (Fig. 4b), but genetic rescue increased the fraction of granulocytic-specifying cells (Fig. 4b, Extended Data Figs. 7b, 9j). Similarly, ADT expression was only subtly different between  $Gfi1^{+/+}$ ,  $Gfi1^{-/-}$  and  $Gfi1^{K403R^{-}}$ ; however,  $Gfi1^{-/-}$  and genetic rescues showed modest but consistent ADT perturbation in LY6G, CD55, CD106, CD115 and F4/80 expression (Fig. 4a, Extended Data Fig. 7b, d).

(top), with representative genes indicated (right) ( $\log_2$ -normalized median values). R/R, R412X/R412X; R $^{-/-}$ , R412X $^{-/-}$ . **e, f**, Enriched Gene Ontology terms from genetically repaired genes in c (e) or genes not genetically repaired in d (f). FDR, false discovery rate; GPCR, G-protein-coupled receptor. **g**, Heat map of genes that are differentially expressed in neutrophils from patients with SCN (right), and their dynamic expression across wild-type mouse clusters (left). **h**, Enriched biological processes from genes in g. Black dots denote control versus female; red dots denote changes conserved to control versus male. TF, transcription factor. **i**, Volcano plot displaying SWATH proteomics data. Black dots denote differential proteins between  $Gfi1^{R412X^{-}}$  and  $Gfi1^{-/-}$  neutrophils. Red dots denote change conserved in  $Gfi1^{R412X^{-}}$ ,  $Gfi1^{K403R^{-}}$  and  $Gfi1^{R412X/R412X}$  neutrophils. **j, k**, Oxidative burst in purified neutrophils from mouse bone marrow (j) or human peripheral blood (k). Data in j and k are mean  $\pm$  s.e.m. from independent biological replicates. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.0001$ , two-tailed t-test for j and k.

Using cluster-specific comparisons between groups in both datasets, we identified genes that are transcriptionally repaired by both genetic rescues (Fig. 4c, Extended Data Fig. 9k) and those that are not repaired by either genetic rescue (Fig. 4d, Extended Data Fig. 9l). Notably, genes enriched in processes that control neutrophil chemotaxis, NADPH oxidation, defence response, translation, and chromosome condensation were not repaired (Fig. 4f, Extended Data Fig. 9l). Indeed, GFI1(R412X) mutant neutrophils did not condense their chromatin (Figs. 1c, j, 3e).

Next, we captured neutrophils from patients with SCN and age-matched donor controls (Fluidigm C1v4) (Extended Data Fig. 10, Supplementary Information). Differentially expressed genes in SCN are dynamically expressed during normal mouse granulopoiesis (Fig. 4g).

and enriched for many of the same GFI1–ZnF6-dependent biological processes (for example, translation, neutrophil chemotaxis and defence response) (Fig. 4h).

To validate defects in translation, we performed SWATH<sup>30</sup> proteomic analysis on purified bone marrow neutrophils. Dissimilar to their increased gene expression (Extended Data Fig. 6f), we find loss of neutrophil granule protein expression (for example, CAMP, MMP8 and MMP9) in all GFI1-mutant neutrophils (Fig. 4i). In addition, levels of NADPH oxidase complex proteins were low (Fig. 4i), which manifested as a blunted oxidative burst response *in vitro* (Fig. 4j). Patients with SCN display innate immune defects despite G-CSF therapy<sup>31</sup>, and their neutrophils exhibit deregulated genes involved in innate immune functions (Fig. 4h), which suggests that their cytokine-rescued neutrophils may be functionally defective. Similar to GFI1-mutant mouse neutrophils (which are also produced under conditions of high G-CSF) (Extended Data Fig. 3b), we found that peripheral blood neutrophils from patients with SCN exhibited a blunted oxidative burst response (Fig. 4k). Our findings provide evidence that genetic or cytokine rescue of granulocyte specification is not sufficient to rescue post-commitment innate effector programs.

## Discussion

Delineating disease-associated genetic variation is the focus of intensive research; however, determining which sequence alterations are pathogenic requires the generation and analysis of genetic models. Exploiting models to resolve the pathobiological effect of mutations is limited by both technology and our understanding of normal biology. Here we describe a transferable workflow that can be harnessed to answer developmental questions across disciplines. Our results suggest that such analyses can reveal cell-state-specific effects of mutations (probably due to accompanying changes in the composition of transcription factors and chromatin accessibility at their target genes), with direct consequences for attempts to repair defects. This has important implications for the study of congenital and acquired genetic changes, especially cancers that are multiclonal. Moreover, this approach could be extended to the analysis of the therapeutic effect of new small molecules, in which bulk cell analyses or current single-cell analytic pipelines might gloss over rare but important cell states.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2227-7>.

1. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
2. Gilbert, S. F. *Developmental Biology*, 6th edition (Sinauer Associates, 2000).
3. DePasquale, E. A. K. et al. cellHarmony: cell-level matching and holistic comparison of single-cell transcriptomes. *Nucleic Acids Res.* **47**, e138 (2019).
4. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e1821 (2019).

5. Xia, J. et al. Prevalence of mutations in *ELANE*, *GFI1*, *HAX1*, *SBDS*, *WAS* and *G6PC3* in patients with severe congenital neutropenia. *Br. J. Haematol.* **147**, 535–542 (2009).
6. Person, R. E. et al. Mutations in proto-oncogene *GFI1* cause human neutropenia and target *ELA2*. *Nat. Genet.* **34**, 308–312 (2003).
7. Wang, H. et al. A reporter mouse reveals lineage-specific and heterogeneous expression of IRF8 during lymphoid and myeloid cell differentiation. *J. Immunol.* **193**, 1766–1777 (2014).
8. Deshmukh, H. S. et al. The microbiota regulates neutrophil homeostasis and host resistance to *Escherichia coli* K1 sepsis in neonatal mice. *Nat. Med.* **20**, 524–530 (2014).
9. Luedi, P. P. et al. Computational and experimental identification of novel human imprinted genes. *Genome Res.* **17**, 1723–1730 (2007).
10. Bonthuis, P. J. et al. Noncanonical genomic imprinting effects in offspring. *Cell Rep.* **12**, 979–991 (2015).
11. Hock, H. et al. Intrinsic requirement for zinc finger transcription factor Gfi-1 in neutrophil differentiation. *Immunity* **18**, 109–120 (2003).
12. Belyaev, N. N. et al. Induction of an IL7-R $\gamma$ c-Kit $^{hi}$  myelolymphoid progenitor critically dependent on IFN-gamma signaling during acute malaria. *Nat. Immunol.* **11**, 477–485 (2010).
13. MacNamara, K. C. et al. Infection-induced myelopoiesis during intracellular bacterial infection is critically dependent upon IFN- $\gamma$  signaling. *J. Immunol.* **186**, 1032–1043 (2011).
14. Mempel, K., Pietsch, T., Menzel, T., Zeidler, C. & Welte, K. Increased serum levels of granulocyte colony-stimulating factor in patients with severe congenital neutropenia. *Blood* **77**, 1919–1922 (1991).
15. Hock, H. et al. Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. *Nature* **431**, 1002–1007 (2004).
16. Karsunky, H. et al. Inflammatory reactions and severe neutropenia in mice lacking the transcriptional repressor Gfi1. *Nat. Genet.* **30**, 295–300 (2002).
17. Zeng, H., Yücel, R., Kosan, C., Klein-Hitpass, L. & Möröy, T. Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. *EMBO J.* **23**, 4116–4125 (2004).
18. Abe, T. et al. Visualization of cell cycle in mouse embryos with Fucci2 reporter directed by Rosa26 promoter. *Development* **140**, 237–246 (2013).
19. Olsson, A. et al. Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature* **537**, 698–702 (2016).
20. Yáñez, A. et al. Granulocyte-monocyte progenitors and monocyte-dendritic cell progenitors independently produce functionally distinct monocytes. *Immunity* **47**, 890–902.e4 (2017).
21. Ng, L. G., Ostuni, R. & Hidalgo, A. Heterogeneity of neutrophils. *Nat. Rev. Immunol.* **19**, 255–265 (2019).
22. Farrell, J. A. et al. Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* **360**, eaar3131 (2018).
23. Zhu, Y. P. et al. Identification of an early unipotent neutrophil progenitor with pro-tumoral activity in mouse and human bone marrow. *Cell Rep.* **24**, 2329–2341.e8 (2018).
24. Drissen, R. et al. Distinct myeloid progenitor-differentiation pathways identified through single-cell RNA sequencing. *Nat. Immunol.* **17**, 666–676 (2016).
25. Kim, M. H. et al. A late-lineage murine neutrophil precursor population exhibits dynamic changes during demand-adapted granulopoiesis. *Sci. Rep.* **7**, 39804 (2017).
26. Evrard, M. et al. Developmental analysis of bone marrow neutrophils reveals populations specialized in expansion, trafficking, and effector functions. *Immunity* **48**, 364–379.e8 (2018).
27. Venkatasubramanian, M., Chetal, K., Schnell, D., Atluri, G. & Salomonis, N. Resolving single-cell heterogeneity from hundreds of thousands of cells through sequential hybrid clustering and NMF. *Bioinformatics* **btaa201** (2020).
28. Fang, R. et al. Fast and accurate clustering of single cell epigenomes reveals *cis*-regulatory elements in rare cell types. Preprint at <https://www.bioRxiv.org/content/10.1101/615179v2> (2019).
29. Zweidler-McKay, P. A., Grimes, H. L., Flubacher, M. M. & Tsichlis, P. N. Gfi-1 encodes a nuclear zinc finger protein that binds DNA and functions as a transcriptional repressor. *Mol. Cell. Biol.* **16**, 4024–4034 (1996).
30. Gillet, L. C. et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell Proteomics* **11**, O111.016717 (2012).
31. Elsner, J., Roesler, J., Emmendorffer, A., Lohmann-Matthes, M. L. & Welte, K. Abnormal regulation in the signal transduction in neutrophils from patients with severe congenital neutropenia: relation of impaired mobilization of cytosolic free calcium to altered chemotaxis, superoxide anion generation and F-actin content. *Exp. Hematol.* **21**, 38–46 (1993).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were blinded to allocation for *C. albicans* and *S. aureus* experiments.

### Human samples

De-identified clinical data and peripheral blood from healthy donors or neutropenic patients was obtained at Cincinnati Children's Hospital Medical Center through informed consent under an approved Institutional Review Board research protocol.

### Mice

*Irf8<sup>eGFP</sup>* mice<sup>7</sup>, *Gfi1<sup>dex23/dex23</sup>* mice<sup>11</sup>, *Rosa26<sup>Fucci2</sup>* mice<sup>18</sup> and *Irf8<sup>gml.2HmJ</sup>* mice<sup>32</sup> were maintained on a C57BL/6 background, and *Ptprc<sup>a</sup>* (Boyl) mice were purchased from Jackson Laboratories. Mice were maintained under specific pathogen-free conditions by Cincinnati Children's Hospital Medical Center (CCHMC) Veterinary Services and all procedures were approved by the Children's Hospital Research Foundation Institutional Animal Care and Use Committee (Protocol Number IACUC2017-0021).

To generate *Gfi1<sup>N382S</sup>*, *Gfi1<sup>K403R</sup>* and *Gfi1<sup>R412X</sup>* mice, two pairs of ZFN RNAs that recognize target sites flanking *Gfi1* coding exon 6 were designed and generated by Millipore-Sigma (Genome Editing division). mRNAs encoding site-specific ZFNs (50 ng  $\mu\text{l}^{-1}$ ) were introduced into one-cell mouse oocytes by standard pronuclear injection approach, and injected oocytes were transferred to a pseudo pregnant surrogate mother. To introduce specific mutations, oocytes were co-injected with ZFN mRNAs as well as a homologous donor construct (2 ng  $\text{ml}^{-1}$ ) containing the R412X mutation. Positive founder pups were identified using mutation-specific primers and mated to C57BL/6 mice to generate stable heritable knock-in lines. Constructs consisted of 1.5-kb 5' and 0.8-kb 3' arms of homology flanking the mutant *Gfi1* exon 6. The ZFN target sequences used were: 5'-GTCCTTACCTCCCTTccggaaGCTGCTGGAGGAGATGAA-3' and 5'-CGTTGCCACCATGCTTgtctaaCAGCTGGCTAAG-3', in which the capital letters denote nucleotides actually bound by right and left ZFN proteins. Positive founders and subsequent progeny were genotyped by PCR using the primers: 5'-CAGAAAGCACACAGGCTTCA-3' and 5'-GATGAGCTTGACACTGG-3' then subsequent restriction enzyme digestion using BamHI for *Gfi1<sup>N382S</sup>*, KpnI for *Gfi1<sup>K403R</sup>* and EcoRI for *Gfi1<sup>R412X</sup>*. Expression of the mutant alleles was confirmed by TOPO cloning/TOP10 transformation (Thermo Fisher) of RT-PCR products generated using the same primers that were used for genotyping, followed by Sanger sequencing of individual clones with T3 and T7 primers and analysis using ApE (v.2.0.49.0). The *Gfi1<sup>R412X</sup>* mouse line is available from The Jackson Laboratory as JAX stock 035156.

### Flow cytometry and cell sorting

Mice were euthanized using carbon dioxide followed by cervical dislocation. Peripheral blood was collected into EDTA coated tubes, and hind limb bones (femurs, tibias and the iliac crest) were obtained immediately after euthanasia and stored in cold FACS buffer (1% FBS, 0.01%  $\text{NaN}_3$  in DPBS) under sterile conditions. Bones were flushed using a syringe for transplantation or crushed using a mortar and pestle for all other applications, then passed through a 40- $\mu\text{m}$  cell strainer (BD) to obtain single-cell suspensions for downstream applications. Before analytical flow cytometry, erythrocytes were lysed using ACK buffer (Gibco), and cells were washed in FACS buffer and resuspended in FACS buffer containing DAPI (Thermo Fisher) or 7-AAD (BD). Flow cytometry analyses were conducted on a FACS LSRII or FACS LSRFortessa (BD) and analysed using FlowJo software.

To enrich for mouse stem/progenitor populations, freshly isolated bone marrow cells were incubated with CD117 Microbeads or with components of the Mouse Lineage Depletion kit and separated on an AutoMACS Pro separator (Miltenyi) according to manufacturer's

specifications. Alternatively, mouse LSK cells were purified by first using the SCA-1 Multisort Kit (Miltenyi) for separation on an AutoMACS Pro separator then cells were subsequently incubated with CD117 microbeads and separated again on an AutoMACS Pro separator. Mouse bone marrow neutrophils were isolated using the EasySep Mouse Neutrophil Enrichment Kit (Stem Cell Technologies) according to manufacturer's specifications. Human peripheral blood neutrophils were enriched from whole blood using the EasySep Direct Human Neutrophil Isolation Kit (StemCell Technologies) according to manufacturer's specifications.

### Peripheral blood analysis

Peripheral blood cells were stained in FACS buffer with a mix of antibodies: CD16/CD32 (clone 2.4G2, BD), CD3-APC (clone 145-2C11, BD), CD45R-PE (clone RA3-6B2, BD), CD11b-PacificBlue (clone M1/70, BioLegend), and Ly6g-PerCP-Cy5.5 (clone 1A8, BioLegend). For analysis of transplant chimaeras, cells were also stained with CD45.1-BV605 (clone A20, Biolegend) and CD45.2-AlexaFluor700 (clone 104, Biolegend). To quantify neonatal mouse blood numbers, samples were supplemented with 5.24- $\mu\text{m}$  AccuCount blank particles (Spherotech) immediately before analysis. To purify peripheral blood neutrophils for cytopsin analysis, cell sorting was performed on a FACSaria II (BD) and cells were collected in a solution of DPBS plus 50% FBS (Atlanta Biologicals).

### Mature myeloid bone marrow analysis

Whole bone marrow cells were stained in FACS buffer with a mix of antibodies CD16/CD32 (clone 2.4G2, BD), Ly6B.2-AlexaFluor647 (clone 7/4, Bio-Rad), and F4-80-AlexaFluor488 (clone Cl:A3-1, Bio-Rad) for 30 min on ice.

### Haematopoietic stem-cell analysis

Whole bone marrow cells were stained for 30 min on ice with a cocktail of antibodies: Fc Block (BD), biotin-conjugated anti-CD3e (clone 145-2C11, BioLegend), biotin-conjugated anti-CD4 (clone RM4-5, Thermo Fisher), biotin-conjugated anti-CD8 (clone 53-6.7), biotin-conjugated anti-CD11b (clone M1/70, BD), biotin-conjugated anti-CD19 (clone 6.D5, BioLegend), biotin-conjugated anti-CD127 (clone B12-1, BD), biotin-conjugated anti-B220 (clone RA3-6B2, BioLegend), biotin-conjugated anti-Gr1 (clone RB6-8C5, BD), and biotin-conjugated anti-Ter119 (clone TER-119, BioLegend). Cells were then washed twice with FACS buffer and incubated on ice overnight with a cocktail of antibodies: streptavidin-APC-Cy7 (BD), APC-conjugated anti-CD117 (clone 2B8, BioLegend), PE-Cy7-conjugated anti-SCA-1 (clone D7, BD), FITC-conjugated anti-CD34 (clone RAM34, BD), PacificBlue-conjugated anti-CD48 (clone HM48-1, BioLegend), PE-conjugated anti-CD135 (clone A2F10, BioLegend) and Brilliant Violet 510-conjugated anti-CD150 (clone TC15-12F12.2, BioLegend).

### Myeloid progenitor analysis

Before FACS sorting, Miltenyi AutoMacs CD117-enriched bone marrow cells were stained for 1 h at room temperature protected from light with a mix of antibodies: PerCP-eFluor710-conjugated anti-CD16/CD32 (clone 93, Thermo Fisher), biotin-conjugated anti-CD3e (clone 145-2C11, BioLegend), biotin-conjugated anti-CD4 (clone RM4-5, Thermo Fisher), biotin-conjugated anti-CD8 (clone 53-6.7), biotin-conjugated anti-CD11b (clone M1/70, BD), biotin-conjugated anti-CD19 (clone 6.D5, BioLegend), biotin-conjugated anti-CD127 (clone B12-1, BD), biotin-conjugated anti-B220 (clone RA3-6B2, BioLegend), biotin-conjugated anti-Gr1 (clone RB6-8C5, BD), biotin-conjugated anti-Ter119 (clone TER-119, BioLegend), Brilliant Violet 421-conjugated anti-CD34 (clone RAM34, BD), APC-conjugated anti-CD117 (clone 2B8, BioLegend), and PE-Cy7-conjugated anti-SCA-1 (clone D7, BD). Cells were then washed twice with FACS buffer and incubated for 15 min at room temperature with streptavidin-APC-Cy7 (BD). Cells were washed once in FACS buffer, and resuspended in FACS buffer for cell sorting.

# Article

Cell sorting was performed on a FACS Aria II (BD) and cells were collected in a solution of DPBS plus 50% FBS (Atlanta Biologicals).

## GMP-P gate

Before FACS sorting, CD117-enriched bone marrow cells were stained for 1 h at room temperature protected from light with a mix of antibodies: PerCP-eFluor710-conjugated anti-CD16/CD32 (clone 93, Thermo Fisher), biotin-conjugated anti-CD3e (clone 145-2C11, BioLegend), biotin-conjugated anti-CD4 (clone RM4-5, Thermo Fisher), biotin-conjugated anti-CD8 (clone 53-6.7), biotin-conjugated anti-CD19 (clone 6.D5, BioLegend), biotin-conjugated anti-CD127 (clone B12-1, BD), biotin-conjugated anti-B220 (clone RA3-6B2, BioLegend), biotin-conjugated anti-Ter119 (clone TER-119, BioLegend), AlexaFluor700-conjugated anti-CD11b (clone M1/70, BioLegend), Brilliant Violet 785-conjugated anti-LY6C (clone HK1.4, BioLegend), FITC-conjugated anti-LY6G (clone 1A8, BioLegend), Brilliant Violet 421-conjugated anti-CD34 (clone RAM34, BD), APC-conjugated anti-CD117 (clone 2B8, BioLegend), and Brilliant Violet 605-conjugated anti-CD115 (clone T38-320, BD). Cells were then washed twice with FACS buffer and incubated for 15 min at room temperature with streptavidin-APC-Cy7 (BD). Cells were washed once in FACS buffer, and resuspended in FACS buffer for cell sorting. Cell sorting was performed on a FACS Aria II (BD) and cells were collected in a solution of DPBS plus 50% FBS (Atlanta Biologicals).

## Alternative GMP-P gate (incorporating additional markers)

Alternatively, before analysis, CD117-enriched bone marrow cells were stained for 1 h at room temperature protected from light with a mix of antibodies: Brilliant UV 395-conjugated anti-CD16/CD32 (clone 93, BD), biotin-conjugated anti-CD3e (clone 145-2C11, BioLegend), biotin-conjugated anti-CD4 (clone RM4-5, Thermo Fisher), biotin-conjugated anti-CD8 (clone 53-6.7), biotin-conjugated anti-CD19 (clone 6.D5, BioLegend), biotin-conjugated anti-CD127 (clone B12-1, BD), biotin-conjugated anti-B220 (clone RA3-6B2, BioLegend), biotin-conjugated anti-Ter119 (clone TER-119, BioLegend), PE-Cy7-conjugated anti-CD11b (clone M1/70, BioLegend), Brilliant Violet 785-conjugated anti-LY6C (clone HK1.4, BioLegend), PerCP-Cy5.5-conjugated anti-LY6G (clone 1A8, BioLegend), Brilliant Violet 421-conjugated anti-CD34 (clone RAM34, BD), Brilliant Violet 650-conjugated anti-CD117 (clone 2B8, BioLegend), and Brilliant Violet 605-conjugated anti-CD115 (clone T38-320, BD). Further analyses were performed by also individually staining cells with the following antibodies: PE-conjugated anti-CD18 (clone M18/2, BioLegend), PE-conjugated anti-CD27 (clone LG.3A10, BioLegend), PE-conjugated anti-CD55 (clone RIKO-3, BioLegend), PE-conjugated anti-CD84 (clone mCD84.7, BioLegend), AlexaFluor 700-conjugated anti-CD177 (clone 1171A, R&D Systems), PE-conjugated anti-CD106 (Vcam1, clone 429, BioLegend), PE-conjugated anti-CD217 (IL-17R, clone PAJ-17R, Thermo Fisher). Cells were then washed twice with FACS buffer and incubated for 15 min at room temperature with streptavidin-APC-Cy7 (BD).

## Cell culture and viral transduction

Lentiviral particles were generated as described previously<sup>33</sup>. In brief, Lenti-X 293T cells (Clontech) were seeded one day before transfection with individual expression vectors (pLVX-EF1α-IRES-Puro base vector, Clontech), the lentiviral packaging plasmid Δ8.9, and envelope plasmid VSV-G using Transit-LT1 (Mirus) according to manufacturer instructions. Growth medium was changed the next day. Virus-containing supernatant was collected 48 and 72 h after transfection, pooled and concentrated via ultracentrifugation over a 20% sucrose cushion, then aliquoted and frozen at -80 °C.

Freshly isolated mouse LSK bone marrow cells were cultured in IMDM (ThermoFisher Scientific) supplemented with 100 U ml<sup>-1</sup> penicillin-streptomycin (ThermoFisher Scientific), 10% BIT (StemCell Technologies), 20 ng ml<sup>-1</sup> IL-3, 20 ng ml<sup>-1</sup> IL-6, and 50 ng ml<sup>-1</sup> SCF (Miltenyi) with lentiviral particles overnight at 37 °C, 5% CO<sub>2</sub>. Medium was changed and

supplemented with 1.5 µg ml<sup>-1</sup> puromycin (Invivogen) 48 h after transduction, and the IRF8-GFP levels were determined by FACS 72 h later.

## RNAscope imaging

Human CD34<sup>+</sup> cells were isolated from either G-CSF-mobilized peripheral blood or fresh bone marrow aspirate using ClinIMACS CD34+ (Miltenyi, 130-017-501) or FACS sorting. Cells were fixed with 4% neutral buffered formalin at 37 °C for 1 h and cyto-centrifuged onto SuperFrost Plus Slides (Fisher Scientific, 12-550-15). Samples were then processed according to the RNAscope V2 User Manual (ACDBio). Channel 1 of the 3-plex Negative Control Probe (ACDBio, 320871), probing for the bacterial gene *dabB*, was developed to determine background fluorescence. Negligible fluorescent signal was generated by this probe in all samples. Channel 1 of experimental probes for *POLR2A* intron 1 (ACDBio, 588751), *GFI1* intron 3 (ACDBio, 575371), and *GFI1* intron 6 (ACDBio, 588761) were also developed. Samples were stained using the Opal 570 reagent (Perkin Elmer, FP1488001KT) at a dilution of 1:1,500. Slides were imaged with a Nikon A1 confocal on an Eclipse TiE inverted microscope with a 60× Plan Apo IR DIC-Water Immersion objective. Z-stacks were acquired with a 0.11-µm pixel size and 0.25-µm Z interval. Z-stacked data were analysed using Imaris 9 with transcripts per cell defined using 'spot' detection. Transcripts per cell were quantified using Imaris XT MatLab algorithms. Cells with no transcripts were excluded from final analysis. Images shown are maximum intensity projections of Z-stacks.

## GFI1(R412X) iPS cell generation and differentiation

iPS cell lines were generated as described previously<sup>5</sup>. CRISPR-Cas9 was used to introduce the R412X mutation to the *GFI1* locus of an iPS cell line (iPSC286C11) derived from a healthy individual. A guide RNA targeting the last intron of *GFI1* (5'-GTGACTCCGTTCTAATTAG-3') was designed according to the web tool (<http://CRISPOR.org>) and cloned into a modified pX458 vector (Addgene 48138) that expresses an optimized sgRNA scaffold and a high-fidelity eSpCas9(1.1)-2A-GFP. The editing activity of the plasmid was validated in HEK293T cells by T7E1 assay. The donor plasmid was constructed by inserting the DNA fragments containing desired mutations and homologous arms to a pUC57 vector carrying a *loxP*-flanked EF1a-GFP-2A-Puro cassette<sup>34</sup>. For gene editing, a single-cell suspension of iPSC286C11 cells was prepared using accutase and 0.75 × 10<sup>6</sup> cells were reverse-transfected in a 12-well Matrigel-coated dish with 0.6 µg of the gRNA-containing pX458 vector, 0.6 µg of donor plasmid, and 0.18 µg of pRetroSuper-p53sh (a gift from J. Mulloy) using the manufacturer's recommended protocol (TransIT-LT1; Mirus). One day after transfection, mTeSR1 containing 1 µg ml<sup>-1</sup> puromycin was added to cells for 48 h to eliminate non-transfected cells. Five days after transfection, a single-cell suspension of surviving cells was prepared using accutase. Cells were replated at 0.25 × 10<sup>6</sup> per well in a matrigel-coated 6-well dish and cultured for 4 days in mTeSR1 containing 10% CloneR (StemCell Technologies), using the manufacturer's recommended protocol. Cells were subsequently fed daily with mTeSR1 containing 1 µg ml<sup>-1</sup> puromycin for a further 1–2 weeks before colonies with stereotypical iPS cell morphology were manually excised and expanded for genotyping. Correctly targeted clones were identified by PCR, enzyme digestion and Sanger sequencing and confirmed to have normal karyotype. To remove the *loxP*-flanked EF1a-GFP-2a-Puro cassette, cells were incubated with TAT-Cre recombinase according to manufacturer's recommendations (Millipore Sigma). GFP<sup>+</sup> cells were subsequently purified by FACS sorting.

Granulocytes were generated from iPS cells using a three-step differentiation process. First, CD34<sup>+</sup>CD45<sup>+</sup> haematopoietic progenitors were derived using the STEMdiff haematopoietic kit (Stem Cell Technologies). Floating CD34<sup>+</sup>CD45<sup>+</sup> cells were collected and cultured in RPMI 1640 medium (Gibco) containing 10% FBS (Hyclone), 50 ng ml<sup>-1</sup> SCF, 10 ng ml<sup>-1</sup> IL-3 and 10 ng ml<sup>-1</sup> GM-CSF (Peprotech) and cultured for 5 days. After 5 days, culture medium was transitioned to RPMI 1640 medium supplemented with 10% FBS and 50 ng ml<sup>-1</sup> G-CSF (Peprotech) and cultured for another 5 days. Cells were stained with PE-conjugated

anti-CD45 (clone HI30, BD), APC-conjugated anti-CD34 (clone 581, BD), APC-Cy7-conjugated anti-CD11B (clone ICRF44, BD), Viobline-conjugated anti-CD14 (clone REA599, Miltenyi), and PE-Cy7-conjugated anti-CD16 (clone REA423, Miltenyi) before FACS analysis. To analyse *Gfi1* expression, cells were lysed in TriZol (Thermo Fisher) and total RNA was reverse transcribed using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Amplification of cDNA was performed using GoTaq Green Master Mix (Promega) using 5'-CAGGAACGGAGCTTGACTG-3' and 5'-GAAGGAGGAGCAACCTGGTA-3' primers before sequencing using 5'-CAAGAGGTCACTCACACTGTC-3' or 5'-TCTGGAAAGTCAGAAC GGAGT-3' primers.

### Immunoblot analysis

Freshly isolated Lin<sup>neg</sup> bone marrow cells were lysed in Laemmli buffer, briefly sonicated, and boiled at 100 °C for 5 min before being subjected to SDS-PAGE. The resolved proteins were transferred onto nitrocellulose membranes (GE Healthcare) that were subsequently blocked in 5% nonfat dried milk (Carnation) in TBS-T (0.1% Tween 20, Sigma Aldrich) for 1 h at room temperature with agitation. After washes with TBS-T, membranes were incubated overnight at 4 °C with agitation in a solution of 5% BSA Fraction V/TBS-T (Fisher Scientific) containing goat anti-Gfi1 (R&D Systems) antibody diluted at 1:1,000. After a wash with TBS-T, membranes were incubated for 2 h at room temperature with agitation in a solution of 2% non-fat dried milk in TBS-T containing a 1:5,000 dilution of HRP-conjugated anti-goat secondary antibody (Invitrogen). Following three washes with TBS-T, chemiluminescent detection of blotted proteins was performed using ECL (Thermo Fisher) and detected using X-ray film (Lab Scientific). Next, membranes were washed thoroughly in TBS-T, then stripped in Restore Plus (Thermo Fisher) for 10 min at room temperature with agitation. After washing in TBS-T, membranes were blocked as before, and incubated as described above with a 1:5,000 dilution of mouse anti-β-actin (Sigma Aldrich) and then a 1:10,000 dilution of HRP-conjugated anti-mouse secondary antibody (GE Healthcare) before chemiluminescent detection was achieved as before.

### Cytospin

Freshly prepared cells were diluted in 200 µl of FACS buffer and spun onto VWR VistaVision Histobond slides (VWR) for 3 min at 900 rpm using a Cytospin 4 apparatus (Thermo Scientific). Slides were then fixed and stained using the Camco stain pak (Cambridge Diagnostic Products), and once dry, sealed using Cytoseal 60 (Thermo Scientific) and microscope cover glass (Globe Scientific).

### CFU unit assays

Freshly isolated Lin<sup>neg</sup> or KIT<sup>+</sup> cells were plated at 1,000 cells ml<sup>-1</sup> in M3434 (StemCell Technologies) and the differential colony number was determined 7 days later.

### Bone marrow transplantation

A single-cell suspension of 1 × 10<sup>6</sup> unfractionated bone marrow cells from 6-week-old mice (CD45.2) was injected into lethally irradiated (11.75 Gy) CD45.1 congenic recipients.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

### Data availability

scRNA-seq, CITE-seq, scATAC-seq and ChIP-seq data are deposited as Gene Expression Omnibus (GEO) SuperSeries GSE120409 and in Synapse (<https://www.synapse.org/#!Synapse:syn16806696>). Proteomic data are available via ProteomeXchange ([www.ebi.ac.uk/pride/archive/](http://www.ebi.ac.uk/pride/archive/)) with identifier PXD010943. Extensive processed and primary data are provided and organized in Synapse, including 10x Genomics count matrices, Fluidigm expression, single-cell populations, quality control metrics, 10x Genomics Cell Ranger outputs (summary report and loupe browser files), genomic coordinate peak files and differential expression results (cellHarmony).

### Code availability

All described code is provided in Github (<https://github.com/nsalomonis/altanalyze>).

32. Holtschke, T. et al. Immunodeficiency and chronic myelogenous leukemia-like syndrome in mice with a targeted mutation of the ICSBP gene. *Cell* **87**, 307–317 (1996).
33. Meyer, S. E. et al. DNMT3A haploinsufficiency transforms FLT3ITD myeloproliferative disease into a rapid, spontaneous, and fully penetrant acute myeloid leukemia. *Cancer Discov.* **6**, 501–515 (2016).
34. Nayak, R. C. et al. Pathogenesis of ELANE-mutant severe neutropenia revealed by induced pluripotent stem cells. *J. Clin. Invest.* **125**, 3103–3116 (2015).

**Acknowledgements** We thank H. Singh for Gfi1 antisera, H. C. Morse for supplying *Irf8*-GFP mice, members of the Cincinnati Children's Hospital Medical Center (CCHMC) DNA sequencing and Genotyping Core and the Research Flow Cytometry Core (supported in part by National Institutes of Health (NIH) grants AR-47363, NIH DK78392 and NIH DK90971), and A. Sproles for assistance with cytokine analyses; S. Smith, H. Chi Liang and K. Rangel in the CCHMC Gene Expression Core for generating scRNA-seq libraries; T. Newkold for technical assistance; and H. Singh, J. Cancelas, S. Kogan, R. Kopan, J. Wells and T. Gruber for questions and discussions. This work was partly funded by support from NIH DP1AI131080 (S.S.W.) T32 ES007250 (to D.E.M.), S10RR027015 (to K.D.G.), and R01HL122661 (H.L.G.).

**Author contributions** D.E.M., G.P., S.S.W. and H.L.G. designed experiments. D.E.M., A.O., G.P., R.S., S.C. and P.D. performed experiments and also analysed data with K.G. K.Z., J.Y. and K.C.M. provided patient samples and data. L.R.T. and C.L. generated and cultured iPSC cells. J.M. and D.J.K. generated the *Gfi1*<sup>N392S</sup>, *Gfi1*<sup>K403R</sup> and *Gfi1*<sup>R412X</sup> mice. N.S. conceived and developed the cellHarmony software with considerable input from H.L.G. S.H. and N.S. developed the gene viewer. K.F., B.S., K.C., K.L.N., N.S. and H.L.G. analysed bioinformatics data. D.E.M., N.S. and H.L.G. wrote the paper.

**Competing interests** D.E.M. is currently employed by Eli Lilly and Company and K.L.N. is employed by BioLegend Inc. The remaining authors declare no competing interests.

### Additional information

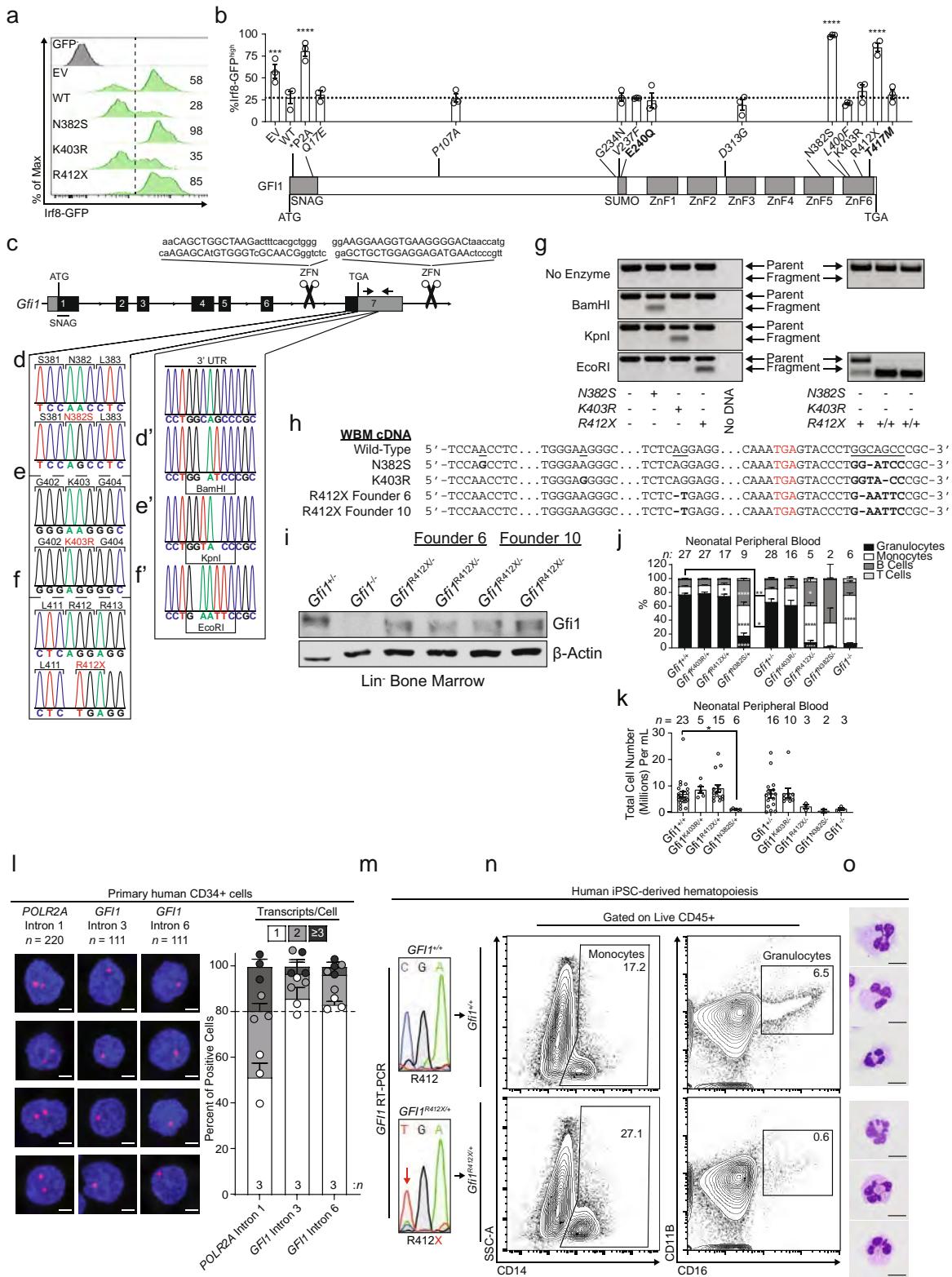
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2227-7>.

**Correspondence and requests for materials** should be addressed to N.S. or H.L.G.

**Peer review information** *Nature* thanks Klaus Ley, Cornelis Murre and Leonard Zon for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

# Article



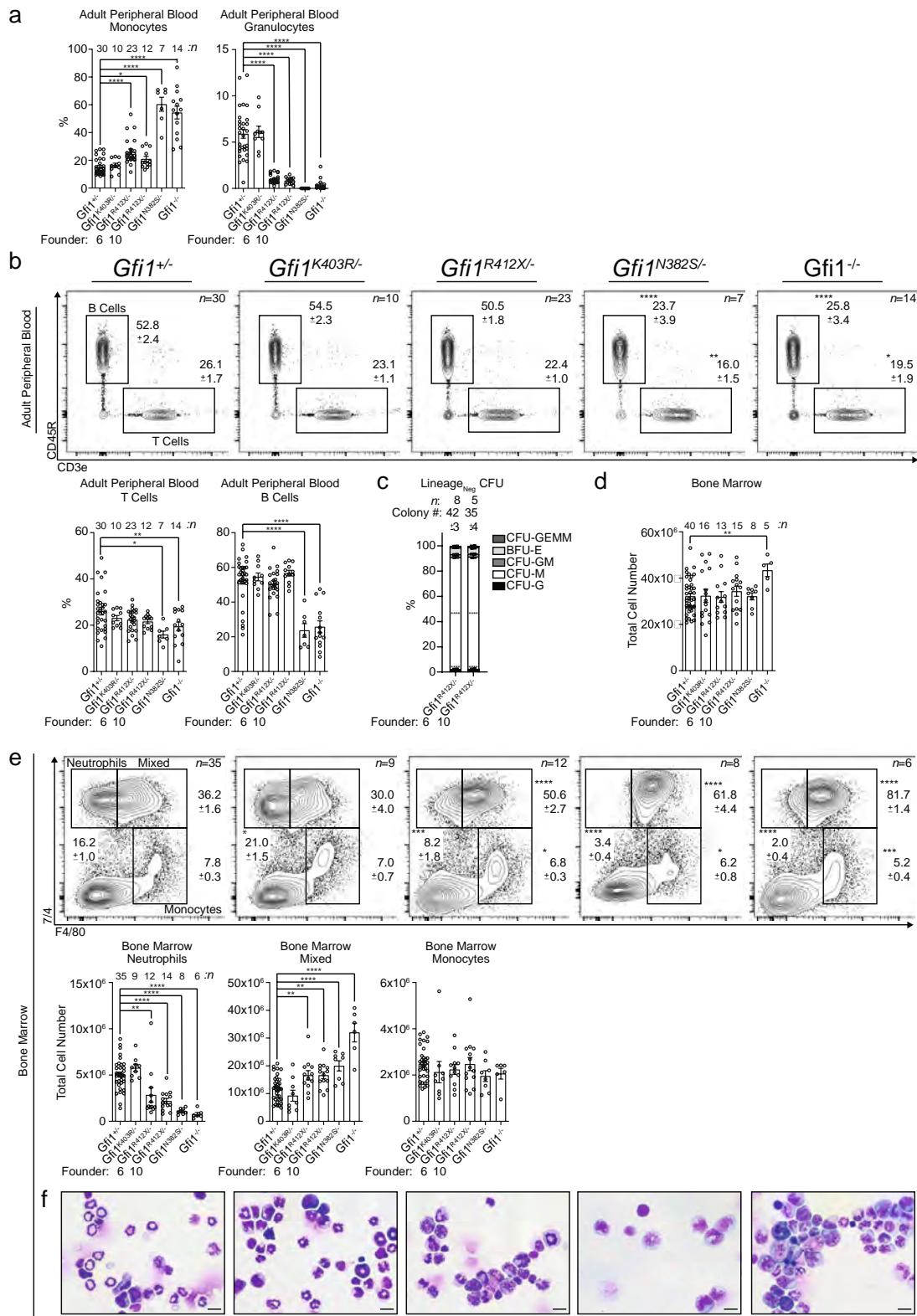
**Extended Data Fig. 1** | See next page for caption.

**Extended Data Fig. 1 | Functional assessment of GFI1 variants derived from patients with SCN and generation of GFI1 ZnF-mutant mice.**

**a**, Representative FACS plots of lentiviral-transduced LSK cells isolated from adult *Irf8-eGFP* transgenic mice with the percentage of IRF8-eGFP<sup>high</sup> cells indicated. **b**, Graphical summary of FACS analysis of lentiviral-transduced LSK cells isolated from adult *Irf8-eGFP* transgenic mice (top) with locations of the variants mapped to the GFI1 protein (bottom). ATG, start codon; EV, empty vector; TGA, stop codon; \*P2A, mutation not found in patients; SNAG; SNAIL/GFI1 family domain; SUMO, sumoylation domain. Italic font denotes other variants detected in the same patient; bold font denotes variants also found in patients diagnosed with a malignancy; grey blocks denote characterized protein domains. **c**, Schematic of the *Gfi1* locus annotated with relevant features. Line with small arrows denote intronic regions; numbered blocks denote exons; black blocks denote coding regions; grey blocks denote noncoding regions; large arrows denote noncoding region used for genotyping, ZFN, zinc-finger nucleases. **d–f**, Schematic of the nucleotide changes made to the coding region and 3' untranslated region (UTR) to introduce the N382S (**d**, **d'**), K403R (**e**, **e'**), and R412X (**f**, **f'**) mutations. **g**, Representative genotyping of the ZnF-mutant mice with or without restriction enzyme digestion. **h**, Representation of Sanger sequencing analysis

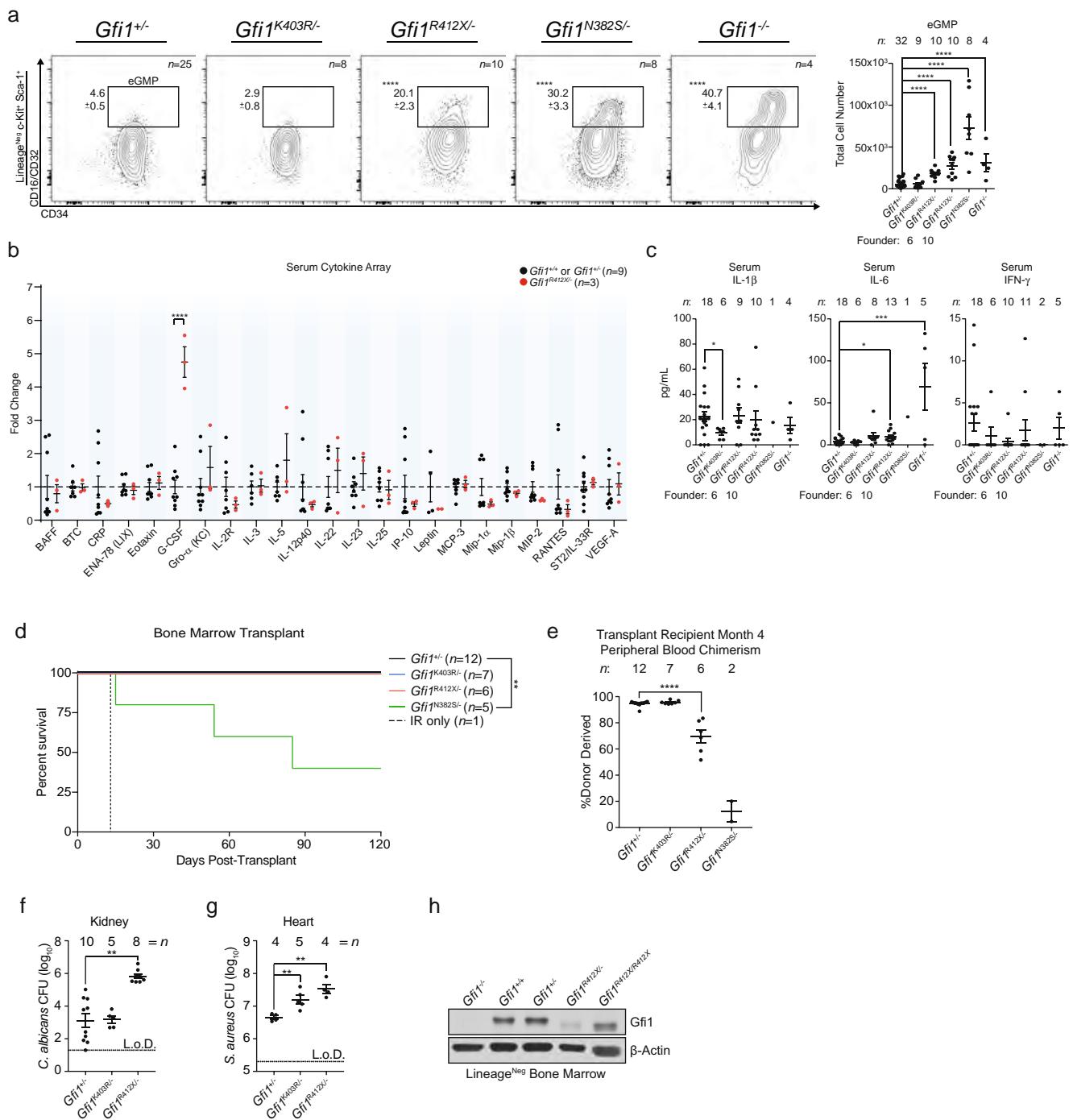
of cDNA from adult whole bone marrow (WBM). Targeted wild-type nucleotides are underlined; mutated nucleotides are in bold; and red nucleotides indicate the location of the stop codon. **i**, Immunoblot analysis of two R412X founder lines using bone marrow lysates from adult mouse Lin<sup>neg</sup> bone marrow. **j**, Graphical summary of FACS analysis of peripheral blood from neonatal mice. **k**, Total cell counts per millilitre of neonatal peripheral blood as determined by FACS. **l**, Representative RNAscope images (left) and transcript quantification (right) of the indicated transcripts in primary human CD34<sup>+</sup> cells. The number of cells scored (left) and the number of donors tested (right) are indicated. **m**, Representative electropherogram plots from Sanger sequencing of *GFI1* RT-PCR products derived from iPS cells. The arrow indicates the nucleotide substitution made to introduce the R412X mutation. **n**, Representative FACS plots of iPS cells at the end of the 10-day haematopoietic differentiation protocol. **o**, Representative cytospin images of neutrophils derived from iPS cells. Data in **a**, **b** and **l** are representative of three biological replicates; data in **j** and **k** are representative of individual biological replicates; data in **i** display two biological replicates for each founder line; data in **m–o** are representative of two independent experiments. Data are mean (**a**) or mean  $\pm$  s.e.m. (**b**, **j–l**). \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ , two-tailed *t*-test. Scale bars, 2  $\mu$ m (**l**) and 10  $\mu$ m (**o**).

$*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ ,  $****P < 0.0001$ , two-tailed *t*-test. Scale bars, 2  $\mu$ m (**l**) and 10  $\mu$ m (**o**).



**Extended Data Fig. 2 | Characterization of GFI1 ZnF-mutant mice at the steady-state.** **a**, Graphical representation of FACS analysis of peripheral blood from adult mice. **b**, FACS plots and graphical representation of FACS analysis of peripheral blood from adult mice. **c**, Graphical summary of CFU assays performed on adult mouse Lin<sup>-</sup> bone marrow from two different founder lines. Total cell counts of whole bone marrow obtained from two femurs and

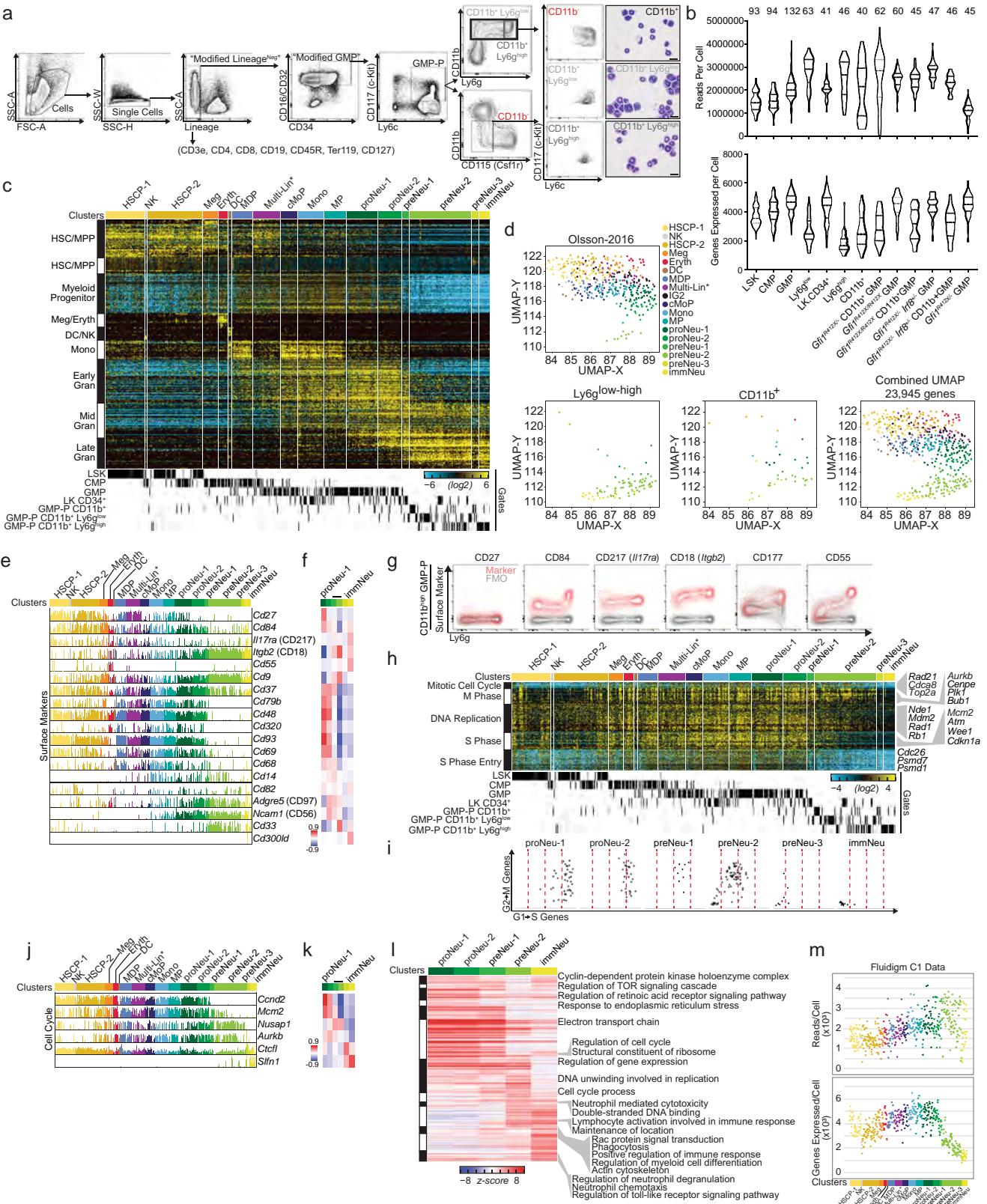
tibias per adult mouse. **e**, FACS plots and quantification of bone marrow populations from adult mice. **f**, Representative cytoplasts of whole bone marrow from adult mice. Data in **f** are representative of three biological replicates. Data are mean  $\pm$  s.e.m. (**a–e**). \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ , two-tailed *t*-test. Scale bar, 10  $\mu$ m.



**Extended Data Fig. 3 | Inflammatory-cytokine-independent emergency granulopoiesis and high G-CSF levels in GFI1 ZnF-mutant mice at steady-state and functional analysis of GFI1 ZnF mutants.** **a**, FACS plots and quantification of bone marrow cells from adult Lin<sup>neg</sup>, KIT<sup>+</sup>, SCA-1<sup>+</sup> mice. **b**, Quantification of adult peripheral blood cytokine levels at steady-state by cytokine array. **c**, Quantification of individual adult peripheral blood cytokine levels at steady-state by Luminex analysis. **d**, Survival analysis of lethally irradiated B6J recipients of adult whole bone marrow from adult GFI1-mutant

donors. **e**, Graphical representation of FACS analysis of peripheral blood chimerism in transplant recipients from **d** four months after transplant. **f,g**, Graphical representations of the total number of colonies in mouse organs obtained 24 h after infection with an LD<sub>50</sub> dose of *C. albicans* (**f**) or 5 × 10<sup>7</sup> CFU of *S. aureus* (**g**). **h**, Immunoblot analysis of adult mouse Lin<sup>neg</sup> bone marrow. Data are mean ± s.e.m. \*P < 0.05, \*\*P < 0.01, \*\*\*P < 0.001, \*\*\*\*P < 0.0001, two-tailed t-test (**a–c, e–g**) or two-sided Mantel-Cox test (**d**). Data in **h** are representative of three independent experiments.

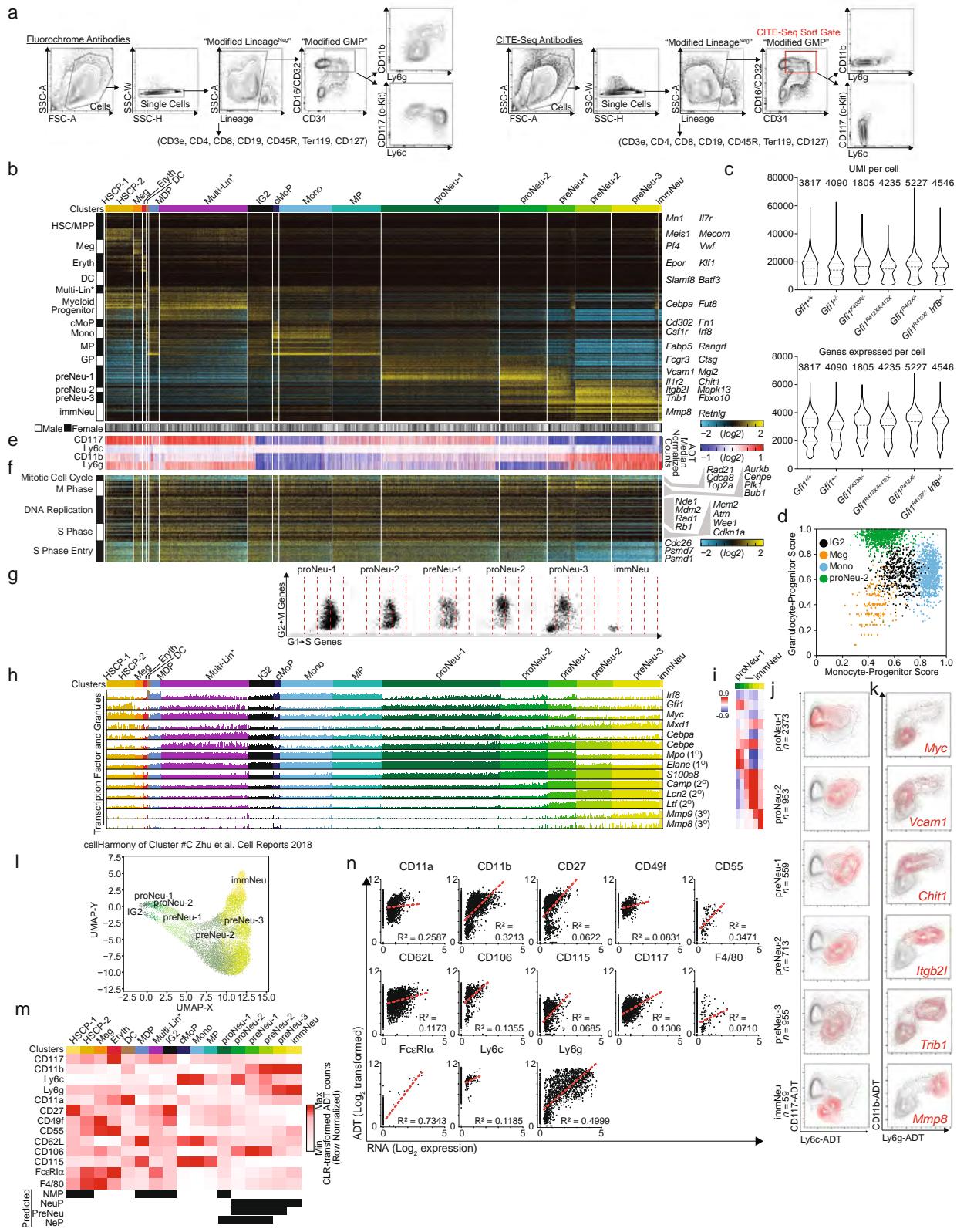
# Article



**Extended Data Fig. 4** | See next page for caption.

**Extended Data Fig. 4 | FACS and Fluidigm-based analysis of steady-state terminal granulopoiesis.** **a**, FACS plots demonstrating the gating strategy used for sorting. The populations sorted for scRNA-seq are indicated in shades of grey with accompanying cytopspins. **b**, Violin plots of the gene and read-level metrics for each of the indicated libraries. Dashed lines indicate mean, lower and upper quartiles. Sample size ( $n$  = number of cells) is displayed at the top. **c**, Heat map of gene expression defined by ICGS (Fluidigm C1, excluding cell-cycle genes) in scRNA-seq data ( $n$  = 516 cells). Each column represents a single cell and each row represents a single gene. ICGS clusters are annotated (top). NK, natural killer T-cell progenitor. **d**, Joint UMAP plot of scRNA-seq data from **c**, separated as previously described<sup>19</sup> and new Fluidigm captures show no detectable batch effects. **e**, Bar chart of the heat map in **c** displaying the incidence and amplitude of selected genes. **f**, Heat map of correlation between gene expression and each displayed cluster as generated by MarkerFinder (AltAnalyze software). **g**, FACS plots comparing expression of LY6G with the indicated surface marker. FMO, fluorescence minus one control. **h**, Heat map of cell cycle gene expression in ICGS-defined clusters in scRNA-seq data ( $n$  = 509 cells). Each column represents a single cell and each row represents a single gene. Gene expression clusters were generated in AltAnalyze and the ICGS clusters are annotated (top). FACS gates are annotated (bottom). LK CD34<sup>+</sup>, Lin<sup>neg</sup> KIT<sup>+</sup> SCA-1<sup>neg</sup> CD34<sup>+</sup>. Key genes are indicated (right). **i**, Scatter plot representation of scRNA-seq data from **h** comparing the gene expression of G1-to-S phase transition genes with G2-to-M phase transition genes in each cell. Each point represents a single cell. **j**, Bar chart of the heat map in **h**, displaying the incidence and amplitude of selected genes. **k**, Heat map of correlation between gene expression and each displayed cluster as generated by MarkerFinder. **l**, Heat map of enrichment for Gene Ontology biological processes enriched in the granulocytic clusters from the Fluidigm scRNA-seq data from **c** with key processes indicated (right). **m**, Scatter plot representation of scRNA-seq data from **c**, in which each point represents a single cell. Reads per cell indicate RNA-seq by expectation–maximization (RSEM) transcript-aligned read counts for each cell library. Genes expressed per cell indicate the number of genes with a transcripts per million (TPM)  $>1$  for each single-cell library. Data in **a** and **g** are representative of three biological replicates; data in **f** and **k** display Pearson correlation values.

# Article

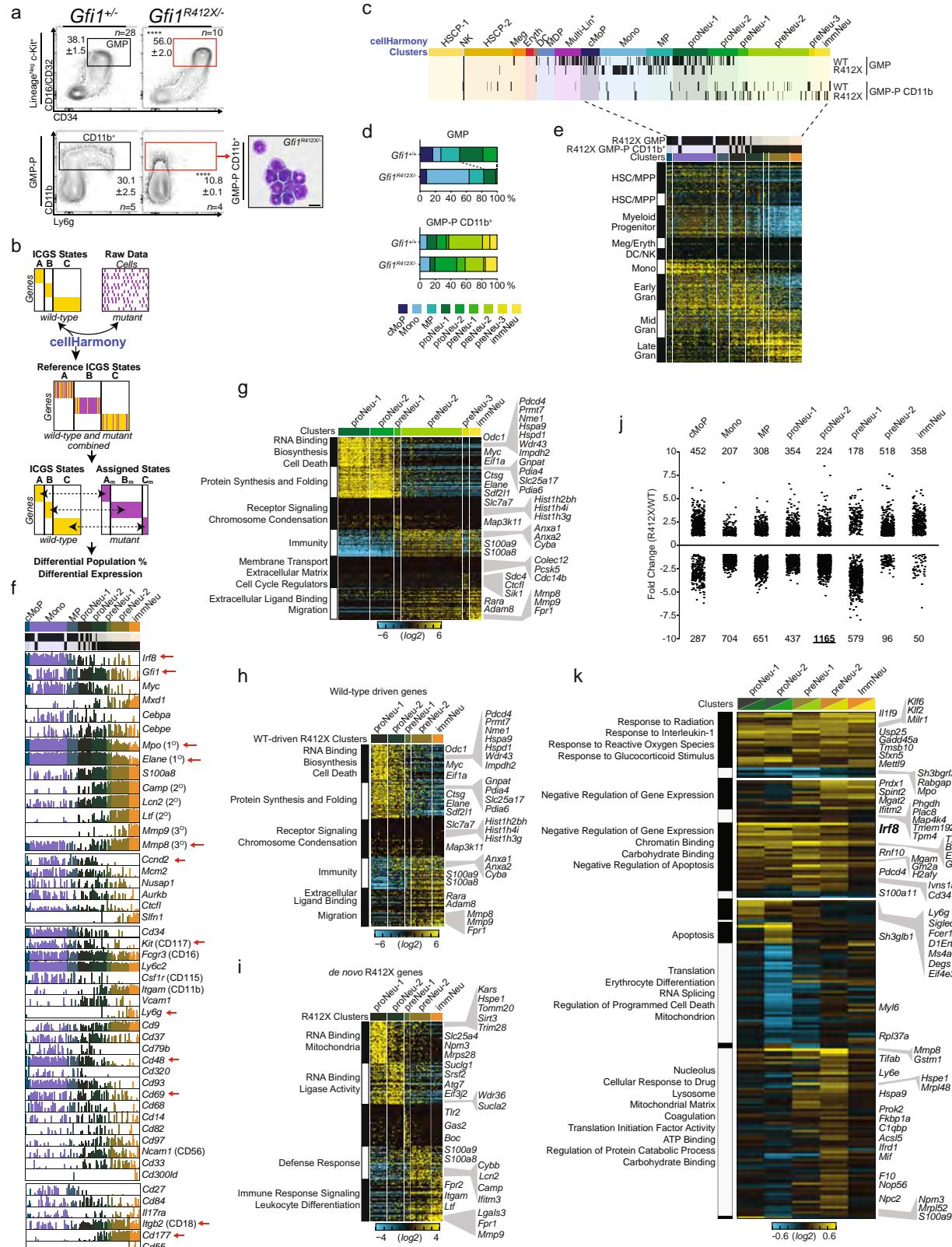


**Extended Data Fig. 5** | See next page for caption.

**Extended Data Fig. 5 | CITE-seq analysis of steady-state terminal granulopoiesis.** **a**, FACS plots demonstrating the gating strategy used for sorting. The population sorted for scRNA-seq is indicated in red. **b**, Heat map of gene expression for cellHarmony-assigned cell populations from CITE-seq 10x Genomics captures (male and female mice,  $n = 11,132$  cells) compared with ICGS-defined clusters from the Fluidigm scRNA-seq data. De novo marker genes (MarkerFinder) for each assigned cluster from the 10x Genomics data are shown (top). Each column represents a single cell and each row represents a single gene. Multi-Lin\*, multi-lineage primed. The gender of the host mouse of each cellular barcode (bottom) and example MarkerFinder genes in common between Fluidigm and 10x Genomics data are indicated (right). **c**, Violin plots of the gene and read-level metrics for each of the indicated libraries. Dashed lines indicate mean, lower, and upper quartiles. Sample size ( $n$  = number of cells) displayed (top). **d**, Lineage-priming scores for monocytic and granulocytic specification. Scatter plot displaying assigned scores for cellHarmony-assigned neutrophil progenitors (proNeu-2), monocytic progenitors (Mono), bi-potential monocytic–granulocytic intermediates (IG2) and megakaryocyte progenitors (Meg) from **b** (see Methods). Each point represents a single cell. **e**, Heat map of row-normalized ADT UMI counts ( $\log_2$ -transformed, median subtracted) for each corresponding cell from **b**. **f**, Heat map of cell cycle gene expression displaying the same cells as **b** and the

same genes as Extended Data Fig. 4h. **g**, Scatter plot representation of scRNA-seq data from **f** comparing the gene expression of G1-to-S phase transition genes with G2-to-M phase transition genes in each cell. Each point represents a single cell. **h**, Bar chart of the heat map in **b**, displaying the incidence and amplitude of selected genes. **i**, Heat map of correlation between gene expression and each displayed cluster as generated by the MarkerFinder feature of AltAnalyze. **j**, **k**, Plots of CITE-seq ADT UMI counts, in which grey indicates all captured cells, and red indicates either cells classified by ICGS clusters (**j**) or the top 1% of UMIs expressing the indicated gene (**k**). **l**, tSNE of 15,968 published cell-barcode<sup>23</sup> coloured according to cellHarmony assigned cell populations (10x Genomics reference) where each dot represents a single cell. Minor contaminant populations (<30 cells per cluster) were excluded. **m**, Heat map of row normalized CLR-transformed ADT counts of the indicated cell surface proteins (left) displayed as an average of all cells in the indicated clusters (top). **n**, Correlation plots between ADT UMI counts and the expression of genes (cellular barcode normalized UMI counts) encoding the corresponding proteins detected via CITE-seq. Each dot represents a single cell. Linear trend lines for all cells with RNA expression > 0 are indicated by dotted red lines with corresponding coefficients of determination displayed. Data in **i** display Pearson correlation values.

# Article



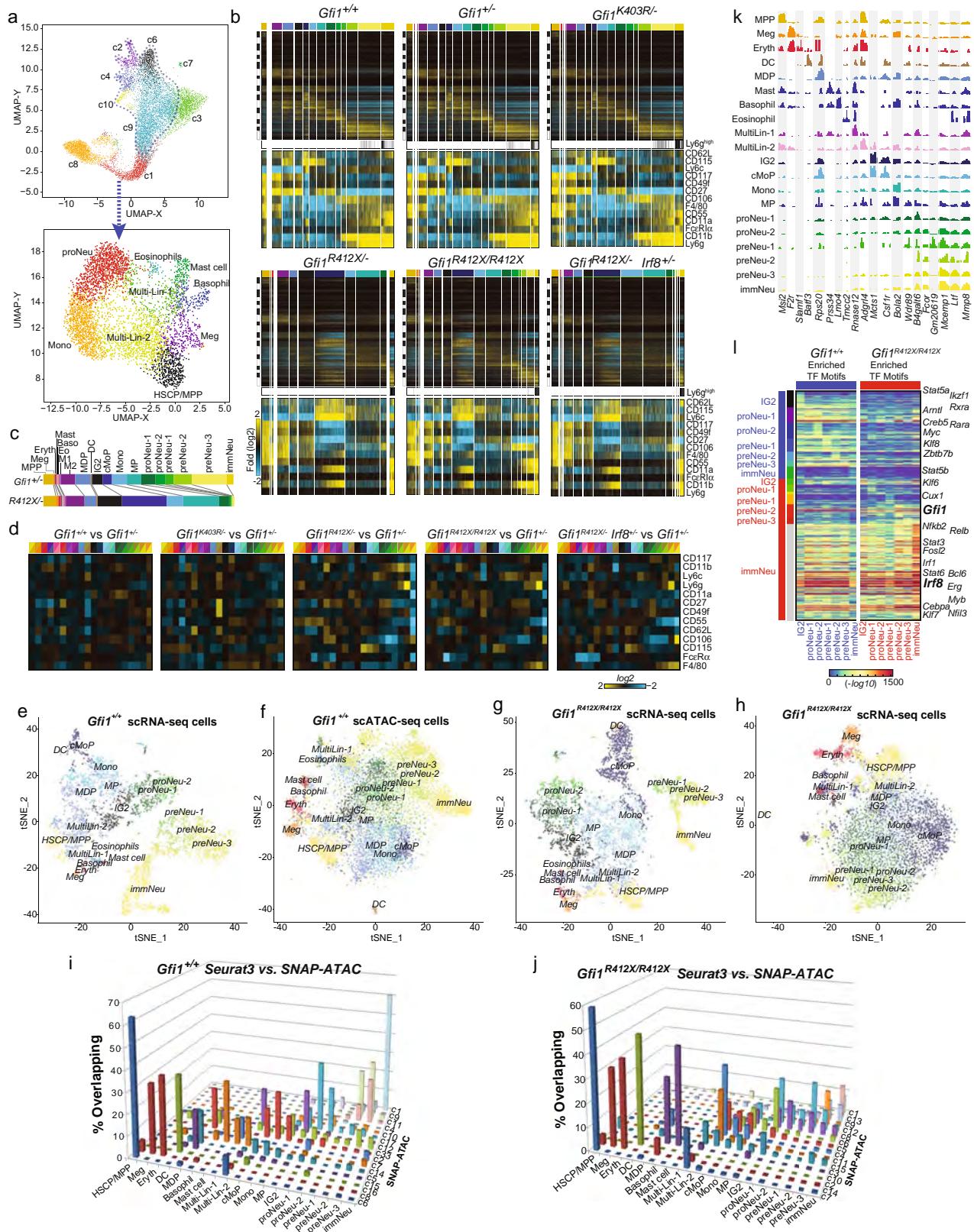
**Extended Data Fig. 6** | See next page for caption.

**Extended Data Fig. 6 | Fluidigm-based transcriptional analysis of *Gfi1*<sup>R412X/-</sup> cells.**

**a**, FACS plots of adult mouse bone marrow (left) and representative cytospin of FACS-sorted adult mouse bone marrow (right). The populations sorted for scRNA-seq are indicated in red. **b**, Schematic summary of the cellHarmony algorithm. **c**, cellHarmony-assigned ICGS states of combined wild-type and *Gfi1*<sup>R412X/-</sup> Fluidigm scRNA-seq data ( $n=624$  cells, excluding cell-cycle genes). Each tick mark represents data from a single cell. Gene expression clusters were generated in AltAnalyze and the ICGS clusters are annotated (top) (Extended Data Fig. 4c). FACS gates are annotated (right). **d**, Population distribution of Fluidigm scRNA-seq data from **c** and Extended Data Fig. 4c. **e**, cellHarmony heat map of wild-type-aligned *Gfi1*<sup>R412X/-</sup> cell gene expression as ordered in **c** (dotted lines). Each column represents a single cell and each row represents a single gene from Extended Data Fig. 4c. The FACS gates and ICGS clusters are annotated (top) and gene clusters are indicated (left). **f**, Bar chart of the heat map in **e**, displaying the incidence and amplitude for selected genes. Arrows indicate differential expression of the adjacent gene, as compared to the wild-type gene. **g**, Heat map of MarkerFinder cell

population-specific genes expression from the Fluidigm scRNA-seq data ( $n=191$  cells) with enriched pathway associated genes (right) and statistically enriched Gene Ontology biological processes (left). **h, i**, Heat maps ( $n=62$  cells) of wild-type-driven genes (same genes as in **g**) (**h**) or de novo MarkerFinder genes for cellHarmony classified *Gfi1*<sup>R412X/-</sup> cell populations (**i**) annotated with enriched pathway associated genes (right) and statistically enriched Gene Ontology biological processes (left). **j**, Representation of the scRNA-seq data from **k**, displaying the fold change in gene expression in *Gfi1*<sup>R412X/-</sup> cells compared to wild-type cells in the indicated clusters, with the number of genes up- or downregulated displayed. Each point represents a single gene. **k**, Heat map of Fluidigm differentially expressed GFI1-target genes (excluding cell-cycle genes), in which each column represents a single cell and each row represents a single gene, with key genes (right) and enriched biological processes (left) indicated. Cytospin data in **a** are representative of two biological replicates. Data in **a** are mean  $\pm$  s.e.m. \*\*\*\* $P<0.0001$ , two-tailed  $t$ -test in **a**.

# Article

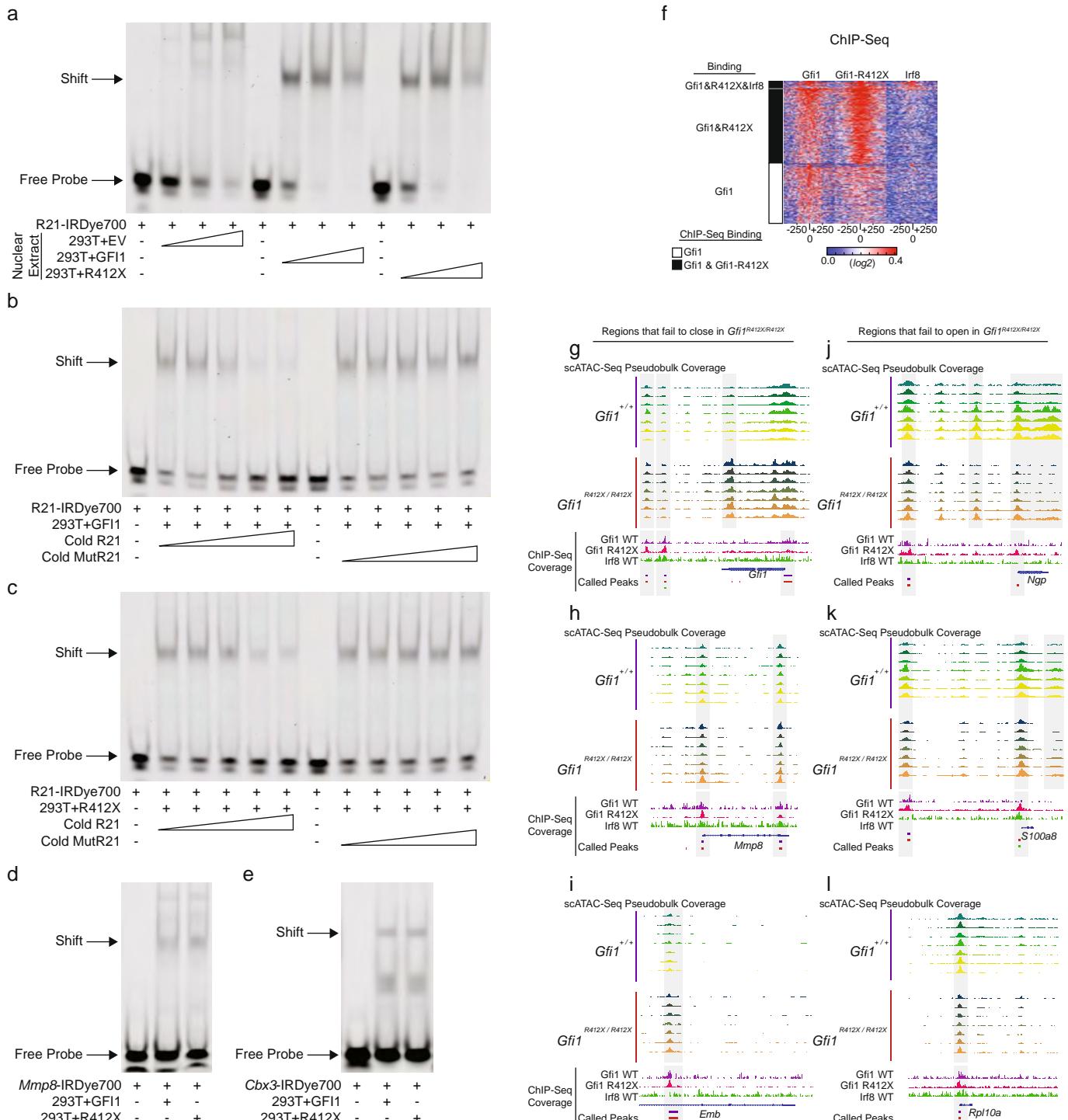


**Extended Data Fig. 7** | See next page for caption.

**Extended Data Fig. 7 | Population analysis of 10x Genomics-based scRNA-seq data and label transfer from CITE-seq transcriptome to scATAC-seq cells.** **a**, Identification of additional cell populations from wild-type scRNA-seq data (ICGS2 unsupervised analysis, male–female CITE-seq datasets). UMAP projections of scRNA-seq data, in which the cell barcodes within the outlined region (top) were analysed for additional heterogeneity through a second ICGS2 analysis (bottom). **b**, Heat maps of additional CITE-seq captures with cell assignments from cellHarmony. cellHarmony classifications were derived using the refined cluster annotation assignments from **a** and Extended Data Fig. 6c. Each panel displays a scRNA-seq heat map of MarkerFinder genes from the cellHarmony reference (top), in which each column represents a single cell and each row represents a single gene. Cell barcodes captured from an independent FACS sort of LY6G<sup>high</sup>CD11b<sup>high</sup> GMP-P cells are indicated by a black bar (middle). Relative expression of median-normalized ADTs (right) are shown in the bottom heat maps. **c**, Assigned cell-population frequencies for *Gfi1*<sup>+/+</sup> and *Gfi1*<sup>R412X/-</sup> CITE-seq (modified GMP gate) cells data sets from cellHarmony. **d**, Heat maps of differential ADT expression compared to *Gfi1*<sup>+/+</sup> for the indicated markers (right). **e**, t-distributed stochastic neighbour embedding (t-SNE) plot of CITE-seq transcriptome for the *Gfi1*<sup>+/+</sup> sample. **f**, t-SNE plot of scATAC-seq for

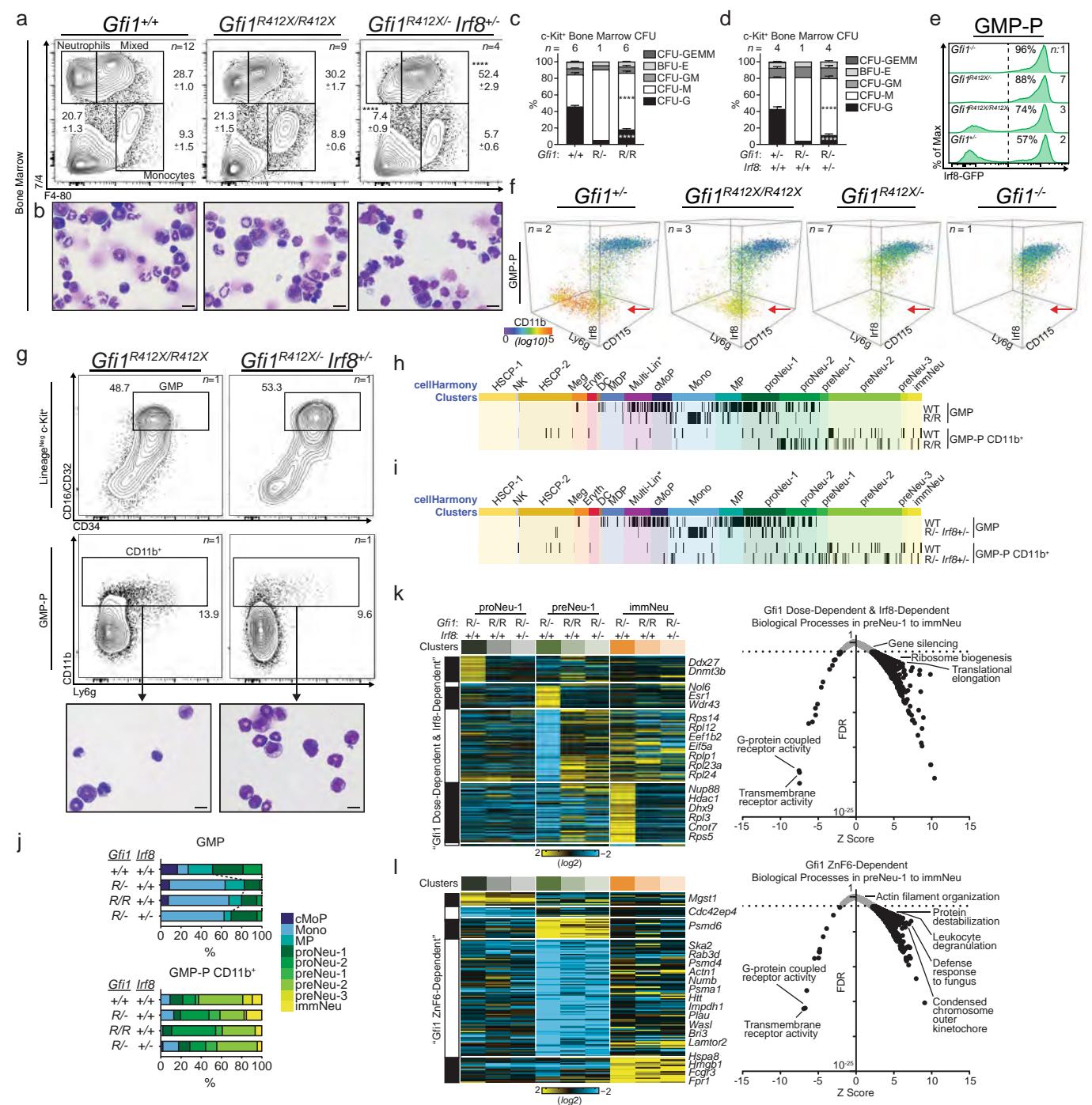
the *Gfi1*<sup>+/+</sup> sample. **g**, t-SNE plot of CITE-seq transcriptome for the *Gfi1*<sup>R412X/R412X</sup> sample. **h**, t-SNE plot of scATAC-seq for the *Gfi1*<sup>R412X/R412X</sup> sample. **i**, Comparison of Seurat-transferred CITE-seq labels to unsupervised scATAC-seq cell population prediction methods (SNAP-ATAC) for *Gfi1*<sup>+/+</sup>. Percentage of overlapping cells for all pairwise comparisons between SNAP-ATAC clusters to CITE-seq clusters derived from Seurat label transfer. **j**, Comparison of Seurat-transferred CITE-seq labels to unsupervised scATAC-seq cell population prediction methods (SNAP-ATAC) for *Gfi1*<sup>R412X/R412X</sup>. Percentage of overlapping cells for all pairwise comparisons between SNAP-ATAC clusters to CITE-seq clusters derived from Seurat label transfer. **k**, Cell-cluster combined scATAC-seq marker peaks associated with *Gfi1*<sup>+/+</sup> CITE-seq annotated cell populations (Seurat 3 label transfer using cicero gene activity scores). Each row is a cluster (left) and each column is a locus within 50 kb of a CITE-seq marker gene for each cluster (bottom), in which the coloured bars represent the normalized read count coverage. **l**, Heat maps of transcription factor-motif-enrichment probabilities ( $-\log_{10}$ -transformed) in accessible regions of *Gfi1*<sup>+/+</sup> (blue) or *Gfi1*<sup>R412X/R412X</sup> (red) cell populations (indicated left). Each row represents a transcription factor motif from the Cisbp2 database (indicated right). Each dot in **a** and **e–h** represents a single cell barcode that is pseudo-coloured for its ICGS2 predicted cluster.

# Article



**Extended Data Fig. 8 | In vitro DNA-binding analysis and in vivo ChIP-seq and scATAC-seq analysis of GFI1(R412X).** **a–e**, EMSA using nuclear extracts of 293T cells transfected with a GFI1, GFI1(R412X), or empty expression vector. Arrows indicate bound (shift) and free probe. All data represent one experiment. **a–c**, A probe containing a high-affinity GFI1-binding site (R21) was incubated with a titration of nuclear extracts (**a**) or a titration of cold competitors with (R21) or without (MutR21) a high-affinity GFI1-binding site (**b**, **c**). **d**, **e**, Nuclear extracts were incubated with probes containing a GFI1-binding site taken from the *Mmp8* (**d**) or *Cbx3* (**e**) locus. These loci contain a GFI1 ChIP-seq peak and the corresponding genes are differentially expressed in *Gfi1*<sup>R412X/-</sup>.

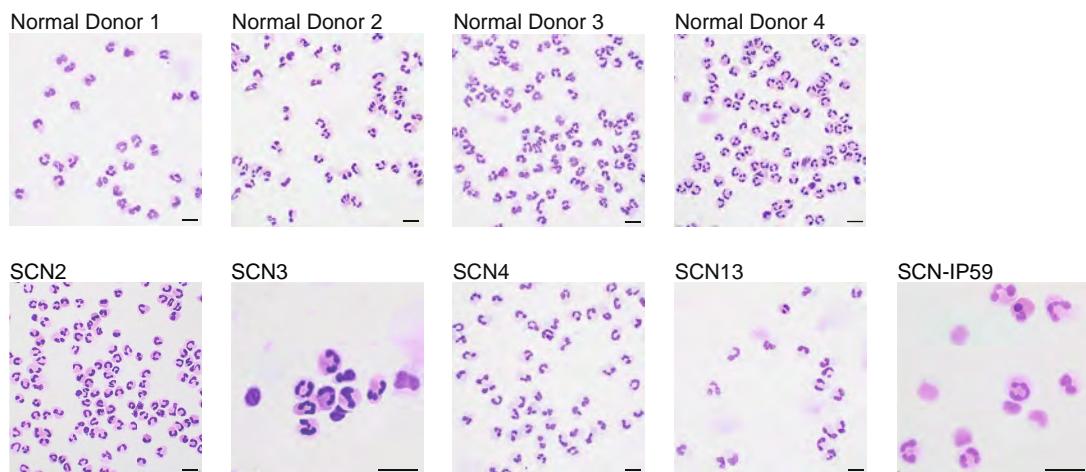
Fluidigm scRNA-seq data. **f**, Heat map of ChIP-seq read coverage at GFI1-specific peaks, in which each row represents one peak that was called in the wild-type and each column represents a DNA-base position of the peak centred on each GFI1 called peak in the indicated ChIP samples. The cluster colour (left) indicates loci bound by GFI1 alone (white) or by GFI1 and GFI1(R412X) (black). **g–l**, Schematics of the indicated genomic loci displaying scATAC-seq pseudobulk accessibility in the colour-coded clusters (top) or ChIP-seq reads and called peaks (bottom), with differentially accessible regions bound by GFI1 shaded.



**Extended Data Fig. 9 | Fluidigm-based transcriptional analysis of *Gfi1*<sup>R412X/-</sup> genetic rescues.** **a, b**, FACS plots (a) and representative cytospins (b) of bone marrow from adult mouse. **c, d**, Graphical summary of CFU assays performed on KIT<sup>+</sup> bone marrow cells. **e**, FACS plots of mouse bone marrow cells isolated from *Irf8*-eGFP transgenic mice with the mean percentage of IRF8-GFP<sup>high</sup> cells indicated. **f**, Three-dimensional FACS plots of bone marrow isolated from adult *Irf8*-eGFP mice. Events are pseudocoloured for CD11b expression and arrows indicate changes in the incidence of granulocytic populations. **g**, FACS plots and representative cytospins of adult mouse bone marrow populations sorted for scRNA-seq. **h, i**, cellHarmony-assigned

*Gfi1*<sup>R412X/R412X</sup> and *Gfi1*<sup>R412X/- Irf8<sup>-/-</sup> cells to wild-type reference Fluidigm scRNA-seq populations (**h**,  $n = 86$  cells, **i**,  $n = 88$  cells). Each tick mark represents a single-cell library. ICGS clusters are annotated (top). FACS gates are annotated (right). **j**, Population distribution of Fluidigm scRNA-seq data from **h**, **i** and Extended Data Figs. 4c, 6c, **j**. **k**, Fluidigm scRNA-seq heat maps (left) of genes that are genetically repaired in each cluster, with key genes indicated. Adjacent plot (right) of enriched biological processes in the indicated clusters. Scale bars, 10  $\mu$ m. Cytospin data in **b** are representative of three biological replicates. Data are mean  $\pm$  s.e.m. (**a, c, d**) or mean (**e**). \*\*\*\* $P < 0.0001$ , two-tailed  $t$ -test.</sup>

## Article



**Extended Data Fig. 10 | G-CSF rescued neutrophils from patients with SCN are morphologically normal but functionally defective.** Representative cytospins of purified human peripheral blood neutrophils. Data are representative of one experiment per donor. Scale bars, 10  $\mu$ m.

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated
- Clearly defined error bars  
*State explicitly what error bars represent (e.g. SD, SE, CI)*

*Our web collection on [statistics for biologists](#) may be useful.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

FACS data was collected using FACSDiva version 8.0, and imaging data was collected using Nikon NIS Elements or Licor Image Studio.

Data analysis

All described code is provided in Github (<https://github.com/nsalomonis/altanalyze>). Software used for data analysis includes: RSEM version 1.3.0, ICGS2 (AltAnalyze version 2.1.3), Cell Ranger workflow version 2.1.0, R version 3.5, samtools version 1.9, macs2 version 2.1.2, bedtools version 2.27.1, IGV version 2.5.3, Ensembl Biomart database version 96, cellranger-atac count program version 1.0.1, Seurat version 3.0.2, Cicero version 1.0.15, homer version 4.11, SeqGeq version 1.1, and ApE version 2.0.49.0.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

scRNA-Seq and CITE-Seq data is deposited as GEO SuperSeries GSE120409 and in Synapse (<https://www.synapse.org/#/Synapse:syn16806696>). Proteomic data are available via ProteomeXchange ([www.ebi.ac.uk/pride/archive/](http://www.ebi.ac.uk/pride/archive/)) with identifier PXD010943. Extensive processed and primary data are provided and organized in Synapse, including 10x Genomics count matrices, Fluidigm expression, single-cell populations, quality control metrics, 10x Genomics Cell Ranger outputs (summary report, loupe browser files), genomic coordinate peak files and differential expression results (cellHarmony)."

## Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](http://nature.com/authors/policies/ReportingSummary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. Instead, sample sizes were rationalized by weighing sufficient replication (to determine the extent of biological variation) with reduction of total animals used.
Data exclusions	No biological data was excluded. Bioinformatically, single cell RNA-Seq data of low quality was excluded according to pre-established criteria to ensure robust and accurate downstream analyses of the data as a whole.
Replication	Results were confirmed by analysis of individual biological replicates, and all attempts at replication were successful.
Randomization	Sample allocation was not random. Instead, biological controls were included in all experiments.
Blinding	The investigators were blinded to allocation for <i>C. albicans</i> and <i>S. aureus</i> experiments. Other blinding was not possible since genotyping was necessary for all other experiments.

## Reporting for specific materials, systems and methods

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Unique biological materials
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants

### Methods

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

### Antibodies used

For immunoblots, goat anti-Gfi1 (R&D Systems), mouse anti-beta-Actin (Sigma Aldrich), HRP-conjugated anti-goat secondary antibody (Invitrogen), and HRP-conjugated anti-mouse secondary antibody (GE Healthcare) were used.

For FACS, CD16/CD32 (clone 2.4G2, Becton, Dickinson and Company), CD3-APC (clone 145-2C11, Becton, Dickinson and Company), CD45R-PE (clone RA3-6B2, Becton, Dickinson and Company), CD11b-PacificBlue (clone M1/70, BioLegend), and Ly6g-PerCP-Cy5.5 (clone 1A8, BioLegend), CD45.1-BV605 (clone A20, Biolegend) and CD45.2-AlexaFluor700 (clone 104, Biolegend) CD16/CD32 (clone 2.4G2, Becton, Dickinson and Company), and Ly-6B.2-AlexaFluor647 (clone 7/4, Bio-Rad), F4-80-AlexaFluor488 (clone Cl:A3-1, Bio-Rad) Fc Block (Becton, Dickinson, and Company), biotin-conjugated anti-CD3e (clone 145-2C11,

BioLegend), biotin-conjugated anti-CD4 (clone RM4-5, Thermo Fisher), biotin-conjugated anti-CD8 (clone 53-6.7), biotin-conjugated anti-CD11b (clone M1/70, Becton, Dickinson, and Company), biotin-conjugated anti-CD19 (clone 6.D5, BioLegend), biotin-conjugated anti-CD127 (clone B12-1, Becton, Dickinson, and Company), biotin-conjugated anti-B220 (clone RA3-6B2, BioLegend), biotin-conjugated anti-Gr1 (clone RB6-8C5, Becton, Dickinson, and Company), and biotin-conjugated anti-Ter119 (clone TER-119, BioLegend), streptavidin-APC-Cy7 (Becton, Dickinson, and Company), APC-conjugated anti-CD117 (clone 2B8, BioLegend), PE-Cy7-conjugated anti-Sca-1 (clone D7, Becton, Dickinson, and Company), FITC-conjugated anti-CD34 (clone RAM34, Becton, Dickinson, and Company), PacificBlue-conjugated anti-CD48 (clone HM48-1, BioLegend), PE-conjugated anti-CD135 (clone A2F10, BioLegend) and Brilliant Violet 510-conjugated anti-CD150 (clone TC15-12F12.2, BioLegend), PerCP-eFluor710-conjugated anti-CD16/CD32 (clone 93, Thermo Fisher), Brilliant Violet 421-conjugated anti-CD34 (clone RAM34, Becton, Dickinson and Company), AlexaFluor700-conjugated anti-CD11b (clone M1/70, BioLegend), Brilliant Violet 785-conjugated anti-Ly6c (clone HK1.4, BioLegend), FITC-conjugated anti-Ly6g (clone 1A8, BioLegend), Brilliant Violet 605-conjugated anti-CD115 (clone T38-320, Becton, Dickinson, and Company), Brilliant UV 395-conjugated anti-CD16/CD32 (clone 93, Becton, Dickinson, and Company), PE-Cy7-conjugated anti-CD11b (clone M1/70, BioLegend), PerCP-Cy5.5-conjugated anti-Ly6g (clone 1A8, BioLegend), Brilliant Violet 650-conjugated anti-CD117 (clone 2B8, BioLegend), PE-conjugated anti-CD18 (clone M18/2, BioLegend), PE-conjugated anti-CD27 (clone LG.3A10, BioLegend), PE-conjugated anti-CD55 (clone RIKO-3, BioLegend), PE-conjugated anti-CD84 (clone mCD84.7, BioLegend), Alexa Fluor 700-conjugated anti-CD177 (clone 1171A, R&D Systems, Minneapolis, MN, USA), PE-conjugated anti-CD106 (Vcam1, clone 429, BioLegend), and PE-conjugated anti-CD217 (IL-17R, clone PAJ-17R, Thermo Fisher) were used.

For CITE-Seq: TotalSeq antibodies (anti-CD117 (clone 2B8), anti-Ly-6C (clone HK1.4), anti-CD11b (clone M1/70), anti-Ly-6G (clone 1A8), anti-CD11a (clone M17/4), anti-CD27 (clone LG.3A10), anti-CD49f (clone GoH3), anti-CD55 (clone RIKO-3), anti-CD62L (clone MEL-14), anti-CD106 (clone 429), anti-CD115 (clone AFS98), anti-FceR1a (clone MAR-1), anti-F4/80 (clone BM8), and TotalSeq-A anti-mouse hashtag antibodies (Hashtags 1-6, BioLegend) were used.

For Gfi1 ChIP-Seq: Rabbit anti-sera against Gfi1 (previously described in Spooner et al, Nat Immunol 2013) was used.

#### Validation

The goat anti-Gfi1 antibody was empirically validated for detection of murine Gfi1 protein in an immunoblot assay by using cell lysates from Gfi1<sup>-/-</sup> mice. No signal was detected at 55 kD in the lane loaded with cell lysates from Gfi1<sup>-/-</sup> mice. All other antibodies were validated by the source company/lab.

## Eukaryotic cell lines

#### Policy information about [cell lines](#)

Cell line source(s)	LentiX 293T cells were purchased from Takara Bio.
Authentication	No cell line authentication was performed.
Mycoplasma contamination	Cell lines were not tested for mycoplasma contamination.
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	No commonly misidentified cell lines were used.

## Animals and other organisms

#### Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Male and female mice were maintained on a C57BL/6 background and were analyzed between 5 and 12 weeks of age.
Wild animals	The study did not include wild animals.
Field-collected samples	The study did not involve samples collected from the field.

## Human research participants

#### Policy information about [studies involving human research participants](#)

Population characteristics	For identification of GFI1 mutations, de-identified clinical data was obtained at Cincinnati Children's Hospital Medical Center through informed consent under an approved institutional review board research protocol. Additional covariate-relevant population characteristics were not made available through the IRB-approved protocol.  For analysis of peripheral blood neutrophils, both male and female neutropenic patients with ages ranging from newborn to 15 years of age were obtained at Cincinnati Children's Hospital Medical Center through informed consent under an approved institutional review board research protocol. Additional covariate-relevant population characteristics (age, gender, genetic variants detected, diagnosis, dose of G-CSF received) are provided in the Supplementary Information.
Recruitment	De-identified data and peripheral blood from healthy donors or neutropenic patients was obtained at Cincinnati Children's Hospital Medical Center through informed consent under an approved institutional review board research protocol. Neutropenia is a rare disease, so all patients being seen in the clinic were asked to participate.

## ChIP-seq

### Data deposition

- Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

#### Data access links

*May remain private before publication.*

ChIP-Seq data is deposited as GEO SuperSeries GSE120409 and in Synapse (<https://www.synapse.org/#!Synapse:syn16806696>). Reviewers may access the GEO data using the "cxfeeqrjclcd" token and the Synapse data using the Username: tempviewer1 and Password: l6GawO3N.

#### Files in database submission

GSM4225637 Gfi1 GMP ChIP-Seq (Gfi1 R412X-R412X mice)

#### Genome browser session (e.g. [UCSC](#))

[https://www.altanalyze.org/ucsc/hub\\_Gfi1\\_R412X\\_mouse\\_model\\_scatacseq\\_and\\_chipseq.txt](https://www.altanalyze.org/ucsc/hub_Gfi1_R412X_mouse_model_scatacseq_and_chipseq.txt)

### Methodology

#### Replicates

No replicates were used.

#### Sequencing depth

For this sample, 34,222,568 reads were generated, of which 32,165,666 aligned to mm10, 26,353,725 were -q25, 18,465,798 unique alignments were made (without duplicate reads) and there were 18,461,755 final tags (-chrM, >chomsize).

#### Antibodies

Rabbit anti-sera against Gfi1 (previously described in Spooner et al, Nat Immunol 2013) was used.

#### Peak calling parameters

Read mapping: BWA (default parameters) Peak Calling: macs2 (version 2.1.2) (default parameters)

#### Data quality

The 75-nt single-end (SE75) sequence reads generated by Illumina sequencing (using NextSeq 500) are mapped to the genome using the BWA algorithm with default settings. Alignment information for each read is stored in the BAM format. Only reads that pass Illumina's purity filter, align with no more than 2 mismatches, and map uniquely to the genome are used in the subsequent analysis. In addition, duplicate reads ("PCR duplicates") are removed. Peaks were called using MACS (v.2.1.0) with default values. Peak filtering was performed by removing false ChIP-Seq peaks as defined within the ENCODE blacklist using bedtools (version 2.27.1). In total, 19039 peaks are at FDR 5% and above 5-fold enrichment.

#### Software

Software: samtools (version 1.9), macs2 (version 2.1.2), bedtools (version 2.27.1), IGV (version 2.5.3), IGV\_tools (version 2.7), AltAnalyze (version 2.1.3), and UCSC binary tools

#### Custom code:

In order to help normalize for total reads captured, BAM alignment files were downsampled (samtools view -s) ("downsample\_chips.sh")

Next, peaks were called (macs -callpeak) ("downsampled\_run\_macs.sh")  
Then, peaks were processed to remove blacklisted peaks from the ENCODE project (bedtools intersect -v) and to find merged locations between the 3 sets analyzed within 200bp of each other. (bedtools merge -d 200) Bedgraph files were also generated to display the data on a browser (bedtools genomecov) and transformed into bigwig files (UCSC binary tool bedGraphToBigWig) as well as .tdf files (igvtools count). Then the base pair position coverage value was determined for the merged peaks in order to plot the coverages +/- 200 bp from the center of the peak (bedtools coverage -d -counts). ("downsampled\_process\_macs2\_peaks.txt").

The coverage positions were organized using a custom R script. ("make\_coverages.r")

Then the coverages were plotted. (AltAnalyze hierarchical clustering option)

## Flow Cytometry

### Plots

#### Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

#### Sample preparation

Mice were euthanized using carbon dioxide followed by cervical dislocation. Peripheral blood was collected into EDTA-coated tubes, and hind limb bones (femurs, tibias, and the iliac crest) were harvested immediately after euthanasia and stored in cold FACS buffer (1% FBS, 0.01% NaN3 in DPBS) under sterile conditions. Bones were flushed using a syringe for transplantation or crushed using a mortar and pestle for all other applications, then passed through a 40 µm cell strainer (Becton, Dickinson and Company) to obtain single cell suspensions for downstream applications. Prior to analytical flow cytometry, erythrocytes were lysed using ACK buffer (Gibco), then cells were washed in FACS Buffer and resuspended in FACS Buffer containing DAPI (Thermo Fisher) or 7-AAD (Becton, Dickinson, and Company).

Human peripheral blood was collected in EDTA-coated tubes for neutrophil purification.

Instrument

Flow cytometry analyses were conducted on a FACS LSRII or FACS LSRFortessa (Becton, Dickinson, and Company).

Cell sorting was performed on a FACSAria II (Becton, Dickenson, and Company).

Software

FACS data was analyzed using FlowJo software (version 10.1 or 10.2, FlowJo, Ashland, OR, USA).

Cell population abundance

No post-sort analysis was performed.

Gating strategy

For all human and murine analyses, DAPI+ or 7AAD+ cells were excluded, then debris was excluded using a FSC-A vs SSC-A gate, then doublets were excluded using FSC-W vs FSC-H or SSC-W vs SSC-H gates for all downstream gating. Fluorescent minus-one controls were used in some circumstances to assist in discriminating between positive and negative signal, while other gating was performed according to previously published strategies.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

## Article

# LEM2 phase separation promotes ESCRT-mediated nuclear envelope reformation

<https://doi.org/10.1038/s41586-020-2232-x>

Received: 8 March 2019

Accepted: 26 February 2020

Published online: 29 April 2020

 Check for updates

Alexander von Appen<sup>1,7</sup>, Dollie LaJoie<sup>2,7</sup>, Isabel E. Johnson<sup>1,7</sup>, Michael J. Trnka<sup>3</sup>, Sarah M. Pick<sup>4</sup>, Alma L. Burlingame<sup>3</sup>, Katharine S. Ullman<sup>2,✉</sup> & Adam Frost<sup>1,5,6,✉</sup>

During cell division, remodelling of the nuclear envelope enables chromosome segregation by the mitotic spindle<sup>1</sup>. The reformation of sealed nuclei requires ESCRTs (endosomal sorting complexes required for transport) and LEM2, a transmembrane ESCRT adaptor<sup>2–4</sup>. Here we show how the ability of LEM2 to condense on microtubules governs the activation of ESCRTs and coordinated spindle disassembly. The LEM motif of LEM2 binds BAF, conferring on LEM2 an affinity for chromatin<sup>5,6</sup>, while an adjacent low-complexity domain (LCD) promotes LEM2 phase separation. A proline–arginine-rich sequence within the LCD binds to microtubules and targets condensation of LEM2 to spindle microtubules that traverse the nascent nuclear envelope. Furthermore, the winged-helix domain of LEM2 activates the ESCRT-II/ESCRT-III hybrid protein CHMP7 to form co-oligomeric rings. Disruption of these events in human cells prevented the recruitment of downstream ESCRTs, compromised spindle disassembly, and led to defects in nuclear integrity and DNA damage. We propose that during nuclear reassembly LEM2 condenses into a liquid-like phase and coassembles with CHMP7 to form a macromolecular O-ring seal at the confluence between membranes, chromatin and the spindle. The properties of LEM2 described here, and the homologous architectures of related inner nuclear membrane proteins<sup>7,8</sup>, suggest that phase separation may contribute to other critical envelope functions, including interphase repair<sup>8–13</sup> and chromatin organization<sup>14–17</sup>.

In late anaphase, as the nuclear envelope surrounds the chromatin disc, LEM2 concentrates at the nuclear envelope core proximal to spindle microtubules. This is in contrast to the localization pattern of the non-core marker lamin B2 and to the even distribution of LEM2 around the nuclear envelope during interphase<sup>18,19</sup> (Fig. 1a, Extended Data Fig. 1a). To determine how LEM2 concentrates within the core, we used live-cell imaging to monitor mutant LEM2–mCherry (mChr) constructs alongside GFP-tagged tubulin. Because, as expected, we found that the LEM domain of LEM2 bound with high affinity to BAF (Extended Data Fig. 1b–d), we tested whether a four-amino-acid substitution in the LEM domain of LEM2 (LEM2<sub>m21</sub>) would disable BAF binding<sup>20</sup>. This mutation was sufficient to disrupt accumulation of LEM2 within the nuclear envelope and its subsequent core enrichment in anaphase cells, consistent with the role of BAF as an early nuclear envelope reformation factor and a constituent of the core region<sup>5</sup> (Fig. 1a, Extended Data Fig. 1c). Unexpectedly, removing the adjacent LCD (LEM2<sub>ΔLCD</sub>; ΔAA 43–202) also compromised nuclear envelope enrichment, despite the presence of an intact LEM domain. This observation suggests that the LEM domain and the LCD are each necessary, but not sufficient alone, for proper targeting (Fig. 1a). By contrast, a deletion within the C-terminal winged helix (WH) domain of LEM2 (LEM2<sub>ΔWH</sub>; Δ amino acids (AA) 415–485) did not affect core localization of LEM2 (Fig. 1a).

Phase separation by LCDs, or the formation of biomolecular condensates, is a mechanism for organizing dynamic processes through

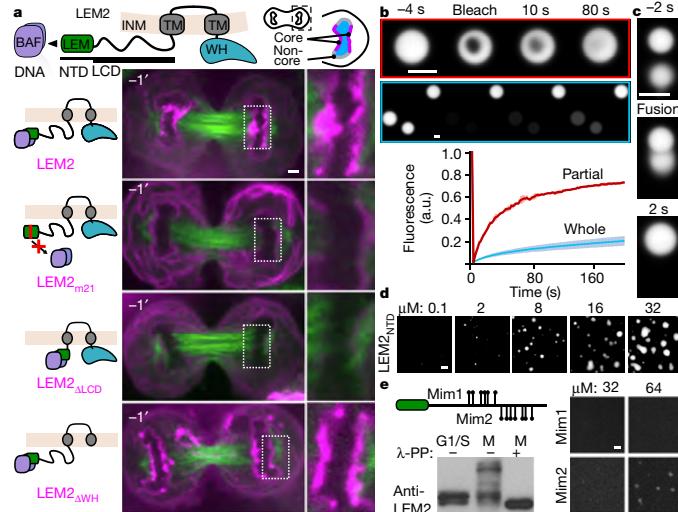
the creation of membraneless, biochemically specialized compartments<sup>21–23</sup>. To characterize the properties and roles of the LEM2 LCD in core localization, we purified the N-terminal domain (NTD) of LEM2, which contains the LEM and LCD domains (LEM2<sub>NTD</sub>; AA 1–208). The LCD is notable for its basic sequence (*pI* = 11.8), which results from overrepresentation of arginine, in addition to overrepresentation of proline and other residues (Extended Data Fig. 1e). After purification, LEM2<sub>NTD</sub> spontaneously formed spherical droplets with liquid-like properties under physiological conditions. These LEM2<sub>NTD</sub> droplets could undergo complete fusion and inter- and intra-droplet diffusion on a timescale of seconds (Fig. 1b–d, Supplementary Video 1).

We next sought to address whether and how cells exploit the ability of LEM2 to phase separate in mammalian open mitosis (in which the nuclear envelope breaks down). Phosphorylation regulates both the cell cycle and phase transitions<sup>24,25</sup>, and LEM2 has several annotated phosphorylation sites within its LCD<sup>26</sup>. To test whether LEM2 could be regulated by phosphorylation, we arrested HeLa cells in G1/S phase or mitosis and compared the phosphorylation status of LEM2 in each state. As expected, LEM2–mChr appeared to be more strongly phosphorylated in mitosis than in G1/S, and this was reversible with phosphatase treatment (Fig. 1e).

Constructs in which two subregions of the LCD had been deleted, each containing annotated phosphorylation sites (an SY-rich region and a PR-rich region, LEM2<sub>ΔSY</sub>–mChr ΔAA 75–123 and LEM2<sub>ΔPR</sub>–mChr

<sup>1</sup>Department of Biochemistry and Biophysics, University of California, San Francisco, CA, USA. <sup>2</sup>Department of Oncological Sciences, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA. <sup>3</sup>Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA, USA. <sup>4</sup>Faculty of Chemistry and Pharmacy, University of Freiburg, Freiburg, Germany. <sup>5</sup>Chan Zuckerberg Biohub, San Francisco, CA, USA. <sup>6</sup>Quantitative Biosciences Institute, University of California, San Francisco, San Francisco, CA, USA. <sup>7</sup>These authors contributed equally: Alexander von Appen, Dollie LaJoie, Isabel E. Johnson. <sup>✉</sup>e-mail: katharine.ullman@hci.utah.edu; adam.frost@ucsf.edu

# Article

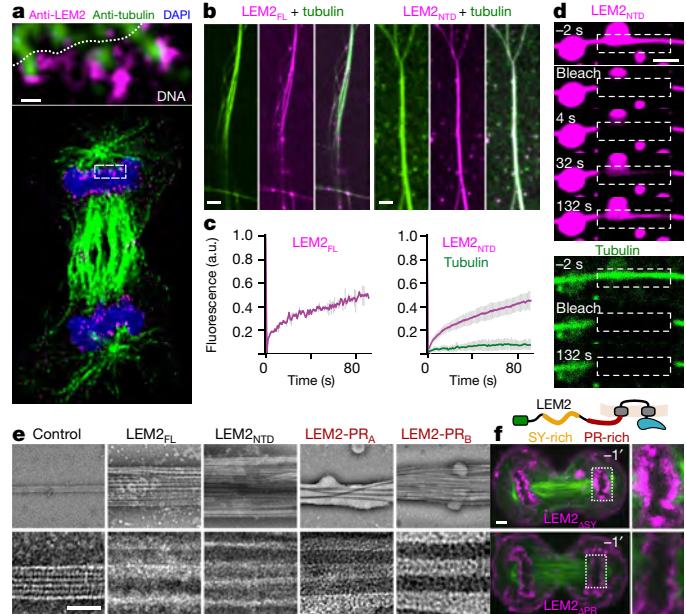


**Fig. 1 | Targeting of LEM2 to the nuclear envelope core at anaphase chromatin discs depends on BAF binding and a LCD that can form liquid-like droplets.** **a**, Top, cartoon of LEM2 sequence motifs and cellular localization. TM, transmembrane. Bottom, live-cell imaging of GFP-tubulin and LEM2-mCh constructs. Time 0 is time of complete cleavage furrow ingression (CFI). Representative of results from three biological replicates except for  $\Delta$ LCD image, which is representative of two biological replicates. **b**, Top, partial and whole droplet fluorescence recovery after photobleaching (FRAP). Bottom, quantification shows mean  $\pm$  s.d. of  $n = 3$  independent samples. **c**, Real-time fluorescence imaging of LEM2<sub>NTD</sub> droplet fusion. Representative of two independent experiments. **d**, Concentration-dependent droplet formation of purified LEM2<sub>NTD</sub>. Representative of three independent experiments. **e**, Top, schematic indicating sites of LEM2<sub>NTD</sub> phosphomimetic mutations in two constructs (Mim1 and Mim2). Bottom, LEM2 immunoblot assessing the migration pattern of full-length LEM2-mCh following separation by Phos-tag SDS-PAGE; lysates treated with lambda phosphatase ( $\lambda$ -PP) as indicated. Representative of two biological replicates, with one and three technical replicates per biological replicate. For immunoblot source data, see Supplementary Fig. 1. Right, fluorescence imaging of purified LEM2<sub>NTD</sub>-phosphomimetic constructs. Representative of two technical replicates. Scale bar, 2  $\mu$ m (a–e).

ΔAA 145–213) displayed fewer phospho-LEM2 species, suggesting that mitotic phosphosites lie within the deleted subsequences (Extended Data Fig. 1f). Notably, phosphomimetic mutations that spanned either sub-region prevented purified LEM2<sub>NTD</sub> from forming droplets near the physiological regime (Fig. 1e), which suggests that cell-cycle-dependent phosphorylation governs the ability of LEM2 to condense. A synthetic peptide from the SY-rich sequence also phase separated, albeit at a notably higher concentration than the LCD (Extended Data Fig. 1g). Finally, complex coacervation with polyanions such as DNA and RNA also promoted monodisperse LEM2<sub>NTD</sub> to condense into spherical liquid-like droplets—a property shared with other PR-rich proteins<sup>27</sup> (Extended Data Fig. 1h). Thus, the LEM2 LCD promotes anaphase core condensation in cells and exhibits both spontaneous liquid-like phase separation and complex coacervation properties in vitro.

## LEM2 forms a liquid-like coat on microtubules

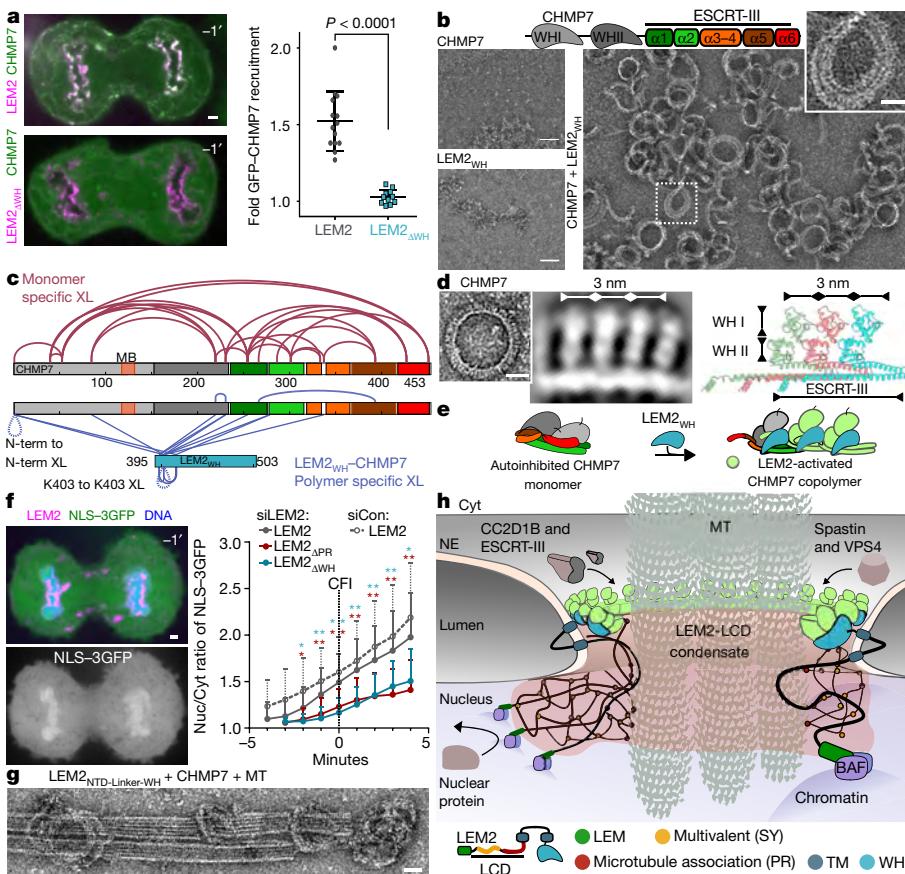
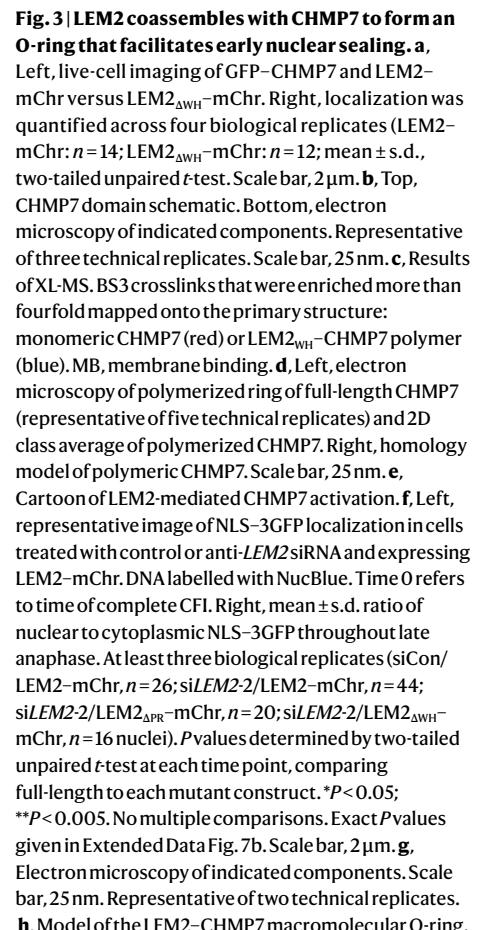
Super-resolution stimulated emission depletion (STED) imaging of immunostained HeLa cells revealed that endogenous LEM2 enriched specifically in toroid-like shapes around spindle microtubules at the chromatin surface (Fig. 2a, Extended Data Fig. 2a). To investigate whether LEM2 binds microtubules directly, we tested binding in vitro with fluorescently labelled proteins and observed that detergent-solubilized full-length LEM2 (LEM2<sub>FL</sub>) stabilized and bundled microtubules under physiological conditions (Fig. 2b). After reconstitution into



**Fig. 2 | LEM2 concentrates around spindle microtubules and its LCD forms a liquid-like coating around microtubules.** **a**, STED imaging of endogenous LEM2 in late anaphase. Scale bar, 150 nm. Representative of data from four cells. **b**, Left, fluorescence imaging of full-length LEM2 (LEM2<sub>FL</sub>) and microtubules. Representative of two technical replicates. Right, fluorescence imaging of LEM2<sub>NTD</sub> and microtubules. Representative of ten technical replicates. **c**, FRAP analysis of LEM2<sub>FL</sub> and LEM2<sub>NTD</sub>-coated microtubule bundles. Mean  $\pm$  s.e.m. for  $n = 5$  (LEM2<sub>FL</sub>) or  $n = 17$  (LEM2<sub>NTD</sub>) independent samples. **d**, FRAP analysis of a LEM2<sub>NTD</sub>-coated microtubule bundle. Images representative of two independent examples. **e**, Electron micrographs of microtubules alone or with indicated portions of purified LEM2. LEM2-PR<sub>A</sub>, AA 145–165; LEM2-PR<sub>B</sub>, AA 188–213. Representative of three technical replicates. Scale bar, 25 nm. **f**, Top, schematic architecture of LEM2 highlighting the SY-rich and PR-rich regions. Bottom, live-cell imaging of indicated LEM2-mCh deletion constructs and GFP-tubulin. Time 0 refers to time of complete CFI. Representative of three biological replicates. Scale bar, 2  $\mu$ m (b, d, f).

proteoliposomes, moreover, membrane-embedded LEM2<sub>FL</sub> could still stabilize and bundle microtubules (Extended Data Fig. 2b). Finally, isolated LEM2<sub>NTD</sub> was sufficient to bundle microtubules (Fig. 2b–e, Extended Data Fig. 2c–f). We quantified microtubule bundling using light scattering and corroborated the assay using negative stain electron microscopy. LEM2<sub>NTD</sub>, which contains the LCD, bundled microtubules in a concentration-dependent and saturable manner, with a half-maximal scattering concentration ( $K_{\text{apparent}}$  value) of 1.3  $\mu$ M (Extended Data Fig. 2c, d). Electron microscopy revealed that LEM2-bundled microtubules lost their characteristic tubulin lattice fine structure, which appeared to be occluded by an amorphous LEM2 coating (Fig. 2e). Furthermore, LEM2<sub>NTD</sub> droplets could form coacervates with unpolymerized tubulin and form microtubule bundles upon addition of GTP and MgCl<sub>2</sub> (Extended Data Fig. 2e). These data demonstrate that the LEM2 LCD binds directly to microtubules to cause bundling in vitro.

Considering the liquid-like phase separation and complex coacervation properties of the LEM2 LCD, we investigated the state of microtubule-bound LEM2 in vitro. Using video fluorescence microscopy, we observed that both LEM2<sub>FL</sub> and LEM2<sub>NTD</sub> bound microtubule bundles dynamically, recovering fluorescence at similar rates after photobleaching, in contrast to tubulin in the microtubule lattice, which did not recover (Fig. 2c, Extended Data Fig. 2f). Notably, photobleached regions of LEM2<sub>NTD</sub> on microtubule bundles recovered fluorescence through coaxial flow, or a liquid-like gradient, rather than in a uniform recovery pattern across the bleached area (Fig. 2d,



Extended Data Fig. 2g, Supplementary Video 2). Thus, the LCD of LEM2 coats microtubules in a liquid-like phase.

We further identified two peptides within the PR-rich region that were sufficient to induce microtubule bundling in vitro (Fig. 2e, Extended Data Fig. 3a–c). In vivo, a LEM2 mutant in which this proline- and arginine-enriched sequence was deleted (LEM2<sub>ΔPR</sub>-mChr) initially targeted to the nascent nuclear envelope; however, in contrast to LEM2<sub>ΔSY</sub>-mChr, it did not enrich at the microtubule-containing core (Fig. 2f, Extended Data Fig. 3d). Thus, direct microtubule-binding by this region is required for core-localized LEM2 condensation during anaphase. Finally, and consistent with phosphoregulation, phosphomimetic mutations that spanned the microtubule-interacting PR-rich motif and its flanking region impaired microtubule bundling in vitro (Extended Data Fig. 3e, f).

## A molecular O-ring seals the nuclear envelope

As predicted by direct binding between the WH domain of LEM2 (LEM2<sub>WH</sub>; AA 395–503) and CHMP7<sup>4</sup>, LEM2<sub>ΔWH</sub>-mChr was unable to recruit GFP-CHMP7 to the nuclear envelope (Fig. 3a). Furthermore, following depletion of endogenous LEM2, overexpression of small inhibitory RNA (siRNA)-resistant LEM2<sub>ΔWH</sub>-mChr did not rescue the recruitment of IST1, an ESCRT-III protein downstream of LEM2 and CHMP7 that is required to recruit the microtubule-severing enzyme SPASTIN<sup>2</sup> (Extended Data Fig. 4a). Overexpression of siRNA-resistant LEM2-mChr led to premature and enhanced recruitment of endogenous IST1 to the nascent nuclear envelope (Extended Data Fig. 4b), which indicates that LEM2 guides both the spatial and temporal patterns of CHMP7-dependent activation of ESCRT.

Incubation of monomeric full-length CHMP7 with LEM2<sub>WH</sub>, but not with LEM2<sub>NTD</sub>, triggered the assembly of looping protein polymers

with an inner diameter of 50–100 nm (Fig. 3b, Extended Data Fig. 4c). To investigate the mechanism of CHMP7 autoinhibition and release by LEM2<sub>WH</sub>, we used quantitative cross-linking mass spectrometry (XL-MS) to isotopically distinguish between monomers and LEM2<sub>WH</sub>-induced polymers (Extended Data Fig. 5a, Supplementary Table 1). We identified 24 cross-links that were specific to monomeric CHMP7. Hybrid peptide mapping revealed that the N- and C-terminal regions of CHMP7 fold together, bringing the  $\alpha 1$ – $\alpha 3$  helices of the ESCRT-III domain of CHMP7 into proximity with its N-terminal WH domains (Fig. 3c, Extended Data Fig. 5, Supplementary Table 2). These interactions in the monodisperse CHMP7 sample were substantially reduced or absent in LEM2<sub>WH</sub>-induced CHMP7 polymers (Fig. 3c, Supplementary Tables 1, 2), consistent with a conformational change into the open and polymeric state as visualized by electron microscopy (Fig. 3b). In total, 19 cross-links were enriched in the LEM2<sub>WH</sub>-induced polymeric sample, mapping to interactions between LEM2<sub>WH</sub> and the  $\alpha 1$ – $\alpha 3$  helices of CHMP7. Polymer-specific cross-links between the N-terminal amines of neighbouring CHMP7 molecules, and between adjacent LEM2<sub>WH</sub> subunits, are consistent with copolymerization of both proteins (Fig. 3c, Extended Data Fig. 5). Consistent with the XL-MS findings, CHMP7 and LEM2<sub>WH</sub> copellet with a 1:1 stoichiometry, and mutations of conserved residues suggested by the hybrid peptides within the CHMP7  $\alpha 1$ – $\alpha 3$  helices impaired co-polymerization of LEM2<sub>WH</sub> and CHMP7 (Extended Data Fig. 6a–e).

Model membranes were also sufficient to trigger polymerization of full-length CHMP7 (Extended Data Fig. 6f). Two-dimensional alignment and image averaging of membrane-induced CHMP7 polymers revealed a repeating unit that comprised a continuous, polymeric strand studded with repeating perpendicular spikes (Fig. 3d, Extended Data Fig. 6g). The periodic dimensions of this polymer matched those of homologous structures for ‘open’ human

# Article

CHMP1B and yeast Snf7, suggesting that these structures comprise an ESCRT-III filament decorated with protruding ESCRT-II-like tandem WH domains<sup>28–30</sup> (Fig. 3d, e, Extended Data Figs. 5e, f, 6f, g). Consistent with this model, a truncated CHMP7 fragment comprised of only the ESCRT-III domain (CHMP7<sub>ESCRT-III</sub>; AA 229–453) spontaneously polymerized during purification into rings that lacked perpendicular spikes, but had otherwise comparable dimensions (Extended Data Fig. 6h). The spontaneous polymerization of CHMP7<sub>ESCRT-III</sub> further supports an autoinhibitory function for the N-terminal WH domains of CHMP7. In summary, the LEM2 WH domain activates autoinhibited CHMP7 monomers through a domain-replacement mechanism, and this triggers their coassembly.

To investigate the roles of the LEM2–CHMP7 partnership at the spindle–chromatin interface, we expressed LEM2 deletion mutants alongside an import cargo (NLS–3GFP) to assess nuclear compartmentalization during anaphase. Cells expressing siRNA-resistant LEM2–mChr first had measurable enrichment of NLS–3GFP in reforming nuclei shortly after localization of LEM2 to the core, when spindle microtubules still present barriers to sealing the nuclear membrane (Figs. 1a, 3f, Extended Data Figs. 1a, 7a, b). Cells that expressed LEM2<sub>ΔSY</sub>–mChr, which concentrated within the core normally, showed mild defects in compartmentalization (Fig. 2f, Extended Data Fig. 7a, b). By contrast, expression of LEM2<sub>ΔPP</sub>–mChr, which did not concentrate within the core, considerably attenuated accumulation of NLS–3GFP, pointing to the importance of the ability of LEM2 to specifically condense around spindle microtubules (Figs. 2f, 3f, Extended Data Fig. 7a, b). Finally, expression of LEM2<sub>ΔWH</sub>–mChr also prominently impaired timely accumulation of NLS–3GFP. This shows that the WH domain has a key role in compartmentalization, presumably owing to its ability to recruit CHMP7 and subsequent ESCRT pathway proteins (Figs. 1a, 3a, f, Extended Data Fig. 7a, b). Accordingly, depletion of either LEM2 or CHMP7 impaired nuclear compartmentalization for an extended period, with defects persisting for at least 30 min after complete ingression of the cleavage furrow (Extended Data Fig. 7c).

Extended siRNA-mediated depletion of LEM2 causes nuclear malformations and is eventually lethal<sup>19</sup>. To identify whether defects in nuclear morphology arise as a result of errors during mitotic exit, we synchronized cells following LEM2 depletion and imaged their progression from anaphase to late telophase. While LEM2-depleted cells progressed to anaphase without noticeable nuclear defects, a strong nuclear and tubulin morphology phenotype began to emerge in late anaphase and persisted through telophase (Extended Data Fig. 8a, Supplementary Video 3). Orthogonal views of LEM2-depleted cells revealed that aberrant microtubule bundles tunneled through the nucleus within a channel lined by nuclear envelope (Extended Data Fig. 8b). The appearance of DNA damage in telophase U2OS cells depleted of LEM2 underscores the important ramifications of these nuclear integrity phenotypes (Extended Data Fig. 8c).

Finally, to test the central hypothesis that the LCD and WH activities of LEM2 cooperate with CHMP7 to self-assemble around spindle microtubules, we engineered a simplified LEM2 construct bearing the extra-luminal domains of LEM2 connected by a flexible linker, LEM2<sub>NTD-LINKER-WH</sub>. When incubated with microtubules, LEM2<sub>NTD-LINKER-WH</sub> triggered copolymerization of CHMP7 and the resulting rings could be seen to loop upon and around microtubule bundles (Fig. 3g, Extended Data Fig. 8d). Consistently, and underscoring the importance of their functional coordination, disruption of the microtubule-interacting or CHMP7-binding domains of LEM2 led to persistent interphase nuclear morphology phenotypes in cells (Extended Data Fig. 8e–g).

Our results suggest the following model: LEM2 concentrates within the nuclear envelope core through direct microtubule binding, and its LCD condenses into a liquid-like phase that wets the surface of spindle microtubules. The WH domain of LEM2 and CHMP7 then copolymerize to complete the formation of a macromolecular ‘O-ring’ seal that establishes a temporary barrier between the new nucleoplasm and

the cytoplasm, before triggering complete spindle disassembly and membrane fusion (Fig. 3h, Supplementary Discussion).

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2232-x>.

- Ungrich, R. & Kutay, U. Mechanisms and functions of nuclear envelope remodelling. *Nat. Rev. Mol. Cell Biol.* **18**, 229–245 (2017).
- Vetri, M. et al. Spastin and ESCRT-III coordinate mitotic spindle disassembly and nuclear envelope sealing. *Nature* **522**, 231–235 (2015).
- Olmos, Y., Hodgson, L., Mantell, J., Verkade, P. & Carlton, J. G. ESCRT-III controls nuclear envelope reformation. *Nature* **522**, 236–239 (2015).
- Gu, M. et al. LEM2 recruits CHMP7 for ESCRT-mediated nuclear envelope closure in fission yeast and human cells. *Proc. Natl. Acad. Sci. USA* **114**, E2166–E2175 (2017).
- Haraguchi, T. et al. BAF is required for emerin assembly into the reforming nuclear envelope. *J. Cell Sci.* **114**, 4575–4585 (2001).
- Margalit, A., Segura-Totten, M., Gruenbaum, Y. & Wilson, K. L. Barrier-to-autointegration factor is required to segregate and enclose chromosomes within the nuclear envelope and assemble the nuclear lamina. *Proc. Natl. Acad. Sci. USA* **102**, 3290–3295 (2005).
- Herrada, I. et al. Purification and structural analysis of LEM-domain proteins. *Methods Enzymol.* **569**, 43–61 (2016).
- Barton, L. J., Soshnev, A. A. & Geyer, P. K. Networking in the nucleus: a spotlight on LEM-domain proteins. *Curr. Opin. Cell Biol.* **34**, 1–8 (2015).
- Raab, M. et al. ESCRT III repairs nuclear envelope ripples during cell migration to limit DNA damage and cell death. *Science* **352**, 359–362 (2016).
- Denais, C. M. et al. Nuclear envelope rupture and repair during cancer cell migration. *Science* **352**, 353–358 (2016).
- Halfmann, C. T. et al. Repair of nuclear ruptures requires barrier-to-autointegration factor. *J. Cell Biol.* **218**, 2136–2149 (2019).
- Thaller, D. J. et al. An ESCRT-LEM protein surveillance system is poised to directly monitor the nuclear envelope and nuclear transport system. *eLife* **8**, e45284 (2019).
- Penfield, L. et al. Dynein-pulling forces counteract lamin-mediated nuclear stability during nuclear envelope repair. *Mol. Biol. Cell* **29**, 852–868 (2018).
- Mekhail, K. & Moazed, D. The nuclear envelope in genome organization, expression and stability. *Nat. Rev. Mol. Cell Biol.* **11**, 317–328 (2010).
- Barrales, R. R., Forn, M., Georgescu, P. R., Sarkadi, Z. & Braun, S. Control of heterochromatin localization and silencing by the nuclear membrane protein Lem2. *Genes Dev.* **30**, 133–148 (2016).
- Ebrahimi, H., Masuda, H., Jain, D. & Cooper, J. P. Distinct ‘safe zones’ at the nuclear envelope ensure robust replication of heterochromatic chromosome regions. *eLife* **7**, e32911 (2018).
- Pieper, G. H., Sprenger, S., Teis, D. & Oliferenko, S. ESCRT-III/Vps4 controls heterochromatin-nuclear envelope attachments. *Dev. Cell* <https://doi.org/10.1016/j.devcel.2020.01.028> (2019).
- Brachner, A., Reipert, S., Foisner, R. & Gotzmann, J. LEM2 is a novel MAN1-related inner nuclear membrane protein associated with A-type lamins. *J. Cell Sci.* **118**, 5797–5810 (2005).
- Ulbert, S., Antonin, W., Platani, M. & Mattaj, J. W. The inner nuclear membrane protein Lem2 is critical for normal nuclear envelope morphology. *FEBS Lett.* **580**, 6435–6441 (2006).
- Lee, K. K. et al. Distinct functional domains in emerin bind lamin A and DNA-bridging protein BAF. *J. Cell Sci.* **114**, 4567–4573 (2001).
- Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
- Alberti, S., Gladfelter, A. & Mittag, T. Considerations and challenges in studying liquid–liquid phase separation and biomolecular condensates. *Cell* **176**, 419–434 (2019).
- Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).
- Rai, A. K., Chen, J. X., Selbach, M. & Pelkmans, L. Kinase-controlled phase transition of membraneless organelles in mitosis. *Nature* **559**, 211–216 (2018).
- Huguet, F., Flynn, S. & Vagnarelli, P. The role of phosphatases in nuclear envelope disassembly and reassembly and their relevance to pathologies. *Cells* **8**, E687 (2019).
- Hornbeck, P. V. et al. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43**, D512–D520 (2015).
- Boeynaems, S. et al. Spontaneous driving forces give rise to protein–RNA condensates with coexisting phases and complex material properties. *Proc. Natl. Acad. Sci. USA* **116**, 7889–7898 (2019).
- Im, Y. J., Wollert, T., Boura, E. & Hurley, J. H. Structure and function of the ESCRT-II–III interface in multivesicular body biogenesis. *Dev. Cell* **17**, 234–243 (2009).
- McCullough, J. et al. Structure and membrane remodeling activity of ESCRT-III helical polymers. *Science* **350**, 1548–1551 (2015).
- Tang, S. et al. Structural basis for activation, assembly and membrane binding of ESCRT-III Snf7 filaments. *eLife* **4**, e12548 (2015).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and, except where stated, investigators were not blinded to allocation during experiments and outcome assessment.

### Cytoskeletal proteins

Porcine brain tubulin reagents were purchased from Cytoskeleton, including unpolymerized tubulin (T238P), unpolymerized, HiLyte-647-labelled Tubulin (TL670M), and pre-formed microtubules (MT002).

### Purification of His<sub>6</sub>-SUMO-tagged proteins

All purified proteins in this study, except full length LEM2, were expressed in a pCA528 vector (WISP08-174; DNASU Plasmid Repository) in BL21-(DE3)-RIPL *Escherichia coli* cells using an N-terminal His<sub>6</sub>-SUMO affinity tag as described previously<sup>4</sup>. Full-length LEM2 was expressed in a pCA528 vector in C43(DE3) *E. coli* cells using an N-terminal His<sub>6</sub>-SUMO affinity tag. All plasmids are listed in Supplementary Table 3. Expression cultures (3- to 4-l) were grown in ZY-5052 rich auto-induction medium containing kanamycin, shaking (220 rpm) for 3 h at 37 °C, then overnight at 19 °C. Cells were collected by centrifugation.

### Purification of BAF

Purification of full length, human BAF protein (Uniprot ID O75531) was adapted from a published protocol<sup>31,32</sup>. The purification was performed as described, except His<sub>6</sub>-SUMO-BAF was cleaved with His<sub>6</sub>-Ulp1 protease (30 min, room temperature) to suit the use of the His<sub>6</sub>-SUMO affinity tag. Ultimately, Superdex 75 16/60 fractions containing BAF dimer were pooled, concentrated, and flash-frozen in liquid nitrogen and stored as single use aliquots at -80 °C.

### Purification of LEM2 proteins

Protein constructs originating from the human LEMD2 protein (Uniprot ID Q8NC56), including LEM2<sub>1-72</sub>, LEM2<sub>NTD(1-208)</sub>, LEM2<sub>Mim1(S82D,S90D,S96D,Y98D,T100D,Y104D)</sub>, LEM2<sub>Mim2(Y122D,S134D,S138D,S139D,T147D,S166D,S174D,S175D)</sub>, LEM2<sub>WH(395-503)</sub>, and LEM2<sub>NTD(1-208)-LINKER-WH(395-503)</sub> (linker sequence SAAGTGAGSGSAAS), were purified as follows. Collected cells were resuspended in lysis buffer (5 ml per gram cell pellet) containing 5% glycerol, 10 mM imidazole, pH 8.0, 2 µg/ml DNase1, lysozyme and protease inhibitors, with HEPES and KCl concentrations optimized individually for each LEM2 protein construct (20 mM HEPES, pH 7.0, 500 mM KCl for LEM2<sub>1-72</sub> and LEM2<sub>NTD-LINKER-WH</sub>, and 40 mM HEPES, pH 8.0, 350 mM KCl for LEM2<sub>WH</sub>, LEM2<sub>Mim1</sub> and LEM2<sub>Mim2</sub>). Cells were lysed by sonication on ice, and clarified (10,000g, 30 min, 4 °C). Clarified lysate was incubated with Ni-NTA agarose resin (Qiagen) (4 ml bed volume, 1.5 h, 4 °C), washed extensively with lysis buffer, and protein was eluted with five column volumes of lysis buffer supplemented with 350 mM imidazole, pH 8.0. Eluate was spin-concentrated (Vivaspin 20, 3 kDa MWCO, PES) to 5 ml, and dialysed overnight at 4 °C into storage buffer (20–40 mM HEPES, pH 7.0–8.0, 5% glycerol), with KCl concentrations optimized individually per construct (300 mM for LEM2<sub>1-72</sub>, 500 mM for LEM2<sub>NTD-LINKER-WH</sub>, 350 mM for LEM2<sub>WH</sub> and LEM2<sub>Mim1</sub>, 110 mM for LEM2<sub>Mim2</sub>). The His<sub>6</sub>-SUMO tag was cleaved by incubation with His<sub>6</sub>-Ulp1 protease during dialysis. LEM2 proteins were further purified by size exclusion chromatography using the Superdex 75 16/60 column (GE Life Sciences) in storage buffer, and LEM2-containing fractions were pooled, spin-concentrated (Vivaspin 6 3 kDa MWCO, PES), and aliquoted.

LEM2<sub>NTD</sub> was purified similarly, including an additional preparative phase separation step. All steps were carried out at room temperature unless otherwise noted. Harvested cells were resuspended in lysis buffer (30 mM HEPES, pH 7.4, 500 mM KCl, 5% glycerol, 10 mM imidazole, pH 8.0, 2 µg/ml DNase1, lysozyme and protease inhibitors), and lysed by sonication with intermittent chilling on ice. Lysate was

clarified by centrifugation (10,000g, 30 min), incubated with Ni-NTA agarose resin (4 ml bed volume, 1.5 h, room temperature), and washed extensively with lysis buffer. Protein was eluted with lysis buffer supplemented with 350 mM imidazole, pH 8.0. Droplet formation was induced to enrich for His<sub>6</sub>-SUMO-LEM2<sub>NTD</sub>: the imidazole eluate was diluted with ice-cold buffer (40 mM HEPES, pH 7.4, 5% glycerol) to drop the concentration of KCl to 50 mM. After 20 min incubation on ice, protein droplets were pelleted by centrifugation (10,000g, 10 min, 0 °C), washed with two column volumes of ice-cold low-salt buffer (40 mM HEPES, pH 7.4, 50 mM KCl, 5% glycerol), and pelleted again. The pellet was resuspended in 5 ml high-salt buffer (40 mM HEPES, pH 7.4, 500 mM KCl, 5% glycerol), to dissolve protein droplets for the remaining purification steps, and incubated with His<sub>6</sub>-Ulp1 protease (2 h). The mixture was incubated with Ni-NTA resin (2 ml, 1 h). Spin-concentrated, flow-through protein was further purified by size exclusion chromatography using the Superdex 75 16/60 column in high-salt buffer at 4 °C. Pooled, spin-concentrated, LEM2-containing fractions were dialysed overnight into storage buffer (40 mM HEPES, pH 7.4, 350 mM KCl, 5% glycerol), spin-concentrated, and aliquoted.

Concentrated protein aliquots were snap-frozen in liquid nitrogen and stored at -80 °C for future experiments. Aliquots were thawed and used only once. Final protein concentrations were 260 µM LEM2<sub>1-72</sub>, ~110 µM LEM2<sub>NTD</sub>, 30 µM LEM2<sub>NTD-LINKER-WH</sub>, 14 µM LEM2<sub>WH</sub>, 190 µM His<sub>6</sub>-SUMO-LEM2<sub>WH</sub>, 100 µM LEM2<sub>Mim1</sub>, and 110 µM LEM2<sub>Mim2</sub>. Note that concentrations of LEM2<sub>WH</sub> without the tag were limited, and LEM2<sub>WH</sub> was more stable with the His<sub>6</sub>-SUMO tag.

Human, full-length LEM2 (LEM2<sub>FL</sub>) was purified as follows. Collected cells (6-l expression culture) were resuspended in lysis buffer (5 ml per gram cell pellet) containing 25 mM HEPES, pH 7.4, 500 mM KCl, 1 mM DTT, 5% glycerol, lysozyme, 2 µg/ml DNase1, and protease inhibitors. Cells were lysed by sonication on ice and clarified (10,000g, 30 min, 4 °C). Membranes were pelleted and collected from clarified lysate by ultracentrifugation (100,000g, 1 h, 4 °C). LEM2<sub>FL</sub> was extracted from membranes with 1% (w/v) n-dodecyl-B-D-maltopyranoside (DDM, Anatrace D310) with stirring (1 h, 4 °C), and collected in the supernatant following centrifugation (35,000g, 30 min, 4 °C). DDM-extracted supernatant was supplemented with 10 mM imidazole and incubated with Ni-NTA agarose resin (Qiagen) (4 ml bed volume, 1 h, 4 °C), washed extensively with lysis buffer containing 0.1% DDM, and protein was eluted with five column volumes of lysis buffer containing 0.1% DDM and supplemented with 500 mM imidazole, pH 8.0. LEM2<sub>FL</sub>-rich elution fractions were pooled and dialysed overnight at 4 °C to remove imidazole. The His<sub>6</sub>-SUMO tag was cleaved by incubation with His<sub>6</sub>-Ulp1 protease (30 min, room temp) and removed by reverse Ni-NTA chromatography. Cleaved, DDM-solubilized LEM2<sub>FL</sub> was used unconcentrated, or spin-column concentrated (Vivaspin20, 100 kDa MWCO, PES) without freeze-thaw.

### Purification of CHMP7 proteins

Purification procedures for full-length, human CHMP7 (Uniprot ID Q8WUX9) and point mutants CHMP7<sub>R270E+I312E</sub>, CHMP7<sub>R270E+V339D</sub>, and CHMP7<sub>I312E+V339D</sub> were adapted from previously published purifications<sup>4</sup>. The truncated CHMP7<sub>ESCRT-III(229-453)</sub> was purified as follows. Collected cells were resuspended in 5 ml ice-cold lysis buffer (40 mM HEPES, pH 8.0, 800 mM KCl, 5% glycerol, 10 mM imidazole, 2 µg/ml DNase1, 5 mM BME, protease inhibitors, and lysozyme) per gram cell pellet. Cells were lysed by sonication, on ice, and clarified (10,000g, 30 min, room temperature). Clarified CHMP7<sub>ESCRT-III</sub> lysate was incubated with Ni-NTA agarose resin and spontaneously formed a gel composed of protein polymerized into rings, assayed by negative stain electron microscopy. His<sub>6</sub>-SUMO-CHMP7<sub>ESCRT-III</sub> rings were eluted with imidazole (350 mM), and the eluate was collected with low-speed spin (1,000g) as a gel phase on top of the resin. The gel was scooped off and washed three times with buffer (40 mM HEPES, pH 8.0, 800 mM KCl, 5% glycerol, 1 mM DTT), and polymers were collected by centrifugation (5,000g) each time.

# Article

The His<sub>6</sub>-SUMO tag was cleaved by incubation with His<sub>6</sub>-Ulp1 protease (2 h, 4 °C). Cleaved CHMP7<sub>ESCR7-III</sub> was washed three times with buffer and collected by centrifugation each time, and soluble His<sub>6</sub>-SUMO and His<sub>6</sub>-Ulp1 were discarded in the aqueous supernatant. Cleaved eluate was diluted with buffer to give a final protein concentration of about 60 μM.

## Analytical size exclusion chromatography

Binding of LEM2<sub>1-72</sub> to BAF or BAF–dsDNA complex was assayed by gel filtration chromatography. DNA duplex was prepared as previously described<sup>32</sup>. Combinations of purified LEM2<sub>1-72</sub>, BAF dimers, and DNA duplexes were combined in 1:1:2 molar ratio. Following incubation at room temperature for 30 min, 50 μl reaction mixture was applied to a Superose 6 3.2/300 column (GE Life Sciences) in buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 5 mM DTT, and 10% glycerol) equilibrated at 4 °C. The flow rate was 40 μl/min for all experiments. Retention volumes for major peaks absorbing at 280 nm (A280) were recorded. The protein contents of peak fractions were assayed by SDS-PAGE.

## LEM2 LCD peptides

Chemically synthesized peptides bearing an N-terminal FITC-Ahx modification (Extended Data Fig. 3a) were purchased from GenScript, peptide stock solutions were prepared in milli-Q H<sub>2</sub>O except for LEM2<sub>75-122</sub> stocks, which were prepared in DMSO. Phase separation of LEM2<sub>75-122</sub> was induced by dilution in milli-Q H<sub>2</sub>O or buffer (25 mM HEPES, pH 7.4, 150 mM KCl) to 0.2 mg/ml stock.

## Turbidimetry

Microtubule bundling by LEM2 was quantified by turbidity (absorbance at 340 nm) (Tecan Spark 10M spectrophotometer)<sup>33</sup>. Reactions (10 μl total volume) of purified LEM2<sub>(1-72, NTD, WH, Mim1, or Mim2)</sub> protein or chemically synthesized LEM2 peptides, 1 μM αβ-tubulin heterodimers polymerized into microtubules, KCl (75 mM for LEM2 protein reactions, 0.9 mM for LEM2 peptide reactions), 25 mM HEPES, pH 7.4, 0.5 mM MgCl<sub>2</sub>, and 10 μM paclitaxel (Sigma Aldrich), were prepared at room temperature in 384-well non-binding plates (Grendier Bio-One, #781906). Turbidity was measured, with shaking before each read, for reactions in triplicate, and averaged for each condition. Turbidity values for reactions of LEM2 protein/peptide with microtubules were corrected for the turbidity of LEM2 protein/peptide alone, normalized to turbidity of microtubules alone, and plotted (Turbidity a.u.) against [LEM2]. Sigmoidal curve fitting and half-maximum LEM2 concentrations were calculated with GraphPad Prism software. The turbidity of LEM2<sub>Mim1</sub> and LEM2<sub>Mim2</sub> was normalized to a saturating turbidity value equal to 1.0, for easy comparison of half-maximal values within plot. Turbidity data were corroborated by complementary negative stain electron microscopy. Turbidimetry was not performed for peptides LEM2<sub>75-122</sub> (LEM2<sub>SY</sub>) or LEM2<sub>145-165</sub> (LEM2-PR<sub>A</sub>) with microtubules, because these peptides were already turbid in the absence of microtubules.

## Protein labelling for fluorescence imaging

Proteins were fluorescently labelled for microscopy using Microscale Protein Labelling kits from Thermo Scientific: Alexa Fluor 488 (A30006), Alexa Fluor 555 (A30007) and Alexa Fluor 647 (A30009). For imaging experiments, unlabelled protein was supplemented with a minimal amount of fluorescently labelled protein. LEM2<sub>NTD</sub> contains a single lysine residue and could not be efficiently labelled with primary amine-reactive probes. Instead, His<sub>6</sub>-SUMO-LEM2<sub>NTD</sub> was labelled with Alexa-488 for fluorescence microscopy and a minimal amount was used to supplement indicated concentrations of unlabelled LEM2<sub>NTD</sub> for imaging.

## Fluorescence microscopy phase separation, coacervation and microtubule bundling assays

Fluorescently labelled proteins were assayed for phase separation, coacervation and microtubule bundling by spinning disc confocal

microscopy, performed in glass-bottom 384-well plates (Greiner Bio-One #7781892). Reactions for confocal microscopy were prepared in tubes and transferred to plates for imaging. LEM2<sub>NTD</sub>, LEM2<sub>Mim1</sub>, LEM2<sub>Mim2</sub>, and LEM2<sub>75-122</sub> droplets were allowed to settle for 15 min before imaging. LEM2<sub>NTD</sub> droplet formation was induced at room temperature by reducing the salt concentration to 150 mM KCl, keeping the concentrations of other buffer components the same; droplets were allowed to grow for 60 min before imaging. Binding of LEM2<sub>NTD</sub> droplets (8 to 24 μM) or detergent-solubilized LEM2<sub>FL</sub> (8 μM) to microtubules was tested with the addition of fluorescently labelled microtubules (1 μM αβ-tubulin polymerized into microtubules). Microtubules were prepared following the Cytoskeleton protocol, in a molar ratio of 1:7 HiLyte647-tubulin:unlabelled tubulin, in G-PEM buffer (Cytoskeleton BST01). Unpolymerized tubulin, 1:7 labelled:unlabelled molar ratio, and Alexa555-labelled LEM2<sub>WH</sub> were tested for association with LEM2<sub>NTD</sub> droplets at concentrations of 8 μM each. Note that 150 mM KCl promotes microtubule depolymerization in the absence of stabilizing proteins. To assay coacervate formation, LEM2<sub>NTD</sub> (8 μM) was fluorescently imaged directly after reducing the salt concentration for LEM2<sub>NTD</sub> alone or in the presence of 1 μM unpolymerized αβ-tubulin, 30 bases RNA (5'-GGGCCUCGCCTCCAGUGAGGGGCCGGCCC-3'), or 50 bases DNA (5'-AATGTATTGGTGGGGCTGCTCGGGATTGCGGATACGCCCTTGGATTCGC-3').

## Spinning disc confocal microscopy with FRAP

To assay LEM2<sub>NTD</sub> droplet dynamics in FRAP and droplet fusion experiments, 384-well glass-bottom plates were PEGylated using PEG-Silane (LAYSAN BIO, MPEG-SIL-5000) followed by passivation with BSA, according to published protocols, to prevent droplet association to the silica surface<sup>34</sup>. Spinning disc confocal microscopy with FRAP was carried out using a W1-CSU with a Borealis upgrade (Andor) and ILE Laser Launch (laser lines 405 nm, 488 nm, 561 nm, 647 nm; Andor) on a Nikon Eclipse Ti microscope (Nikon Instruments) equipped with a Rapp Optoelectronic UGA-40 (Rapp OptoElectronic) photobleaching unit and operated with MicroManager 2.0beta (Open Imaging). Fluorescence images were collected with an Andor Zyla 4.2 cMOS Camera (Oxford Instruments). Samples were imaged using a CFI Plan Apochromat VC 100× Oil NA 1.4 objective (Nikon Instruments). Excitation wavelengths were: 488 nm for Alex488TM or FITC; 561 nm for Alex555TM; 640 nm for Alex-647. Photobleaching for FRAP experiments was done with a 473-nm laser (Vortran) for 200 ms per bleach event. Images were recorded with a frequency of 1 Hz, starting with at least one image recorded pre-bleach, including the bleach event, and up to 240 s post-bleaching. The ratio of the intensity within bleached regions of interest to background was calculated with ImageJ for each time point and normalized to the intensity at the time of bleaching. Replicates were aligned to the time of bleaching, averaged, and plotted ± s.d.

## Negative stain electron microscopy

Continuous carbon grids (200–400 mesh copper, Quantifoil) were glow-discharged (PELCO EasiGlow, 15 mA, 0.39 mBar, 30 s). Samples (3–5 μl) were stained with 0.75% uranyl formate as described previously<sup>29</sup>. Images were collected with a Tecnai T12 microscope (FEI) with a LaB6 filament, operated at 120 kV, and data were captured with a Gatan Ultrascan CCD camera (Gatan). For microtubule-binding assays, reactions were prepared on grids and incubated for 2 min. Microtubules were used at a concentration of 1 μM αβ-tubulin in pre-formed microtubules. LEM2 constructs were imaged with microtubules (Fig. 2f) at the following concentrations: 6 μM LEM2<sub>FL</sub>, 4 μM LEM2<sub>NTD</sub>, 200 μM LEM2<sub>145-165</sub> (LEM2-PR<sub>A</sub>), 80 μM LEM2<sub>188-212</sub> (LEM2-PR<sub>B</sub>). Reactions of microtubules with LEM2<sub>NTD</sub> or LEM2<sub>188-212</sub>, to corroborate turbidimetry, were prepared as described above. Phosphomimetic constructs, LEM2<sub>Mim1</sub> and LEM2<sub>Mim2</sub>, were tested for microtubule bundling at 8 μM.

### LEM2-induced CHMP7 polymerization

LEM2 and CHMP7 proteins were mixed at concentrations of 4–8  $\mu$ M in a total volume of 50  $\mu$ L. LEM2<sub>WH</sub> or LEM2<sub>NTD-Linker-WH</sub> was present in twofold molar excess to CHMP7 proteins, unless otherwise stated. The buffer was adjusted to between 600 and 800 mM KCl. Reactions were dialysed for 9–12 h into low-salt buffer (30 mM HEPES, pH 8.0, 25 mM KCl, and 1 mM DTT) using Slide-A-Lyzer MINI Dialysis Device, 10K MWCO (Thermo Fisher Scientific). For experiments including microtubules, a 3 $\times$  stock of pre-formed, lyophilized microtubules was prepared in buffer (30 mM HEPES, pH 8.0, 800 mM KCl, 1 mM DTT, 60  $\mu$ M paclitaxel, and 3 mM MgCl<sub>2</sub>), and dialysis buffer was supplemented with 10  $\mu$ M paclitaxel and 1 mM MgCl<sub>2</sub>. After dialysis, polymeric assemblies were concentrated by low speed centrifugation (5,000g, 5 min) and resuspended in 30  $\mu$ L buffer for negative stain electron microscopy, SDS-PAGE, or cross-linking mass spectrometry analysis. To determine the stoichiometry of the CHMP7–LEM2<sub>WH</sub> polymer, CHMP7 (4  $\mu$ M) was titrated with LEM2<sub>WH</sub> (0, 0.08, 1 and 4  $\mu$ M), dialysed, and polymers were collected by low-speed centrifugation. The supernatant was collected such that 10  $\mu$ L was left behind as the pellet fraction. Quantities of polymerized and unpolymerized CHMP7 at each LEM2 condition were measured by Coomassie-stained SDS-PAGE, in which equal volumes of pellet and supernatant fractions were loaded. Intensities were quantified by ImageJ. The data was normalized to samples containing 0 and equimolar amounts of LEM2<sub>WH</sub>.

### Liposome preparation

Lipid solutions (Avanti Polar Lipids) were resuspended in chloroform, and liposomes were prepared as previously described<sup>29</sup>. In brief, lipids (2 mg total) were dried in a glass vial to give final ratio (mole %) of 30% egg phosphatidylserine (PS), 30% egg phosphatidylcholine (PC) and 40% phosphatidylethanolamine (PE). Lipid films were dispersed in buffer (40 mM HEPES, pH 8.0, 150 mM KCl) to produce liposomes at a final concentration of 1 mg/ml, and stored at –80 °C.

### Microtubule bundling assay of full-length LEM2 in the presence of model membranes

A dried lipid film in a glass vial (1 mg total containing final ratio (mole %) 60% egg PC, 24.7% PE, 10% bovine liver phosphatidylinositol, 5% cholesterol, and 0.3% 18:1 Liss Rhod PE from Avanti Polar Lipids) was dispersed in buffer (40 mM HEPES, pH 8.0, 800 mM KCl, 5% glycerol, 0.1% DDM) and sonicated at 30 °C for 10 min to produce detergent-solubilized lipids at a concentration of 1 mg/ml. Equal volumes detergent-solubilized lipids, activated Bio-Beads SM-2 Resin (BioRad), and concentrated LEM2<sub>FL</sub> (1 mg/ml) or LEM2<sub>FL</sub> buffer were combined and incubated overnight at 4 °C. Detergent-depleted reactions were separated from resin and mixed with fluorescently labelled microtubule and G-PEM buffer (Cytoskeleton BST01) to a final KCl concentration of 150 mM and 1  $\mu$ M tubulin. Reactions were incubated for 1 h at room temperature before imaging.

### Membrane-induced CHMP7 polymerization

CHMP7 was dialysed or diluted to reduce the salt concentration from 800 mM to 150 mM KCl (supplemented with 5% glycerol, 40 mM HEPES, pH 8.0, 1 mM DTT) to give a final protein concentration of 0.1 mg/ml. Liposomes and CHMP7 were mixed 40:1 (w/w) and negative stain electron microscopy grids were prepared immediately.

### Electron microscopy data acquisition and 2D classification

Membrane-induced CHMP7 polymers were prepared for negative stain electron microscopy and imaged with a Tecnai T20 microscope (FEI) with a LaB6 filament, operated at 200 kV. 227 micrographs were collected with a TemCam-F816 8k  $\times$  8k camera (TVIPS) using SerialEM software<sup>35</sup>, with a nominal pixel size of 1.57 Å. The defocus was 0.7–1.7  $\mu$ m and the total dose was 20 e<sup>–</sup>/Å<sup>2</sup>. Particles containing a

repeating polymeric unit were picked manually along the polymeric protein chain, yielding 6,094 particles. Specifically, particles were picked from polymers detached from membrane, which were in a favourable orientation for subsequent classification. Particles were picked and 2D-classified using default parameters within RELION version 2.0 software.

### Cross-linking mass spectrometry

Full length CHMP7 (60  $\mu$ g) was polymerized with equimolar His<sub>6</sub>-SUMO–LEM2<sub>WH</sub>, following the described polymerization assay, and crosslinked with 2 mM light crosslinker (H<sub>12</sub>-BS3, Creative Molecules) for 30 min at 30 °C. His<sub>6</sub>-SUMO–LEM2<sub>WH</sub> was used to achieve a higher protein concentration. Full-length, monomeric CHMP7 (60  $\mu$ g) was crosslinked with 2 mM heavy crosslinker (D<sub>12</sub>-BS3, Creative Molecules) under the same conditions. Reactions were quenched (10 mM ammonium bicarbonate, 10 min, room temperature), and light and heavy crosslinked reaction mixtures were combined and processed for mass spectrometry as described previously<sup>36,37</sup>. Crosslinked products were enriched by size-exclusion chromatography (Superdex Peptide, GE Healthcare Life Sciences) as described previously<sup>36</sup> and fractions eluting between 0.9 and 1.4 ml were dried and resuspended in 0.1% formic acid for MS analysis. The fractions starting at 0.9 ml and 1.3 ml were combined before evaporation to make four total SEC fractions.

LC-MS analysis was performed with a Q-Exactive Plus mass spectrometer (Thermo Scientific) coupled with a nanoelectrospray ion source (Easy-Spray, Thermo) and NanoAcquity UPLC system (Waters). Enriched fractions were separated on a 15 cm  $\times$  75  $\mu$ m ID PepMap C18 column (Thermo) using a 60-min gradient from 5–28% solvent B (A: 0.1% formic acid in water, B: 0.1% formic acid in acetonitrile). Precursor MS scans were measured in the Orbitrap scanning from 350 to 1,500 m/z (mass resolution: 70,000). The ten most intense triply charged or higher precursors were isolated in the quadrupole (isolation window: 4 m/z) and dissociated by HCD (normalized collision energy: 24), and the product ion spectra were measured in the Orbitrap (mass resolution: 17,500). A dynamic exclusion window of 20 s was applied and the automatic gain control targets were set to  $3 \times 10^6$  (precursor scan) and  $5 \times 10^4$  (product scan).

Peaklists were generated using Proteome Discoverer 2.2 (Thermo), and crosslinked peptides searched for with Protein Prospector 5.20.23<sup>38</sup>. 85 of the most intense peaks from each product ion spectrum were searched against a database containing His<sub>6</sub>-SUMO–LEM2<sub>WH</sub> along with the sequences of 10 other proteins comprising CHMP subunits and tubulin, concatenated with 10 randomized decoy versions of each of these sequences (121 sequences total). Search parameters were: mass tolerance of 20 ppm (precursor) and 30 ppm (product); fixed modifications of carbamidomethylation on cysteine; variable modifications of peptide N-terminal glutamine conversion to pyroglutamate, oxidation of methionine, and ‘dead-end’ modification of lysine and the protein N terminus by semi-hydrolysed heavy and light BS3; trypsin specificity was used with two missed cleavages and three variable modifications per peptide were allowed. Unique, crosslinked residue pairs were reported at a 1.5% FDR threshold, corresponding to a score difference cutoff of 15.

For quantitative analysis, precursor ion filtering in Skyline 4.1<sup>39</sup> was used to extract light:heavy crosslinked precursor ion signals. Skyline does not directly support crosslinking analysis, so crosslinked peptides were linearized and exported as a spectral library as previously described<sup>40</sup>. Transitions were generated for every light or heavy crosslinked peptide species discovered in the Protein Prospector search and paired with its corresponding heavy or light counterpart. Precursor ion transitions matching the first three isotopes were extracted across all four LC-MS fractions. Each extracted ion chromatogram was manually inspected and the start and end points were adjusted to ensure that the correct peaks were detected and that there were no interfering signals. Isotope dot products were required to exceed 0.85.

# Article

The peak areas were re-imported into R and summarized at the level of crosslinked residues for each light and heavy crosslink. Peak areas were summed for all peptides matching a given crosslink. Finally,  $\log_2$  ratios of the heavy to light peak areas were determined. Filtered cross-links were mapped to the primary protein structure using xiNET<sup>41</sup>.

## Homology modelling and cross-link mapping

Homology models of human CHMP7 and LEM2 domains were created with Phyre2<sup>42</sup> (Supplementary Table 2). Reference structures were selected based on confidence scores and homology to reference structure. Models were validated by mapping cross-linking data to the models using UCSF Chimera<sup>43</sup> together with Xlink analyser<sup>44</sup>.

## Immunostaining for fluorescence microscopy

Cells were fixed at room temperature in 2% paraformaldehyde for 25 min. The primary antibodies used for immunodetection were rabbit anti-IST1<sup>45</sup>, rabbit anti-tubulin (ab18251; abcam), rat anti-tubulin (YL1/2; Accurate Chemical & Scientific), mouse anti-lamin B2 (AbCam; ab8983), mouse anti-SUN2 (gift from B. Burke), and mouse anti-53BP1 (MAB3802; Millipore). After incubation with fluorescently labelled secondary antibodies (anti-rabbit 488, anti-mouse 488, anti-rabbit 647, and anti-rabbit 568; Thermo Fisher), coverslips were mounted with DAPI ProLong Gold (Thermo Fisher) and imaged by widefield microscopy (Zeiss Axioskop 2; Axiovision software), spinning disc confocal microscopy (Nikon Eclipse Ti; Metamorph software), or stimulated emission depletion microscopy (see below). For quantifying IST1 localization: anti-rabbit 488 against rabbit anti-IST1; for assessing DNA damage in telophase cells: anti-mouse 488 against mouse anti-53BP1 and anti-rabbit 568 against rabbit anti-tubulin; for quantifying and imaging the tubulin and nuclear envelope phenotype in telophase cells: anti-mouse 488 against mouse anti-SUN2 (INM marker for orthogonal imaging) or mouse anti-lamin B2 (INM marker for quantitative experiments), and anti-rabbit 568 against rabbit anti-tubulin; for interphase phenotypes: anti-rabbit 488 against rabbit anti-tubulin; for STED: anti-mouse 568 against mouse anti-tubulin and anti-rabbit 647 against rabbit anti-LEM2.

## STED microscopy

STED microscopy was performed on a Leica TCS SP8 STED 3× confocal laser-scanning microscope equipped with a HC PL APO CS2100×/1.40 oil objective. Confocal sections were imaged with Leica LAS X Core software and processed with Huygens Software Suite (SVI). Images were recorded using a 405-nm laser line at 1.4% laser power to image DAPI, and a 572-nm laser line at 5.6% laser power to image tubulin in confocal detector mode. LEM2 was imaged with a 653-nm laser line at 2.5% laser power in STED pulsed detector mode (gate start at 0.3 ns and gate end at 6.5 ns) with a Huygens saturation factor of 5.7. Deconvolved images were contrast enhanced in ImageJ (NIH); raw data are available upon request.

## Quantification from fluorescence microscopy

For illustration, images of anaphase A and B cells were acquired by widefield microscopy at 100× and adjusted so that background fluorescence in the DNA, IST1, and LEM2-mChr channels were comparable between samples. Raw images acquired by widefield microscopy at 100× were used to score the IST1 phenotype in anaphase A (early) and anaphase B (late). IST1 localization to anaphase chromatin masses was assessed in three independent experiments in which images were coded, randomized and scored blindly by three independent scorers. Images were decoded and blind scores were quantified. Each chromatin disc (two per cell) was scored as having extra robust, robust, weak, or no recruitment for IST1. Robust recruitment was characterized by distinctive foci organized at the core of chromatin masses, whereas weak recruitment was characterized by less intense, often fewer, and less organized foci at the chromatin surface, consistent with what has previously been shown<sup>4</sup>. Extra robust was characterized by strikingly

intense IST1 fluorescence, often accompanied by recruitment over most of the chromatin disc. For clarity, the robust and extra robust categories were graphed together as a single category. The majority score was used in cases where the three scores differed.

For time-lapse colocalization experiments tracking GFP-CHMP7, images of anaphase cells were acquired by time-lapse light microscopy (described below) and were selected for scoring at the time of peak LEM2-mCherry or LEM2<sub>ΔWH</sub>-mCherry enrichment at the core region of anaphase chromatin discs. In FIJI, each cell was thresholded for either LEM2-mCherry or LEM2<sub>ΔWH</sub>-mCherry enrichment. The mean fluorescence (arbitrary units) was measured for each region of interest in the GFP-CHMP7 channel. The regions of interest were subtracted from the area of the whole cell to measure the mean fluorescence of cytoplasmic GFP-CHMP7 for each anaphase cell. The plotted values are the mean GFP-CHMP7 fluorescence of the region of interest, as determined by exogenous LEM2 enrichment, normalized to the mean fluorescence of cytoplasmic GFP-CHMP7.

Nuclear accumulation of NLS-3GFP in either late anaphase or telophase was determined by taking the ratio of nuclear to cytoplasmic NLS-3GFP at either 1-min intervals throughout anaphase (after imaging at 15 s intervals) or 30 min after complete cleavage furrow ingression, in telophase (time-lapse microscopy described below). In FIJI, regions of chromatin were defined as regions of interest in the NucBlue channel and used to measure mean fluorescence (arbitrary units) of nuclear NLS-3GFP in late anaphase or the integrated density (the product of area and mean grey value) in telophase cells. Cytoplasmic levels of NLS-3GFP were determined by selecting the whole cell, and subsequently deselecting the regions of chromatin. For anaphase cells, the mean fluorescence of nuclear NLS-3GFP for each disc was then divided by the mean fluorescence of cytoplasmic NLS-3GFP of the same cell. For telophase, the integrated density was used to account for changes in nuclear size due to decondensation.

For the purposes of illustration, interphase and telophase cells were acquired by widefield and time-lapse microscopy and were adjusted so that background fluorescence was comparable between samples. Spinning disc confocal microscopy at 60× was used to acquire z-stacks for orthographic projection. Raw images used to score the nuclear envelope and tubulin phenotype in telophase were acquired by widefield microscopy at 60×. To score DNA damage in telophase, images were acquired by widefield microscopy at 60× and thresholded uniformly in FIJI. Nuclear foci were detected using the find maxima function in FIJI and noise tolerance was held constant for all conditions. To score interphase nuclear circularity, images were acquired by widefield microscopy at 60×. The DNA channel was thresholded in FIJI and each nucleus was defined as a region of interest. The nuclei were then assessed for their circularity (circularity =  $4\pi(\text{area}/\text{perimeter}^2)$ ), as described by others<sup>46</sup>.

## Time-lapse light microscopy analysis

All live imaging experiments were imaged at 60× by spinning disc confocal microscopy using glass-bottomed Mat-Tek dishes. For experiments tracking chromatin, cells were treated with NucBlue (Thermo Fisher) 1 h before imaging, per the manufacturer's instructions. Stable cell lines and plasmids for transient transfection used in this study are described in Supplementary Table 3. Cells stably expressing GFP-lamin B2 and LEM2-mCherry were cultured for 48 h and treated with 200 nM SiR-tubulin (CY-SC006; Cytoskelton) 1 h before imaging. Cells stably expressing GFP-BAF and LEM2-mCherry were also cultured asynchronously. All other experiments were subjected to cell-cycle synchronization, as described below. To track nuclear integrity, cells stably expressing NLS-3GFP (PL19) alone or in combination with the stable expression of siRNA-resistant LEM2-mCherry (PL13), LEM2<sub>ΔSY</sub>-mChr (PL16), LEM2<sub>ΔPR</sub>-mChr (PL17), or LEM2<sub>ΔWH</sub>-mCherry (PL18) were imaged at either 15-s intervals (to assess anaphase compartmentalization) or 3-min intervals for at least 30 min following complete cleavage furrow

ingression (to assess telophase compartmentalization). In the latter experiments, 100 nM SiR-tubulin (CY-SC006; Cytoskelton) was added 1 h before imaging. HeLa cells stably expressing H2B-mCherry and GFP-tubulin were treated with siRNA oligos, as described below. For transient expression experiments, cells were transfected with plasmids using Lipofectamine LTX Plus (Thermo Fisher) for 24 h, then re-plated and subjected to cell-cycle synchronization. Cells stably expressing GFP-tubulin<sup>47</sup> were transiently transfected with siRNA-resistant pCMV(Δ5)-LEM2-mCherry (PL13), pCMV(Δ5)-LEM2<sub>m21</sub>-mCherry (PL14), pCMV(Δ5)-LEM2<sub>ALCD</sub>-mCherry (PL15), pCMV(Δ5)-LEM2<sub>ΔSY</sub>-mCherry (PL16), pCMV(Δ5)-LEM2<sub>ΔPR</sub>-mCherry (PL17), or pCMV(Δ5)-LEM2<sub>ΔWH</sub>-mCherry (PL18). For time-lapse colocalization experiments, HeLa cells stably expressing GFP-CHMP7 were transiently transfected with either siRNA-resistant pCMV(Δ5)-LEM2-mCherry (PL13) or siRNA-resistant pCMV(Δ5)-LEM2<sub>ΔWH</sub>-mCherry (PL18).

#### siRNA-mediated depletion and cell-cycle synchronization

HeLa cells were a gift from M. Powers and U2OS cells were a gift from D. Ayer; both cell lines were authenticated by STR profiling. All HeLa cell lines were generated from the same parental HeLa cells, which tested negative for mycoplasma contamination. U2OS cells were not tested for mycoplasma contamination. HeLa and U2OS cells were plated on fibronectin-coated coverslips in the presence of 10 nM siRNA oligo, delivered by Lipofectamine RNAiMAX Transfection Reagent (Thermo Fisher). Specific sequences used were: siControl (siScr-1)<sup>47,48</sup>, siLEM2-1 (antisense sequence targeting nucleotides 78–98: UUGCGGUAGACAUCCCGGGdTdT<sup>19</sup>), siLEM2-2 (antisense sequence targeting nucleotides 1,297–1,317: UACAUUAUGGAUAGCGCUCCdTdT<sup>19</sup>), and siCHMP7 (antisense sequence targeting nucleotides 650–669: CCCUCUUACACUCAAdTdT)<sup>49</sup>. In experiments testing whether exogenous LEM2 could rescue function after endogenous LEM2 depletion, siLEM2-2 was used to deplete endogenous LEM2 as exogenous LEM2 constructs harbour silent mutations that confer resistance to the siLEM2-2 oligo. For cell-cycle synchronization experiments, 6–8 h after plating, 2 mM thymidine was added for 24 h to arrest cells at G1/S. Cells were then rinsed thoroughly with PBS, followed by the addition of culture medium. Twelve hours after release, cells were imaged live or fixed for microscopy. For experiments tracking interphase phenotypes, cells were fixed for immunostaining sixteen hours after release. To assess the phosphorylation status of exogenous LEM2 at different stages of the cell cycle, cells were arrested at G1/S by treating with 2 mM thymidine for 24 h. G1/S-arrested cells were then either harvested or thoroughly rinsed with PBS before the addition of culture medium containing 100 ng/ml nocodazole. After 16 h of nocodazole treatment, mitotic cells were harvested by shake off. Cell pellets were rinsed in TBS before lysis and separation by Phos-Tag acrylamide SDS-PAGE<sup>50</sup> and subsequent immunoblot (described below).

#### Immunoblots

To verify the efficacy of siRNA treatments and expression of siRNA-resistant constructs, cells were plated in six-well dishes and subjected to the same experimental conditions as those used to generate imaging data. Primary antibodies (anti-LEM2 (HPA017340; Sigma-Aldrich), anti-mCherry (Novus Biologicals, NBP2-25157), anti-tubulin (ab18251; abcam)) were detected with HRP-coupled secondary antibodies (Thermo Fisher) and chemiluminescence. To assess phosphorylation of exogenous LEM2 by mobility shift, whole-cell lysates were prepared in lysis buffer lacking EDTA and phosphatase inhibitors. Aliquots of each lysate were then subjected to either lambda protein phosphatase treatment per the manufacturer's instructions (P07535; NEB) or a control reaction treatment containing only PMP buffer and MnCl<sub>2</sub>. Samples were run on 8% and 10% SDS-PAGE gels prepared with Phos-Tag Acrylamide reagent (AAL-107; Waco) per the manufacturer's instructions. Detection of LEM2 by immunoblot was performed as described above.

#### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

#### Data availability

Raw data and peaklists from the quantitative crosslinking mass spectrometry analysis can be accessed with MassIVE: <ftp://massive.ucsd.edu/MSV000084837/>. Crosslinked peptide spectral assignments are accessible using accession: rgf6hriush at <http://msviewer.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msviewer>. The uncropped blots and gels are provided in Supplementary Fig. 1. Source Data for Figs. 1–3 and Extended Data Figs. 1–4, 5, 7, 8 are provided online.

31. Samwer, M. et al. DNA cross-bridging shapes a single nucleus from a set of mitotic chromosomes. *Cell* **170**, 956–972.e923 (2017).
32. Bradley, C. M., Ronning, D. R., Ghirlando, R., Craigie, R. & Dyda, F. Structural basis for DNA bridging by barrier-to-autointegration factor. *Nat. Struct. Mol. Biol.* **12**, 935–936 (2005).
33. Hamon, L., Savarin, P., Curmi, P. A. & Pastré, D. Rapid assembly and collective behavior of microtubule bundles in the presence of polyamines. *Biophys. J.* **101**, 205–216 (2011).
34. Gibson, B. A. et al. Organization of chromatin by intrinsic and regulated phase separation. *Cell* **179**, 470–484.e421 (2019).
35. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
36. Bui, K. H. et al. Integrated structural analysis of the human nuclear pore complex scaffold. *Cell* **155**, 1233–1243 (2013).
37. Leitner, A. et al. Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. *Mol. Cell Proteomics* **11**, M110.014126 (2012).
38. Trnka, M. J., Baker, P. R., Robinson, P. J., Burlingame, A. L. & Chalkley, R. J. Matching cross-linked peptide spectra: only as good as the worse identification. *Mol. Cell. Proteomics* **13**, 420–434 (2014).
39. Schilling, B. et al. Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline: application to protein acetylation and phosphorylation. *Mol. Cell. Proteomics* **11**, 202–214 (2012).
40. Müller, F., Fischer, L., Chen, Z. A., Auchyannikava, T. & Rappaport, J. On the reproducibility of label-free quantitative cross-linking/mass spectrometry. *J. Am. Soc. Mass Spectrom.* **29**, 405–412 (2018).
41. Combe, C. W., Fischer, L. & Rappaport, J. xiNET: cross-link network maps with residue resolution. *Mol. Cell. Proteomics* **14**, 1137–1147 (2015).
42. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protocols* **10**, 845–858 (2015).
43. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
44. Kosinski, J. et al. Xlink Analyzer: software for analysis and visualization of cross-linking data in the context of three-dimensional structures. *J. Struct. Biol.* **189**, 177–183 (2015).
45. Bajorek, M. et al. Biochemical analyses of human IST1 and its function in cytokinesis. *Mol. Biol. Cell* **20**, 1360–1373 (2009).
46. Donahue, D. A. et al. SUN2 overexpression deforms nuclear shape and inhibits HIV. *J. Virol.* **90**, 4199–4214 (2016).
47. Mackay, D. R., Makise, M. & Ullman, K. S. Defects in nuclear pore assembly lead to activation of an Aurora B-mediated abscission checkpoint. *J. Cell Biol.* **191**, 923–931 (2010).
48. Mackay, D. R., Elgort, S. W. & Ullman, K. S. The nucleoporin Nup153 has separable roles in both early mitotic progression and the resolution of mitosis. *Mol. Biol. Cell* **20**, 1652–1660 (2009).
49. Morita, E. et al. Human ESCRT-III and VPS4 proteins are required for centrosome and spindle maintenance. *Proc. Natl Acad. Sci. USA* **107**, 12889–12894 (2010).
50. Kinoshita, E., Kinoshita-Kikuta, E., Takiyama, K. & Koike, T. Phosphate-binding tag, a new tool to visualize phosphorylated proteins. *Mol. Cell. Proteomics* **5**, 749–757 (2006).
51. Cai, M. et al. Solution NMR structure of the barrier-to-autointegration factor-Emerin complex. *J. Biol. Chem.* **282**, 14525–14535 (2007).

**Acknowledgements** We thank W. I. Sundquist and S. Redding for critically reading the manuscript. For reagents, technical advice, and discussions we thank the Redding, Narlikar, Ullman, and Frost laboratories, as well as the Nikon Imaging Center at UCSF. We thank L. Williams for helping to score cell phenotypes and M. Mendoza and J. Rosenblatt for sharing microscope and other resources. We also thank the UCSF Center for Advanced cryoEM, including A. Myasnikov, D. Bulkley and M. Braunfeld. Our research was supported by NIH grants P50 GM082545, 1DP2-GM10772 (A.F.), 1R01-GM131052 (K.S.U.), and the Huntsman Cancer Foundation (K.S.U.). Shared resources used at the University of Utah were funded in part by the Huntsman Cancer Institute Cancer Center Support Grant NIH P30CA042014. Mass spectrometry analysis at the UCSF Mass Spectrometry Resource was supported by the Dr. Miriam and Sheldon G. Adelson Medical Research Foundation and a shared instrument grant (NIH S10OD016229). A.V.A. was funded by EMBO (ALTF 455-2016) and the German Research Foundation (DFG AP 298/1-1). I.E.J. was funded by the NSF Graduate Research Fellowship (I000232072) and a Mortiz-Heyman Discovery Fellowship. A.F. is a Chan Zuckerberg Biohub investigator and an HHMI faculty scholar.

# Article

**Author contributions** A.v.A., D.L., I.E.J., M.J.T., A.L.B., K.S.U., and A.F. designed research; A.v.A. and I.E.J. purified proteins and performed in vitro fluorescent and negative stain electron microscopy imaging, FRAP, and analysis. A.v.A. performed protein droplet imaging. A.v.A. performed 2D averaging and homology modeling. D.L. performed mammalian cell experiments and analysis. D.L. prepared fixed, immunostained samples for STED and A.v.A. performed STED imaging and analysis. I.E.J. performed and analysed turbidimetry and analytical SEC experiments. A.v.A. performed XL-MS sample preparation and analysis. M.J.T. performed mass spectrometry and analysis. A.v.A. performed and analysed polymer pelleting experiments. S.M.P. purified CHMP7<sub>ESCRT-III</sub>. A.v.A., D.L., I.E.J., M.T., K.S.U., and A.F. analysed data. A.v.A., D.L., I.E.J., K.S.U., and A.F. wrote the manuscript.

**Competing interests** The authors declare no competing interests.

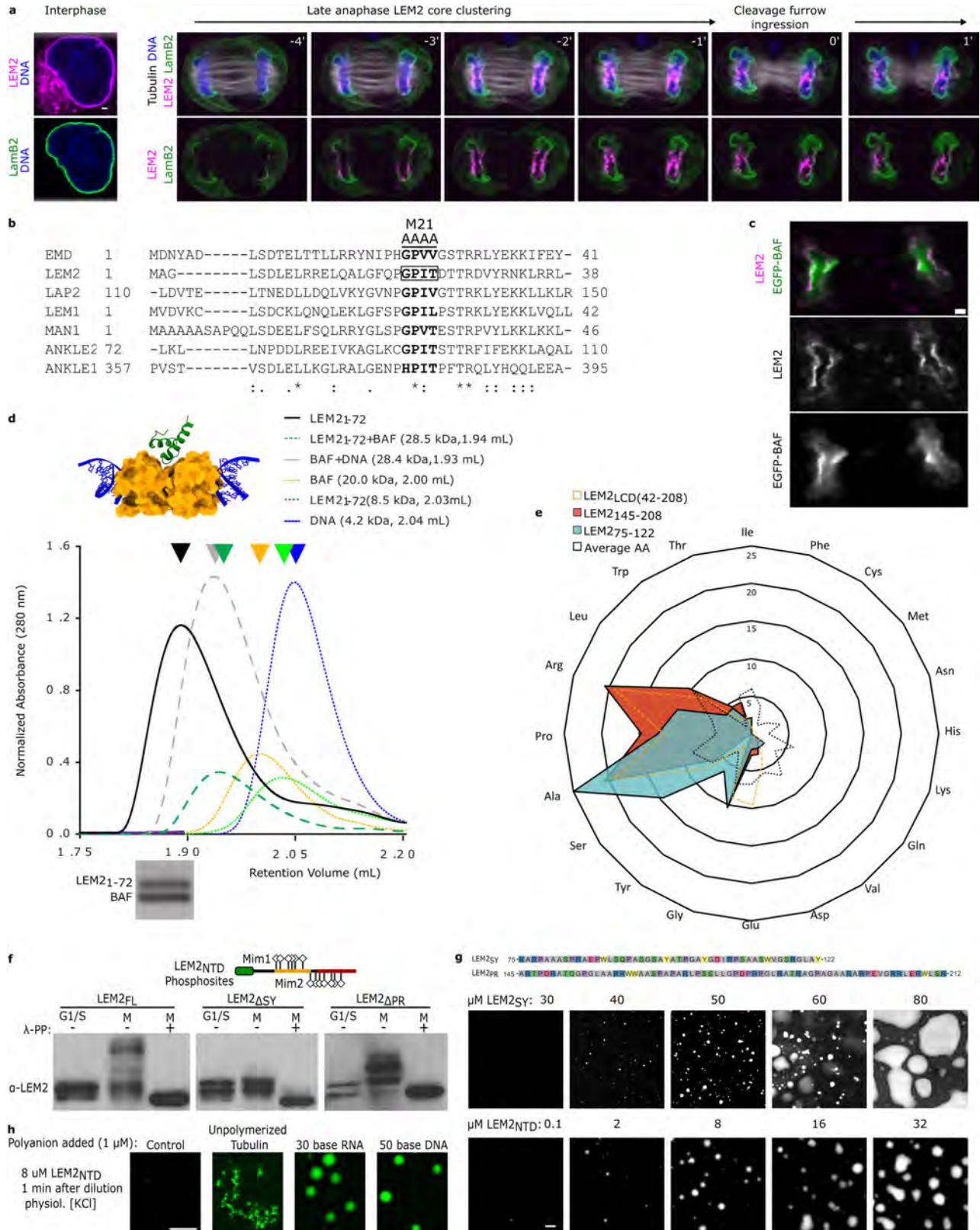
**Additional information**

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2232-x>.

**Correspondence and requests for materials** should be addressed to K.S.U. or A.F.

**Peer review information** *Nature* thanks Stephen Michnick and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

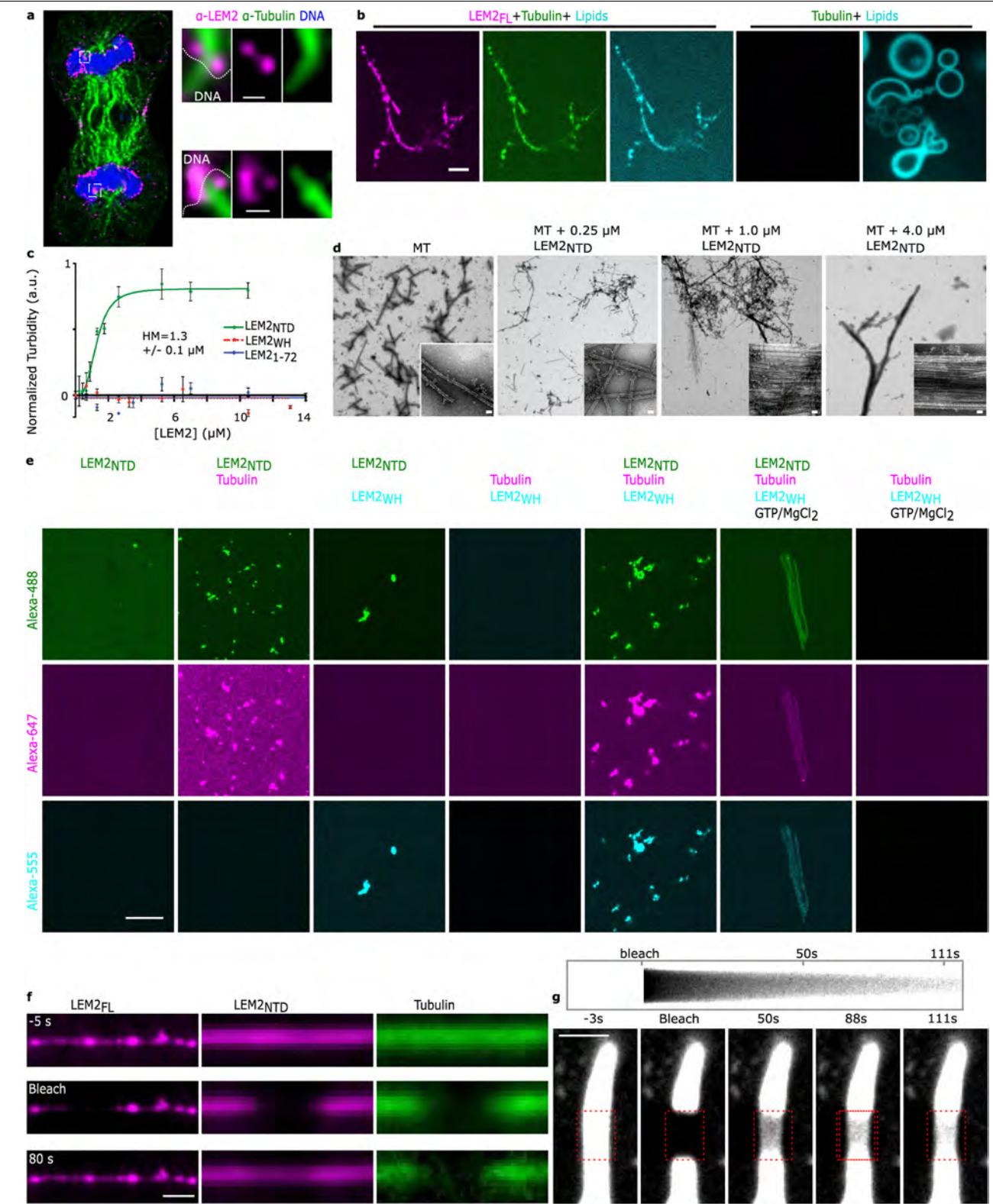


**Extended Data Fig. 1** | See next page for caption.

# Article

**Extended Data Fig. 1 | The N terminus of LEM2 possesses a canonical BAF-binding LCD and an LCD.** **a**, Live-cell imaging of GFP-lamin B2 and LEM2-mChr; DNA is stained with NucBlue and tubulin detected with SiR-tubulin. Time 0 refers to complete CFI. Representative of ten or more cells imaged across at least three biological replicates. Scale bar, 2 μm. **b**, Multiple sequence alignment of LEM domains across LEM family proteins, highlighting a conserved four-amino-acid sequence that, when mutated in emerin (EMD<sub>m24</sub>), disrupts BAF binding<sup>20</sup>. The position of an analogous mutation in LEM2 (LEM2<sub>m21</sub>) is indicated. **c**, HeLa cells stably expressing LEM2-mCherry and EGFP-BAF live-imaged during anaphase. Representative of ten or more cells imaged across at least three biological replicates for both fixed and live-cell imaging. Scale bar, 10 μm. **d**, Top, a homology model for the LEM2<sub>1-72</sub>-BAF-DNA complex<sup>51</sup>, based on Protein Data Bank code (PDB): 2BZF and PDB: 2ODG<sup>32,51</sup>. Middle, absorbance at 280 nm as a function of retention volume (ml) from analytical size exclusion chromatography. Retention volumes for major peaks (arrowheads) and predicted molecular weights for protein or protein-DNA complexes are listed. Bottom, SDS-PAGE of major peak for LEM2<sub>1-72</sub> + BAF + DNA sample. Representative of three technical replicates. For

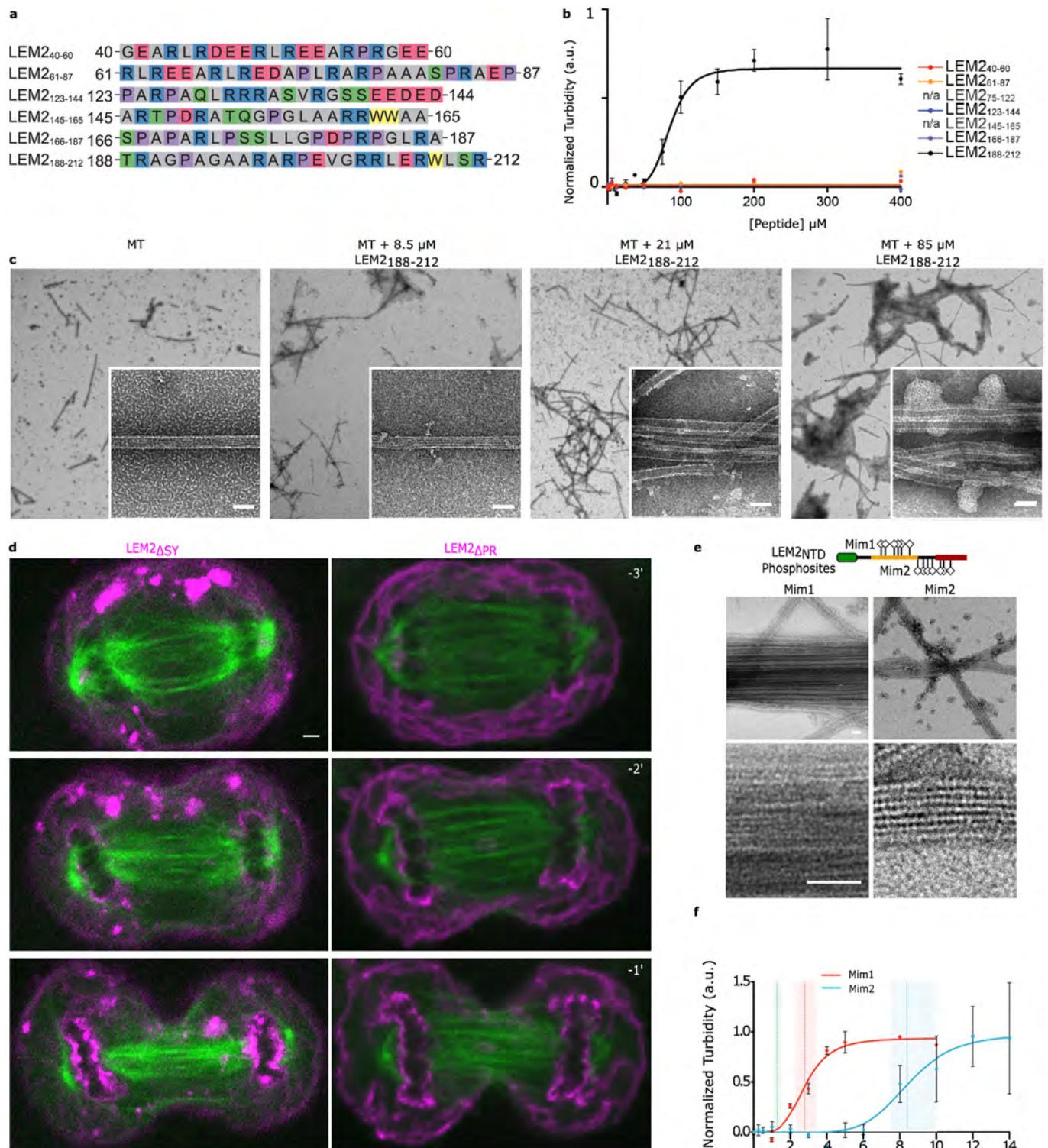
gel source data, see Supplementary Fig. 1. **e**, Percentage amino acid composition for the LEM2 LCD, and the compositions of two subregions, compared to an average amino acid composition. **f**, Schematic of LEM<sub>NTD</sub> with amino acid substitutions (S, T, or Y to D) relative to SY-rich (yellow) and PR-rich (rust) regions, in LEM2<sub>NTD</sub> Mim1 and Mim2 constructs. LEM2 immunoblot assessing the migration pattern of full-length (also shown in Fig. 1e) and mutant LEM2-mChr constructs following separation by Phos-tag SDS-PAGE. Cell lysates were prepared from G1/S- and prometaphase-arrested cells expressing the indicated exogenous LEM2; lysates treated with λ-PP are indicated. Representative data from two biological replicates, with one and three technical replicates per biological replicate. For immunoblot source data, see Supplementary Fig. 1. **g**, Top, amino acid sequences of the peptides corresponding to the SY-rich and PR-rich regions of LEM2. Bottom, concentration-dependent droplet formation by the LEM2<sub>SY</sub> peptide, juxtaposed with similar data collected for full LEM2<sub>NTD</sub> as in Fig. 1d. Representative of three technical replicates. Scale bar, 2 μm. **h**, Fluorescence microscopy of purified LEM2<sub>NTD</sub> with indicated molecular anions. Image representative of two technical replicates. Scale bar, 2 μm.



**Extended Data Fig. 2 | LEM2 wets the surface of microtubules with a liquid-like coat.** **a**, Additional example of STED imaging of endogenous LEM2 localization during late anaphase. Scale bars, 150 nm. Representative of data from four cells. **b**, Fluorescence imaging of indicated combinations of LEM2<sub>FL</sub>-Alexa488 (magenta), tubulin-Alexa647 (green), and lipids labelled with PE-rhodamine (cyan). Scale bar, 2 μm. Representative of two technical replicates. **c**, Light scattering at 340 nm (turbidity) of microtubule bundling by indicated LEM2 constructs. Half maximal concentration of LEM2<sub>NTD</sub> is  $1.303 \pm 0.1 \mu\text{M}$ . Mean  $\pm$  s.e.m. for  $n=3$  independent samples. **d**, Negative stain electron microscopy of microtubules alone or microtubules with indicated

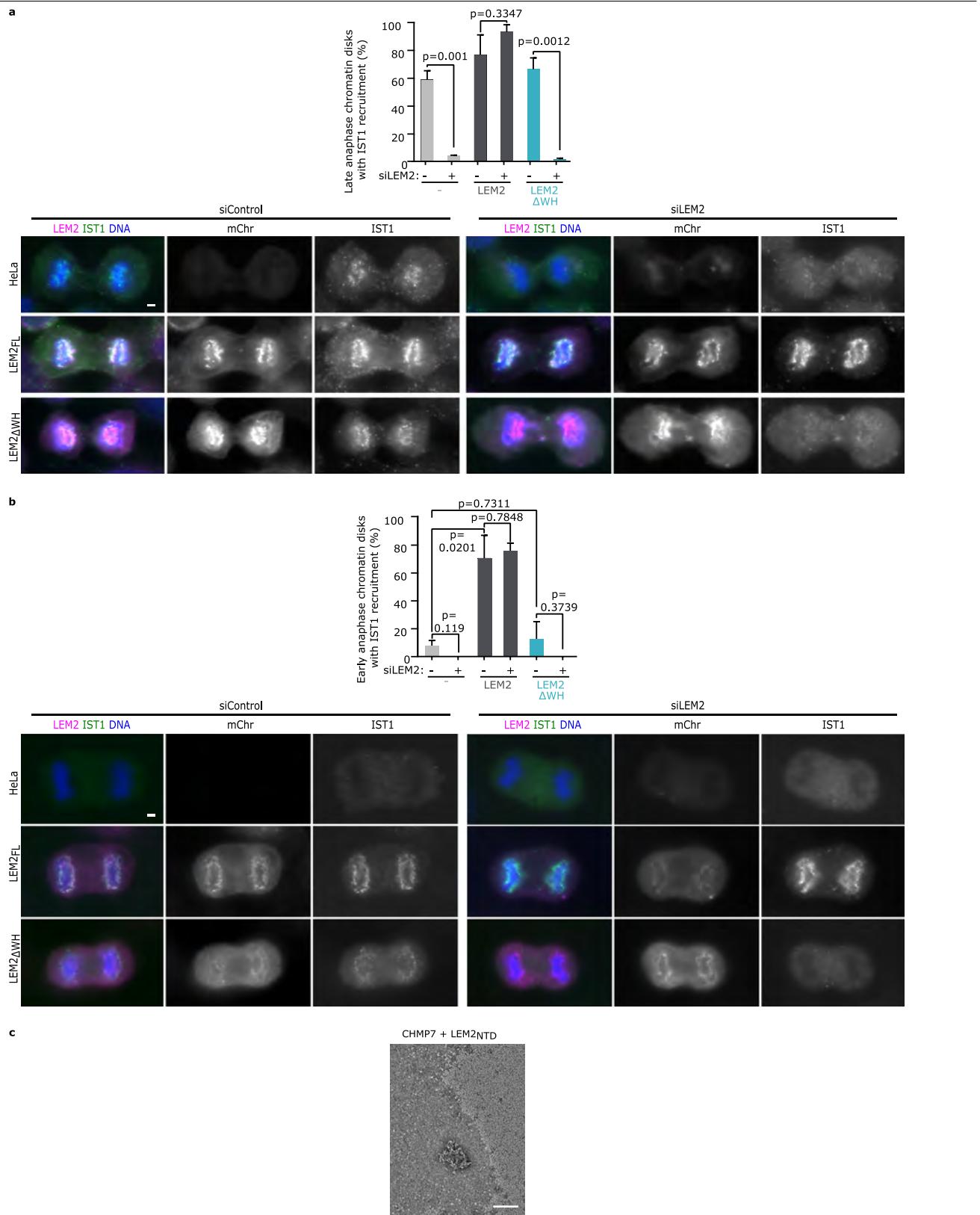
concentrations of LEM2<sub>NTD</sub>, corresponding to the concentrations measured by light scattering. Images representative of at least six; scale bars, 25 nm. **e**, Fluorescence microscopy of LEM2<sub>NTD</sub>-Alexa488 in combination with tubulin-Alexa647, LEM2<sub>WH</sub>-Alexa555, and GTP/MgCl<sub>2</sub>, as indicated. Scale bar, 10 μm. Images are representative of at least three. **f**, Example images for FRAP of LEM2<sub>FL</sub>- and LEM2<sub>NTD</sub>-coated microtubule bundles, representative of five independent samples (LEM2<sub>FL</sub>) or 17 independent samples (LEM2<sub>NTD</sub>). Scale bar, 2 μm. **g**, Top, kymograph across bleached region. Bottom, FRAP of LEM2<sub>NTD</sub>-coated microtubule bundle. Images show fluorescent LEM2<sub>NTD</sub> channel. Scale bar, 2 μm. Representative of eight technical replicates.

# Article



**Extended Data Fig. 3 | LEM2-LCD bundles microtubules in vitro through a regulated microtubule-binding domain.** **a**, Amino-acid sequences of six LEM2 peptides tiling the LCD. LEM2<sub>145-165</sub> is LEM2-PR<sub>A</sub>, and LEM2<sub>188-212</sub> is LEM2-PR<sub>B</sub>. **b**, Light scattering at 340 nm (turbidity) of microtubule bundling by indicated LEM2 peptides. Half maximal concentration of LEM2<sub>188-212</sub> (LEM2-PR<sub>B</sub>) is 85.11 μM. For LEM2<sub>188-212</sub>, data plotted are mean ± s.e.m. for  $n = 3$  technical replicates. LEM2<sub>40-60</sub>, LEM2<sub>61-81</sub>, LEM2<sub>123-144</sub> and LEM2<sub>166-187</sub> did not bundle. **c**, Negative stain electron microscopy of microtubules alone or microtubules with indicated concentrations of LEM2<sub>188-212</sub>, corresponding to turbidity reactions. Scale bars, 25 nm. For each condition, images are representatives of seven. **d**, Live-cell imaging of indicated LEM2<sub>ΔSY</sub>-mCherry and

LEM2<sub>ΔPR</sub>-mCherry deletion constructs (magenta) and GFP-tubulin (green). Time 0 refers to time of complete CFI. Scale bar, 2 μm. Images are representative of three independent experiments. **e**, Negative stain electron microscopy of microtubules co-incubated with the indicated phosphomimetic proteins. Scale bar, 25 nm. Images are representative of two technical replicates. **f**, Light scattering at 340 nm (turbidity) of microtubule bundling by phosphomimetic LEM2 constructs. Vertical lines are half maximum, shaded regions show s.e.; LEM2<sub>NTD</sub> (green) half maximum 1.303 μM (1.144–1.493 μM), LEM2<sub>Mim1</sub> (red) half maximum 2.824 μM (2.283–3.397 μM), LEM2<sub>Mim2</sub> (blue) half maximum 8.442 μM (7.511–10.06 μM). Data plotted are mean ± s.e.m. from three technical replicates.



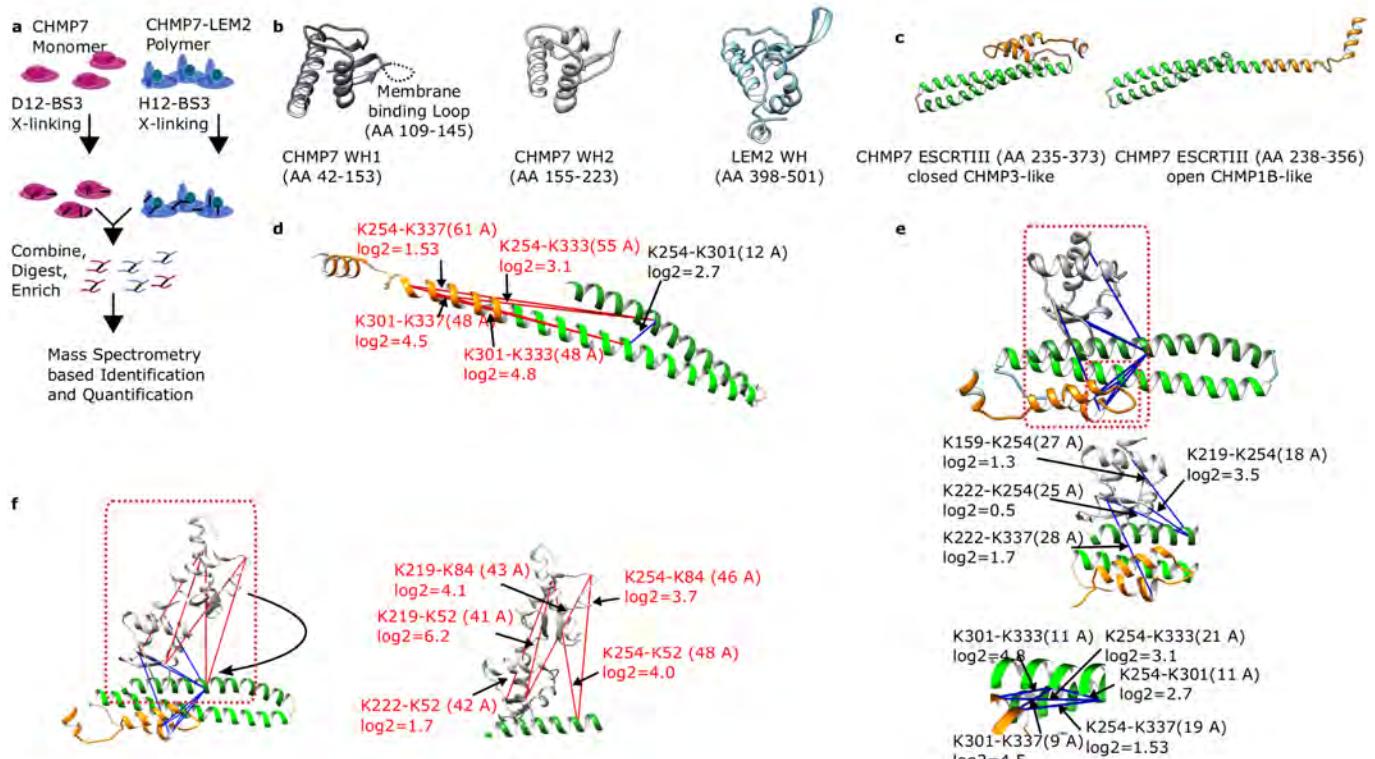
**Extended Data Fig. 4** | See next page for caption.

# Article

## Extended Data Fig. 4 | The LEM2 WH domain is required for recruitment of IST1 to the nascent nuclear envelope and mediates polymer formation with CHMP7.

**a.** Top, quantification of robust IST1 recruitment to chromatin discs in late anaphase, as assessed by blind scoring. Mean  $\pm$  s.e.m. determined from three independent experiments (siControl parental:  $n = 80, 58, 106$ ; siLEM2-2/parental:  $n = 78, 50, 58$ ; siControl/LEM2-mChr:  $n = 46, 42, 62$ ; siLEM2-2/LEM2-mChr:  $n = 78, 51, 62$ ; siControl/LEM2<sub>ΔWH</sub>-mChr:  $n = 138, 52, 42$ ; siLEM2-2/LEM2<sub>ΔWH</sub>-mChr:  $n = 112, 48, 51$ ). Two-tailed unpaired *t*-test, no multiple comparisons. Bottom, representative images by widefield showing localization of endogenous IST1 in late anaphase cells depleted of endogenous LEM2 and expressing the indicated siRNA-resistant LEM2-mChr constructs. Scale bar, 2  $\mu$ m. **b.** Top, quantification of the percent of early anaphase discs

with robust IST1 recruitment, as assessed by blind scoring. Mean  $\pm$  s.e.m. determined from three independent experiments (siControl/parental:  $n = 38, 16, 20$ ; siLEM2-2/parental:  $n = 24, 10, 25$ ; siControl/LEM2-mChr:  $n = 50, 20, 24$ ; siLEM2-2/LEM2-mChr:  $n = 38, 14, 8$ ; siControl/LEM2<sub>ΔWH</sub>-mChr:  $n = 44, 16, 20$ ; siLEM2-2/LEM2<sub>ΔWH</sub>-mChr:  $n = 20, 4, 14$ ). Two-tailed unpaired *t*-test, no multiple comparisons. Bottom, representative images by widefield showing localization of endogenous IST1 in early anaphase cells depleted of endogenous LEM2 and expressing the indicated siRNA-resistant LEM2-mChr constructs. Scale bar, 2  $\mu$ m. **c.** Negative stain electron microscopy corresponding to the CHMP7 polymerization assay showing no polymerization for the control condition CHMP7 + LEM2<sub>NTD</sub>. Representative of two technical replicates.

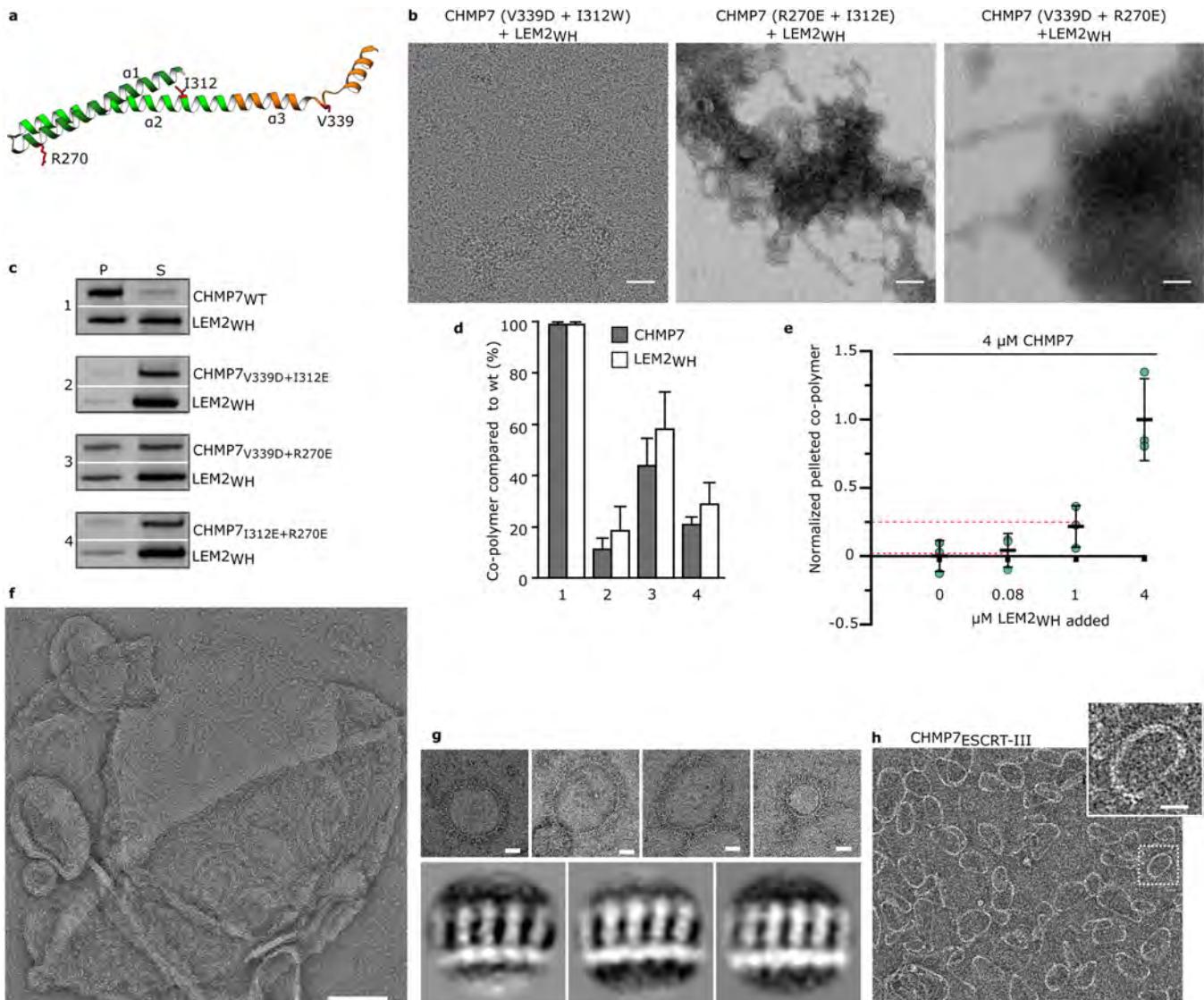


**Extended Data Fig. 5 | Homology modelling and XL-MS consistent with a WH domain-swap mechanism for activation of CHMP7 by LEM2.**

**a**, Workflow of lysine–lysine hybrid peptide mapping using XL-MS. BS3 cross-links surface-accessible lysine residues with  $\text{C}\alpha$ – $\text{C}\alpha$  distances below about 3 nm. **b**, Homology models for the WH domains of CHMP7 and LEM2 from reference structures described in Supplementary Table 2. WH1 of CHMP7 contains a membrane-binding region indicated as a loop<sup>4</sup>. **c**, Homology models of the CHMP7 ESCRT-III-fold in open and closed conformations. Green and orange,  $\alpha$ -helices 1–2 and 3–4, respectively. **d**, Distance restraints identified from XL-MS analysis of the CHMP7 monomer were mapped to open and closed homology models.  $\text{C}\alpha$ – $\text{C}\alpha$  distances over 3 nm are considered violations. Blue, satisfied restraints; red, violated restraints. The open ESCRT-III conformation

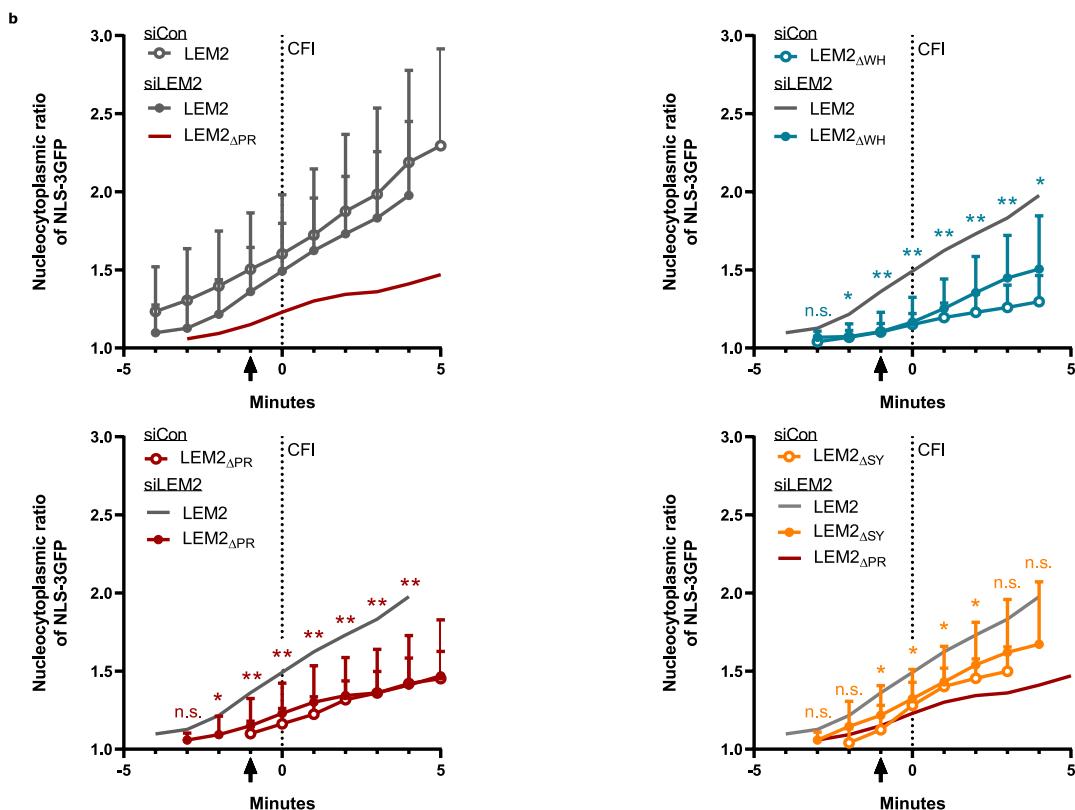
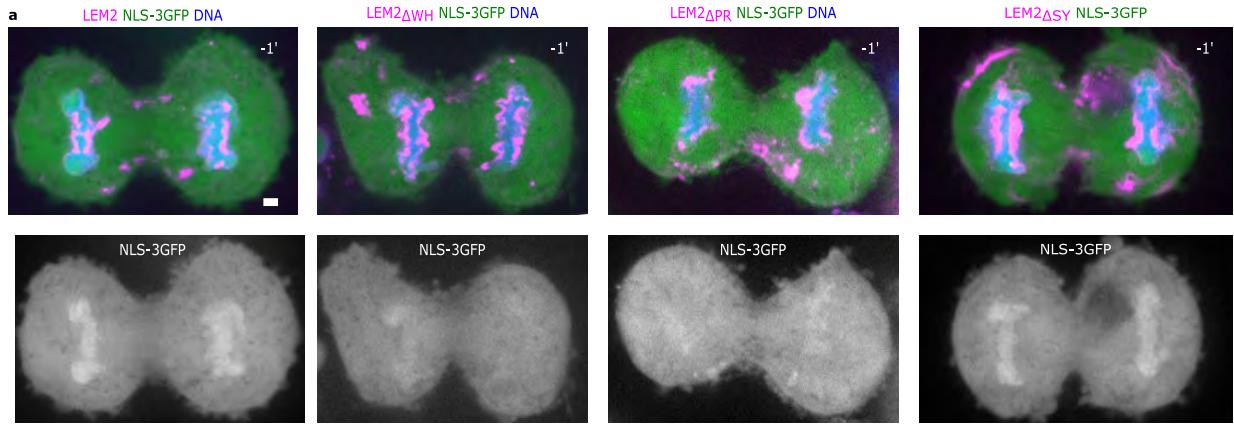
was rejected. **e**, Top, a crystallographic interface between VPS25 and an ESCRT-III helix<sup>28</sup> serves as a template for the XL-MS-informed homology model of the CHMP7 WH2 interaction with the CHMP7 ESCRT-III domain. For details on template structures see Supplementary Table 2. Middle, bottom, all cross-links are satisfied when mapped to the closed CHMP7-ESCRT-III model and agree with domain connectivity. A subset of cross-links was not satisfied when mapping WH1 instead to the same interface (data not shown). **f**, Left, distance restraints identified from XL-MS analysis of CHMP7 monomer mapped to conformation of polymerized CHMP7 consistent with 2D class averages. Violated restraints suggest a hinge region between CHMP7 WH1 and WH2 that allows its WH1 to move closer to the ESCRT-III core (black arrow). Right, violated restraints to WH2.

# Article

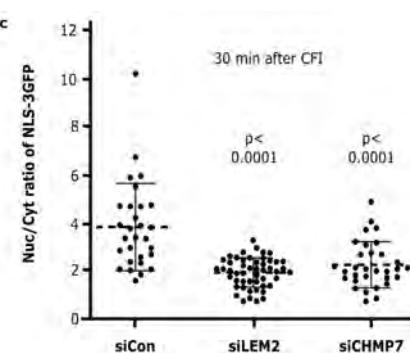


**Extended Data Fig. 6 | Activated CHMP7 forms polymeric rings via its ESCRT-III domain, exposing its tandem WH domains.** **a**, CHMP7 point mutations are indicated in an open ESCRT-III fold, representing polymerized CHMP7. **b**, Negative stain electron microscopy of His<sub>6</sub>-SUMO-LEM2<sub>WH</sub> co-incubated with CHMP7 mutants. Scale bars, 50 nm. Images representative of three technical replicates. **c**, SDS-PAGE of pellet (P) or supernatant (S) following centrifugation of LEM2<sub>WH</sub> incubated with wild-type or mutant CHMP7. For gel source data, see Supplementary Fig. 1. **d**, Quantification of pelleted protein after SDS-PAGE and Coomassie blue staining for mutant versus wild-type proteins. Mean  $\pm$  s.d. quantified from  $n=3$  independent experiments. **e**, SDS-PAGE-based relative quantification of polymerized and

pelleted CHMP7 with different ratios of LEM2<sub>WH</sub> present. Red dashed lines indicate expected fraction of CHMP7 in the pelleted polymer, assuming 1:1 stoichiometric polymer. Mean  $\pm$  s.d. quantified from  $n=3$  independent experiments. For gel source data, see Supplementary Fig. 1. **f**, Negative stain electron microscopy of CHMP7 polymers on a liposome. Scale bar, 100 nm. **g**, Top, negative stain electron microscopy of membrane-induced CHMP7 polymers used for 2D averaging. Scale bar, 20 nm. Bottom, representative 2D class averages from manually picked particles from polymers shown at top. **f, g**, Representative of five technical replicates. **h**, Negative stain electron microscopy of CHMP7<sub>ESCRT-III</sub> (AA 229–453). Representative of three technical replicates. Scale bar, 20 nm.



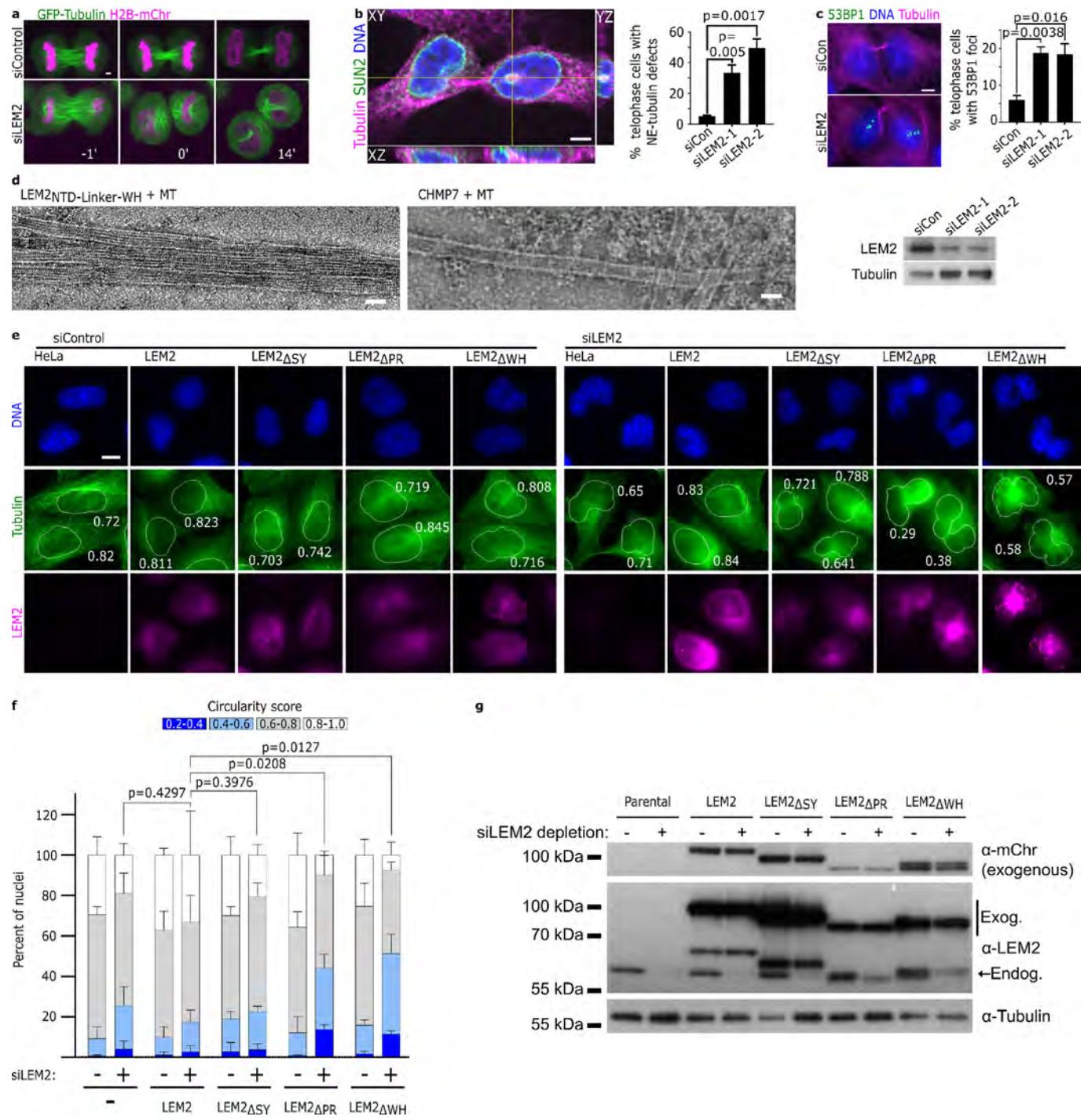
time	p-value		
	$\Delta SY$	$\Delta PR$	$\Delta WH$
-3'	0.2092	0.2094	0.3131
-2'	0.2024	0.0287	0.0146
-1'	0.0292	0.0032	0.0009
0'	0.0167	0.0008	0.0001
1'	0.0146	0.0003	0.0001
2'	0.0299	<0.0001	0.0003
3'	0.0805	<0.0001	0.0037
4'	0.0580	0.0004	0.0070



**Extended Data Fig. 7** | See next page for caption.

## Article

**Extended Data Fig. 7 | LEM2 promotes early nuclear compartmentalization via cooperation between its LCD and WH domains.** **a**, Example images of cells treated with si*LEM2* and expressing NLS–3GFP in combination with siRNA-resistant LEM2 constructs corresponding to the quantification graphs shown in **b** and Fig. 3f. DNA labelled using NucBlue. Time 0 refers to the time of complete CFI. Scale bar, 2  $\mu$ m. **b**, Mean  $\pm$  s.d. nuclear/cytoplasmic ratio of NLS–3GFP fluorescence over time in cells treated with the indicated siRNAs and expressing the indicated siRNA-resistant constructs. Cells imaged at 15-s intervals though 1-min increments are plotted. Data were collected across at least three biological replicates (siCon/LEM2–mChr:  $n = 26$ ; si*LEM2*-2/LEM2–mChr:  $n = 44$ ; siCon/LEM2<sub>ΔSY</sub>–mChr:  $n = 20$ ; si*LEM2*-2/LEM2<sub>ΔSY</sub>–mChr:  $n = 24$ ; siCon/LEM2<sub>ΔPR</sub>–mChr:  $n = 20$ ; si*LEM2*-2/LEM2<sub>ΔPR</sub>–mChr:  $n = 20$ ; siCon/LEM2<sub>ΔWH</sub>–mChr:  $n = 24$ ; si*LEM2*-2/LEM2<sub>ΔWH</sub>–mChr:  $n = 16$ ). Two-tailed unpaired *t*-test was used to determine *P* values comparing deletion mutant lines to the full-length LEM2 line under endogenous LEM2 depletion conditions at each time point. No multiple comparisons. \**P* < 0.05, \*\**P* < 0.005; exact *P* values below. **c**, Quantification of nuclear/cytoplasmic ratio of NLS–3GFP approximately 30 min after complete CFI in parental HeLa cells treated with the indicated siRNAs. Data were collected across three biological replicates and plotted as mean  $\pm$  s.d. (siControl:  $n = 11, 12, 6$ ; si*LEM2*-2:  $n = 18, 18, 14$ ; siCHMP7:  $n = 11, 6, 14$ ). Two-tailed unpaired *t*-test, no multiple comparisons.



**Extended Data Fig. 8** | See next page for caption.

# Article

**Extended Data Fig. 8 | Loss of LEM2 function leads to DNA damage and abnormal nuclear morphologies.** **a**, Live-cell imaging of GFP-tubulin and H2B-mChr in siRNA-treated cells. Images representative of two biological replicates, quantified in **b**. Time 0 refers to the time of complete CFI. Scale bar, 2 μm. **b**, Left, orthogonal view of the tubulin phenotype following LEM2 depletion co-stained for the nuclear envelope protein SUN2. Images representative of three biological replicates. Scale bar, 5 μm. Right, mean ± s.e.m. percentage of telophase cells with nuclear tubulin defects, lined with inner nuclear membrane (as assessed by immunofluorescence of lamin B2). Three biological replicates (siControl:  $n = 75, 35, 37$ ; siLEM2-1:  $n = 72, 35, 37$ ; siLEM2-2:  $n = 45, 34, 34$ ). Two-tailed unpaired *t*-test, no multiple comparisons. **c**, Left, example images of 53BP1 localization by immunofluorescence in telophase U2OS cells following siRNA treatment. Scale bar, 5 μm. Right, mean ± s.e.m. percentage of telophase cells with five or more 53BP1 nuclear foci. Three biological replicates (siControl:  $n = 56, 52, 44$ ; siLEM2-1:  $n = 64, 56, 26$ ; siLEM2-2:  $n = 74, 50, 40$ ). Two-tailed unpaired *t*-test, no multiple comparisons. Bottom, immunoblot confirming depletion of endogenous LEM2 in U2OS cells, using siRNA oligos previously validated in other human cell lines, including HeLa<sup>4,19</sup> (immunoblot source data shown in Supplementary Fig. 1). **d**, Negative stain electron microscopy of indicated combinations of microtubules, LEM2<sub>NTD-linker-WH</sub>, and CHMP7<sub>FL</sub>. Scale bars, 25 nm. Images representative of two technical replicates. **e**, Example images of cells expressing the indicated siRNA-resistant constructs and treated with the indicated siRNAs. Cells were arrested in S-phase and then allowed to progress through one round of division, resulting in an interphase population of cells that had just exited mitosis. We observed an increased number of highly

irregular nuclei in cells expressing either LEM2<sub>ΔPR</sub>-mChr or LEM2<sub>ΔWH</sub>-mChr compared to cells expressing full-length LEM2 or even those depleted of LEM2. Notably, deformed nuclei were commonly associated with microtubule disorganization and aberrant accumulation of LEM2<sub>ΔPR</sub>-mChr and LEM2<sub>ΔWH</sub>-mChr. Representative nuclear, tubulin, and LEM2 phenotypes and the correspondence to nuclear circularity score are shown. Nuclear borders and circularity scores annotated in tubulin channel. Scale bar, 5 μm. These findings suggest that interfering with cooperation between the microtubule-interacting and ESCRT-binding domains of LEM2 alters nuclear morphology, indicating that both activities are necessary, but neither is sufficient for nuclear envelope reformation. Moreover, the presence of one activity without the other is detrimental to nuclear morphology. **f**, Quantification of nuclear circularity in interphase parental HeLa cells and cells expressing the indicated siRNA-resistant LEM2 constructs, treated with the indicated siRNAs. Mean ± s.e.m. from three biological replicates (siControl/parental:  $n = 105, 46, 80$ ; siLEM2-2/parental:  $n = 102, 116, 59$ ; siControl/LEM2-mChr:  $n = 153, 53, 122$ ; siLEM2-2/LEM2-mChr:  $n = 84, 81, 105$ ; siControl/LEM2<sub>ΔPR</sub>-mChr:  $n = 123, 68, 144$ ; siLEM2-2/LEM2<sub>ΔPR</sub>-mChr:  $n = 93, 95, 105$ ; siControl/LEM2<sub>ΔPR</sub>-mChr:  $n = 149, 123, 58$ ; siLEM2-2/LEM2<sub>ΔPR</sub>-mChr:  $n = 49, 31, 42$ ; siControl/LEM2<sub>ΔWH</sub>-mChr:  $n = 116, 96, 94$ ; siLEM2-2/LEM2<sub>ΔWH</sub>-mChr:  $n = 85, 32, 68$ ). Two-tailed unpaired *t*-test comparing circularity scores less than 0.6 (indicated by blue) between the indicated treatments; no multiple comparisons. **g**, Immunoblot showing relative levels of the siRNA-resistant constructs fused with mCherry in parallel with endogenous LEM2. Representative of two technical replicates. For immunoblot source data, see Supplementary Fig. 1.

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

## Software and code

Policy information about [availability of computer code](#)

Data collection

Software used for data collection: Proteome Discoverer 2.2 (Thermo), Protein Prospector 5.20.23, Skyline 4.1, SerialEM 3-5-9, ImageJ 1.52a, Leica Application Suite X 3.1.5, Metamorph Version 7.10.1.161, and Axiovision (AxioVs40) version 4.8.2.0. All software packages used for data acquisition are described in the method section.

Data analysis

ImageJ 1.52a, UCSF Chimera-1.14, Relion 2.0 , Leica Application Suite X 3.7.1, Huygens Professional 17.04, For mammalian cell data: (fiji Is Just) ImageJ 2.0.0-rc-65/1.52o, GraphPad Prism Version 8.3.0

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Raw data and peaklists from the quantitative crosslinking mass spectrometry analysis can be accessed with MassIVE: <ftp://massive.ucsd.edu/MSV000084837/>. Crosslinked peptide spectral assignments are accessible using accession: rgf6hriush at <http://msviewer.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msviewer>. The data that supports the findings of this study are available within Supplementary Information files, and from the corresponding authors upon reasonable request.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](http://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were not predetermined using statistical methods. Sample sizes are consistent with what is published in the field.
Data exclusions	As the experimental parameters were determined prior to data collection, data was excluded only in instances where technical limitations existed. Criteria for exclusion of data was determined prior to quantification and applied uniformly across experiments and samples. Specifically, for live cell imaging, individual cells were excluded if the nuclei moved out of focus or expression of exogenous, fluorescent protein-tagged constructs either fell below a predetermined threshold of expression or was high, resulting in image saturation. Cells with visible lagging chromosomes were also excluded. For nuclear compartmentalization, data was excluded for timepoints where n was less than or equal to 10. Outliers were not excluded in mammalian cell data sets. For turbidity measurements, data was excluded where bubbles in plate prevented accurate light scattering measurements of sample.
Replication	All attempts at replication were successful. For quantified mammalian cell data, no validated replicate experiments were excluded. For qualitative mammalian cell data, validated replicate experiments were not excluded when determining representative phenotypes.
Randomization	Only in the instance where a visual score of IST1 recruitment phenotypes by immunofluorescence was subjective in nature were data randomized. Specifically, all the images across all treatment conditions for each experiment (three replicate experiments) were coded and randomized as a single set of data to be scored visually by three blind scorers (blinding described below) using predetermined criteria. The results of blind scoring was recorded for all images, which were then decoded for quantification. Therefore, random allocation is not relevant – in other words, subjects/data were not assigned into groups for experimentation/study. For this reason, in combination with the fact the experimental samples were all treated in parallel (for each of the three independent experiments), covariates are not relevant.
Blinding	IST1 recruitment was scored blindly, consistent with previously published approaches. Experiments tracking cell phenotypes that were not blinded are consistent with what is published in the field, including using objective quantification methods where-ever possible.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

### Antibodies used

[anti-53BP1; Millipore, MAB3802 (lot 2753986); Mouse monoclonal BP13; reactivity: Human; application: Human cells, IF at 1:1,000;], [anti-Lamin B2; AbCam, ab8983 (lot GR141423); Mouse monoclonal LN43; reactivity: Mouse, Hamster, Human, Xenopus laevis, Zebrafish; application: Human cells, IF at 1:2,000 and WB at 1:1000], [anti-LEMD2; Sigma-Aldrich; HPA017340 (lots B96759 and B106882); Rabbit polyclonal; reactivity: rat, human, mouse; application: Human cells, IF at 1:100 and WB at 1:500], [anti-mCherry; Novus Biologicals, NBP2-25157 (lot 010519); Rabbit polyclonal; reactivity: human, mouse, rat, non-species specific; application: detecting exogenous expression in Human cells, WB at 1:2,000], [anti-Tubulin; AbCam, ab18251 (lots GR175278-1 and GR3198330-1); Rabbit polyclonal; reactivity: Mouse, Rat, Chicken, Cow, Human, Drosophila melanogaster, Indian muntjac, African green monkey, Chinese hamster; application: Human cells, IF at 1:1,000 and WB at 1:20,000], [anti-Tubulin; Accurate Chemical & Scientific, YSRTMCA77G; Rat monoclonal YL1/2; reactivity: Yeast, Birds, Mammals; application: Human cells, IF at 1:1,000], [anti-IST1; Gift from Wes Sundquist (generated by the Sundquist lab, published in Bajorek, et al MBio 2009; doi:10.1091/mbe.E08-05-0475); Rabbit polyclonal; validated reactivity: Human; application: Human cells, IF at 1:500; Ab.], [anti-SUN2; unpurified hybridoma supernatant, gift from Brian Burke (unpublished), Mouse monoclonal 3.1E; validated reactivity: Human; application: Human cells, IF at 1:100]

## Validation

[anti-53BP1; from manufacturer's website: "validated in IF & WB"], [anti-Lamin B2; from manufacturer's website: "Our Abpromise guarantee covers the use of ab8983 in the following tested applications. The application notes include recommended starting dilutions; optimal dilutions/concentrations should be determined by the end user." Tested applications listed: ICC, Flow Cyt, IHC-Fr, WB, IP, ICC/IF.], [anti-LEMD2; from the manufacturer's website: "All Prestige Antibodies Powered by Atlas Antibodies are developed and validated by the Human Protein Atlas (HPA) project ([www.proteinatlas.org](http://www.proteinatlas.org))and as a result, are supported by the most extensive characterization in the industry." Additionally, we previously published further characterization by IF and western, following siRNA depletion: doi: 10.1073/pnas.1613916114], [anti-mCherry; From manufacturer's website: "This mCherry antibody is useful for Immunocytochemistry/Immunofluorescence and Western Blot, where a band can be seen at ~28 kDa. Use in IHC and IHC-P reported in scientific literature (PMID: 27396338 and 28891816 respectively). Use in Live Imaging Microscopy was reported from a verified customer review."], [rabbit anti-Tubulin; From manufacturer: "Our Abpromise guarantee covers the use of ab18251 in the following tested applications. The application notes include recommended starting dilutions; optimal dilutions/concentrations should be determined by the end user." Tested applications listed: ICC/IF, IHC-P, Flow Cyt, WB], [rat anti-Tubulin; From the manufacturer's data sheet: "This product has been reported to work in the following applications." Applications listed by the manufacturer: Immunohistology, ELISA, Immunoprecipitation, Western Blotting, Radioimmunoassay.], [anti-SUN2 was a gift from Brian Burke and we confirmed that the antibody decorated the nuclear membrane, as expected, by immunofluorescence at the nuclear envelope (co-localizing with a validated antibody specific to SUN1 - AbCam ab124770; rabbit monoclonal EPR6554; Abpromise guarantee covers the use of ab124770 by ICC/IF, among other techniques) and that it recognized a single band corresponding to the anticipated size by western blot analysis], [anti-IST1; Antibody specificity confirmed by siRNA depletion followed by ICC and WB in source manuscript: Bajorek, et al MBoC 2009; doi:10.1091/mbo.08-05-0475]

## Eukaryotic cell lines

Policy information about [cell lines](#)

## Cell line source(s)

HeLa cells were a gift from Maureen Powers (Emory University School of Medicine). U2OS were a gift from Don Ayer (University of Utah School of Medicine).

## Authentication

Both cell lines were authenticated by STR profiling.

## Mycoplasma contamination

All HeLa cell lines were generated from the same parental HeLa cells, which tested negative for mycoplasma contamination. U2OS cells were not tested for mycoplasma contamination.

Commonly misidentified lines  
(See [ICLAC](#) register)

No commonly misidentified lines were used in this study.

## Article

# Convergent genes shape budding yeast pericentromeres

<https://doi.org/10.1038/s41586-020-2244-6>

Received: 11 March 2019

Accepted: 25 February 2020

Published online: 29 April 2020

 Check for updates

Flora Paldi<sup>1</sup>, Bonnie Alver<sup>1</sup>, Daniel Robertson<sup>1</sup>, Stephanie A. Schalbetter<sup>2</sup>, Alastair Kerr<sup>1</sup>, David A. Kelly<sup>1</sup>, Jonathan Baxter<sup>2</sup>, Matthew J. Neale<sup>2</sup> & Adele L. Marston<sup>1</sup>✉

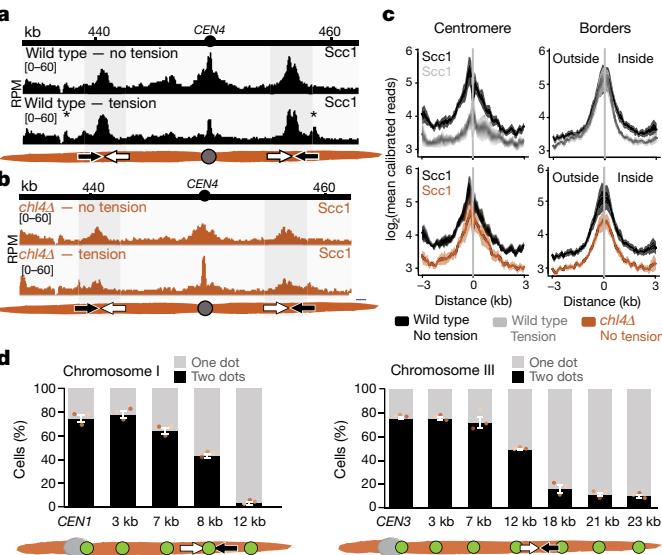
The three-dimensional architecture of the genome governs its maintenance, expression and transmission. The cohesin protein complex organizes the genome by topologically linking distant loci, and is highly enriched in specialized chromosomal domains surrounding centromeres, called pericentromeres<sup>1–6</sup>. Here we report the three-dimensional structure of pericentromeres in budding yeast (*Saccharomyces cerevisiae*) and establish the relationship between genome organization and function. We find that convergent genes mark pericentromere borders and, together with core centromeres, define their structure and function by positioning cohesin. Centromeres load cohesin, and convergent genes at pericentromere borders trap it. Each side of the pericentromere is organized into a looped conformation, with border convergent genes at the base. Microtubule attachment extends a single pericentromere loop, size-limited by convergent genes at its borders. Reorienting genes at borders into a tandem configuration repositions cohesin, enlarges the pericentromere and impairs chromosome biorientation during mitosis. Thus, the linear arrangement of transcriptional units together with targeted cohesin loading shapes pericentromeres into a structure that is competent for chromosome segregation. Our results reveal the architecture of the chromosomal region within which kinetochores are embedded, as well as the restructuring caused by microtubule attachment. Furthermore, we establish a direct, causal relationship between the three-dimensional genome organization of a specific chromosomal domain and cellular function.

To map pericentromere domains, we arrested cells in metaphase in either the presence or the absence of microtubules, and analysed cohesin localization (specifically, the localization of the cohesin sub-unit Scc1) by calibrated chromatin immunoprecipitation followed by sequencing (ChIP-seq). Although cohesin peaks on chromosome arms were comparable with and without microtubules, the signal was reduced over roughly 15 kilobases surrounding centromeres in the presence of microtubule-dependent tension, as reported previously<sup>6–8</sup> (Fig. 1a). The Wpl1/Rad61 protein promotes cohesin turnover before metaphase<sup>9</sup>, but is dispensable for the tension-dependent reduction in pericentromeric cohesin (Extended Data Fig. 1a), suggesting passive removal. Notably, prominent peaks flanking centromeres persisted in the presence of tension, and additional peaks appeared further away from some centromeres (Fig. 1a, asterisks). Pericentromeric cohesin enrichment occurs through specific targeting of cohesin loading to the centromere by a direct interaction between the Ctf19 inner kinetochore subcomplex and the Scc2/Scc4 cohesin loader<sup>10,11</sup>. Existing models posit that cohesin accumulates at positions distinct from its loading sites<sup>12</sup>. Indeed, abolishing kinetochore-driven cohesin loading (by deletion of *CHL4*<sup>5</sup>, encoding a component of the Ctf19 complex) diminished the cohesin peaks flanking centromeres (Fig. 1b), suggesting that some

centromere-loaded cohesin collects at these regions. We henceforth denote these centromere-flanking regions that retain high levels of cohesin under tension and mark the limits of the pericentromere as ‘borders’. Aligning pericentromere borders from all 16 chromosomes, using the centre of the first cohesin peak that persists under tension, confirmed that although cohesin at centromeres is diminished under tension, cohesin at borders or chromosome arm peaks is not, and that Chl4 promotes cohesin association with centromeres and borders, but not with chromosome arms (Fig. 1c and Extended Data Fig. 1b).

Closer inspection of all pericentromere borders revealed the presence of convergent gene pairs—known sites of cohesin accumulation<sup>12</sup>—typically arranged symmetrically around the centromere, and often associated with an additional, centromere-oriented gene on the distal side (Extended Data Fig. 2a, b). Pericentromere size, as measured by distance between borders, ranges from 9.7 kb (chromosome II) to 29.8 kb (chromosome III) with a mean of approximately 17 kb, and shows no correlation with chromosome size (Extended Data Fig. 2c, d). Border convergent gene pairs are more frequently essential than *S. cerevisiae* genes overall and also found at *Candida glabrata* pericentromeres<sup>13</sup> (Extended Data Fig. 2e, f), suggesting a conserved functional arrangement. To determine whether any convergent gene pair has the potential

<sup>1</sup>The Wellcome Centre for Cell Biology, Institute of Cell Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, UK. <sup>2</sup>Genome Damage and Stability Centre, University of Sussex, Brighton, UK. ✉e-mail: adele.marston@ed.ac.uk



**Fig. 1 | Convergent genes mark pericentromere borders.** **a, b,** Representative cohesin (Scc1) enrichment in wild-type ( $n = 3$ ) (**a**) and  $chl4\Delta$  ( $n = 2$ ) (**b**) cells, arrested in metaphase under conditions of tension or no tension. Pericentromere borders are shaded in grey; black and white arrows indicate convergent genes at borders; asterisks indicate additional cohesin peaks under tension. RPM, reads per million mapped (given in square brackets). **c,** Mean calibrated ChIP-seq reads (solid lines), standard errors (dark shading) and 95% confidence intervals (light shading) at all 32 borders and 16 centromeres. **d,** Separation of *tetO*/*TetR*-GFP markers at indicated distances from *CEN1* (left) or *CEN3* (right) in metaphase. For each biological replicate (data points,  $n = 3$ ), 200 cells were scored. Data show means of three replicates, with error bars representing  $\pm$ s.e.m.

for border function, we analysed a strain with an endogenous centromere (*CEN3*) removed and an ectopic centromere (*CEN6*) inserted on the arm of chromosome III<sup>3</sup>. This showed loss of cohesin enrichment at the endogenous pericentromere and borders, and tension-sensitive accumulation of cohesin surrounding ectopic *CEN6* on the arm of chromosome III (Extended Data Fig. 3). Notably, convergent gene pairs flanking the ectopic centromere showed increased cohesin association that persisted under tension, similar to endogenous pericentromere borders (Extended Data Fig. 3).

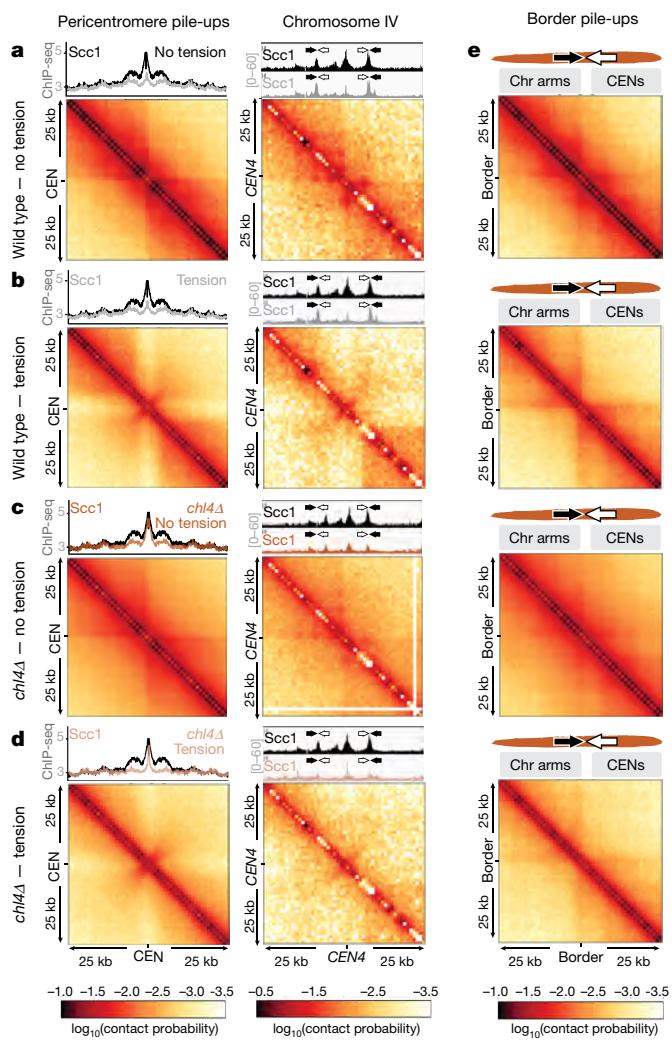
The pericentromeric adaptor protein, shugoshin (Sgo1), promotes the biorientation of sister kinetochores and the proper segregation of chromosomes—in part by recruiting the chromosome-organizing complex condensin to pericentromeres<sup>14,15</sup>—and dissociates from chromosomes in a tension-dependent manner upon biorientation<sup>8</sup>. Pericentromere borders show enrichment of both Sgo1 and condensin (Brn1) (Extended Data Fig. 4a, b); and condensin at borders, but not at core centromeres, is dependent on Sgo1 (Extended Data Fig. 4c, d). Moreover, tension-sensitive Sgo1 resides predominantly at borders (Extended Data Fig. 4e, f), implying that pericentromere borders may elicit the signal that indicates that tension-generating biorientation has been achieved.

Paradoxically, despite the high levels of cohesin, the attachment of sister kinetochores to opposite poles at metaphase causes the separation of sister centromeres, but not of chromosomal arms<sup>16–18</sup>. If borders define the limits of pericentromeres by trapping cohesin to resist the separation of sister chromatids at metaphase, then fluorescent *tetO*/*TetR*-green fluorescent protein (GFP) markers within the pericentromere would be expected to split into two foci at metaphase, whereas markers outside the border should appear as a single focus (Extended Data Fig. 5a). We selected chromosome I, with its clearly delineated border cohesin peaks indicating a small (13.1 kb) pericentromere, and chromosome III, with less defined tension-insensitive

cohesin peaks indicating a large pericentromere (of roughly 29.8 kb), for further analysis (Extended Data Fig. 5b). This difference in pericentromere size predicts differential behaviour of GFP foci at equivalent distances from their centromeres. Indeed, whereas a GFP marker 12 kb from *CEN1* was almost always observed as a single focus at metaphase, a marker 12 kb from *CEN3* frequently split into two foci and, for *CEN3*, single foci prevailed only when markers were positioned 21 kb from the centromere (Fig. 1d and Extended Data Fig. 5c). However, a marker 18 kb from *CEN3*, outside of the annotated border, splits in approximately 16% of cells (Fig. 1d). Similarly, on chromosome I, the second peak of cohesin that persists in the presence of tension appears predominant in border function, because a marker at 7 kb separates in roughly 64% of cells, whereas a marker at 8 kb—within a second, distal cohesin peak—separates in around 43% of cells (Fig. 1d and Extended Data Fig. 5b, c). Therefore, although preferred pericentromere borders exist, they are not fail-safe, and alternative sites of cohesin accumulation lead to cell-to-cell variability in the extent of sister-chromatid separation at metaphase.

Our data suggest that cohesin accumulation at borders defines the domain of chromosomal separation under tension, which we hypothesize defines the structure of the pericentromere. A previous chromosome conformation capture (3C) study<sup>19</sup> observed contacts between the left and right flanks of pericentromere III, and it was suggested to form an intramolecular loop, extending between 11.5 kb and 25 kb. Although this predicted pericentromere size is consistent with our mapping and functional analysis (Fig. 1d and Extended Data Fig. 5), the role of borders remains unclear. To determine pericentromere structure globally and the effect of spindle tension, we performed high-resolution, high-throughput 3C (Hi-C) analysis of metaphase-arrested cells in both the absence (no tension) and the presence (tension) of microtubules to capture unbiased genome-wide interactions. In the absence of tension, consistent with *cis*-looping in mitosis<sup>20,21</sup>, centromere-centred pile-up contact maps of all chromosomes showed a high frequency of *cis* contacts along chromosome arm regions, with core centromeres acting as strong insulators (Fig. 2a, left and Extended Data Fig. 6a). The lower than expected frequency of contacts on the diagonal—between the left and right sides of the centromere—argues against the presence of the previously proposed single intramolecular loop across both sides of the pericentromere<sup>19</sup> (Extended Data Fig. 6b, c). Instead, examination of individual pericentromeres or pile-ups revealed that each side of the core centromere made frequent contacts with the pericentromere on the same side, extending as far as the border, 5–10 kb away (Fig. 2a, right and Extended Data Fig. 7). Both cohesin and condensin can extrude DNA loops *in vitro*<sup>22–24</sup>, and the characteristic Hi-C stripe protruding from the core centromere is suggestive of extrusion of a chromatin loop by a centromere-anchored factor<sup>25,26</sup> (Fig. 2a, right and Extended Data Fig. 7). There is also evidence of longer (20–30 kb) *cis*-looping emanating from directly adjacent to the core centromere into either chromosome arm (Extended Data Fig. 6a). This is consistent with the notion that the usage of convergent gene pairs as boundaries is somewhat stochastic (Fig. 1d). Notably, the strongest Hi-C signal occurs where there is the greatest cohesin density at pericentromere borders (Fig. 2a, Scc1 traces). By contrast, pericentromeric condensin does not appear to be critical for pericentromere structure in the absence of tension. Hi-C maps of *sgo1Δ* mutants (in which pericentromeric condensin is reduced) or *sgo1-3A* mutants (in which, although failing to bind PP2A, condensin is recruited normally)<sup>14,15</sup> showed pericentromeric structures that were virtually indistinguishable from wild type (Extended Data Fig. 8a, b).

The presence of spindle tension changed the conformation of pericentromeres radically, whereas the conformation of chromosome arms was unchanged (Fig. 2b and Extended Data Fig. 6a). Under tension, the centromeres were no longer the point of chromosome arm insulation, and instead border regions formed chromosomal arm loop boundaries roughly 5–10 kb from the core centromeres (Fig. 2b). Inside the borders, the frequency of contacts within, and reaching out of,



**Fig. 2 | The pericentromere is a multilooped structure in mitosis that extends to an open V shape under tension.** Hi-C analysis of wild-type and *chl4Δ* cells arrested in metaphase. **a–d**, Pile-ups (bin size 1 kb) of *cis* contacts 25 kb around all 16 centromeres (left), and contact maps for pericentromere IV (right), for wild-type and *chl4Δ* cells in the absence (**a, c**) or presence (**b, d**) of tension. Above each panel is the median calibrated Scc1 ChIP-seq signal around all 16 centromeres, or for chromosome IV. Arrows denote border convergent genes. **e**, Pile-ups (1-kb bins) of *cis* contacts (25 kb) around the 32 pericentromere borders.

pericentromeres was substantially reduced with a new conformation definable. Contacts across individual centromeres describe an open loop or V-shaped structure, with the core centromere at the apex and the borders at the tips (Fig. 2b, right and Extended Data Fig. 7). Therefore, borders mark the boundary between the pericentromere open loop and the *cis*-loop chromosome arm conformation.

To determine whether cohesin confers boundary function at borders, we analysed *chl4Δ* cells (Fig. 2c, d), in which cohesin enrichment at pericentromere borders is reduced (Fig. 1b, c). Hi-C maps of *chl4Δ* cells in the presence of tension revealed a reduction in both boundary function at borders and the strength of centromere-proximal loops, as shown by less distinct lines and spots on individual chromosomes (Fig. 2d and Extended Data Figs. 6a, 7, particularly evident on chromosomes VII, X and XV), consistent with the increased intersister centromere distance at metaphase in *chl4Δ* cells<sup>5</sup>. In the absence of tension in *chl4Δ*, centromere-proximal loop structures were less defined, indicating that kinetochore-driven cohesin loading promotes *cis*-looping from

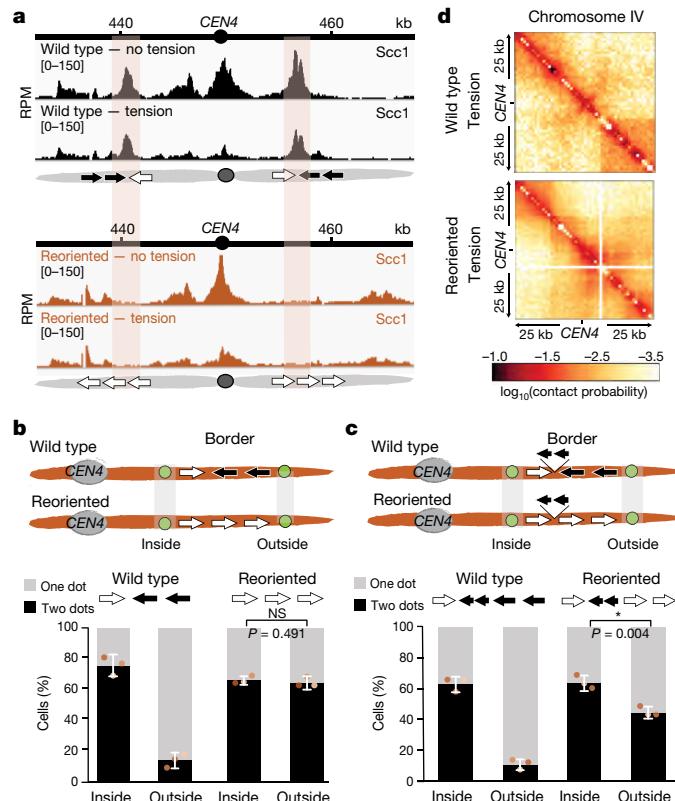
centromeres, as well as establishing boundary function at borders (Fig. 2c and Extended Data Figs. 6a, 7). The maintenance of residual loop structures in *chl4Δ* cells lacking tension is consistent with the presence of residual, non-cohesive cohesin at centromeres of these cells<sup>4,5,11</sup>. Centring the pile-ups on the borders themselves reveals strong isolation of domains proximal and distal to the centromere, which sharpens under tension but is less distinct in *chl4Δ* (Fig. 2e), confirming the boundary function of the borders and their dependence on *CHL4*. Vertical and horizontal lines<sup>25,26</sup> are also suggestive of loop extrusion emanating from borders.

We next investigated which property of borders enables the structural organization of the pericentromere. Because cohesin localization is altered by transcription<sup>12,27,28</sup>, we hypothesized that convergent transcription of border gene pairs leads to cohesin retention, which results in robust intersister-chromatid linkages that isolate domains and resist spindle forces. Indeed, convergent genes at borders show a narrower RNA-sequencing (RNA-seq) density distribution by comparison with all genes, suggesting moderate expression on average (Extended Data Fig. 9a). Analysis of transcriptome-wide RNA polymerase II binding site data<sup>29</sup> revealed that active transcription at convergent gene pairs is typically higher towards, rather than away from, the centromere (Extended Data Fig. 9b). Conversely, convergent genes within the pericentromere show higher expression away from, rather than towards, the centromere (Extended Data Fig. 9c). Consistent with transcription-dependent cohesin positioning<sup>11,27,28</sup>, insertion of a *URA3* cassette between convergent genes at the left border on chromosome IV led to redistribution of cohesin in the direction of transcription (Extended Data Fig. 9d, e).

If convergent genes at borders define pericentromere boundaries, reorienting gene pairs into a tandem arrangement might affect pericentromere behaviour. We engineered such a 'reoriented' strain in which tandem border gene pairs on both sides of *CEN4* transcribe away from the centromere. The reoriented chromosome IV lost cohesin peaks at pericentromere borders, while additional cohesin peaks emerged further away from the centromere, potentially forming new borders (Fig. 3a). Both Sgo1 and condensin (Brn1) associate with the 'new' borders only on reoriented chromosome IV (Extended Data Fig. 10a).

Given that orienting the original border genes in tandem causes other, centromere-distal regions to act as borders, pericentromere size is expected to increase, resulting in an expansion of the region of sister-chromatid separation at metaphase. Consistently, a *tetO*-GFP marker outside of the original border, which infrequently separated in wild-type metaphase cells, was separated in the reoriented strain to a similar extent as a marker inside the original border (Fig. 3b and Extended Data Fig. 10b). However, a marker outside of the new border—roughly 23 kb away from the centromere—was infrequently separated in both wild-type and reoriented strains (Extended Data Fig. 10c). Insertion of a pair of tandemly arranged model genes oriented towards the centromere to restore the convergent gene arrangement at the borders partially rescued separation of the 'outside' marker (Fig. 3c). Therefore, convergent genes set the boundaries at pericentromere borders. Hi-C analysis of the reoriented strain in metaphase in the presence of tension revealed a striking change in structure, specifically of pericentromere IV (Fig. 3d and Extended Data Fig. 10d, e). Boundaries at the original border positions were lost, while centromere-distal regions displayed increased contact frequency, consistent with an expansion of the pericentromeric domain. The typical V shape of endogenous pericentromeres was less apparent: frequent asymmetric contacts were observed within the pericentromere, and the centromere lost its strong insulating effect (Fig. 3d and Extended Data Fig. 10e). Therefore, reorientation of border genes results in a more open, structurally disorganized, pericentromeric structure (Extended Data Fig. 10f).

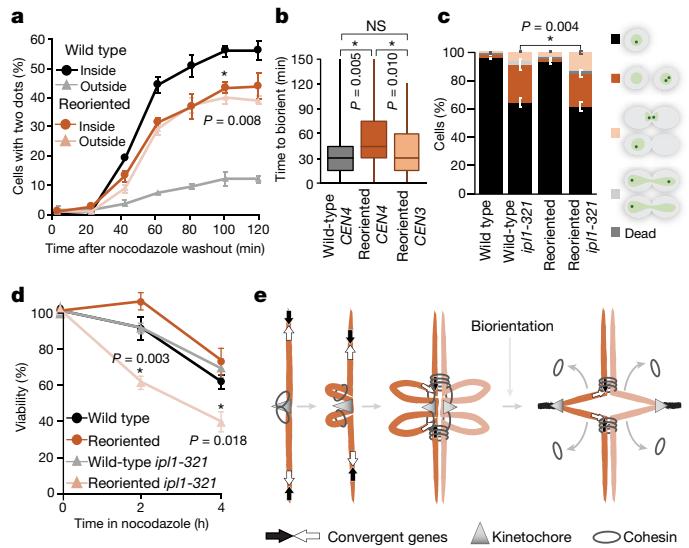
To determine the functional importance of pericentromere boundaries, we assayed sister-chromatid biorientation upon the formation of the metaphase spindle after washing out microtubule-depolymerizing drugs (Extended Data Fig. 10g). Compared with wild-type chromosome IV,



**Fig. 3 | Gene orientation determines pericentromere size.** Border convergent genes on chromosome IV were reversed to a tandem orientation ('reoriented'). **a**, Representative cohesin enrichment in wild-type and reoriented strains ( $n=2$ ). Orange shading indicates the wild-type pericentromere border position. Black and white arrows indicate genes transcribed towards and away from the centromere, respectively. **b**, Separation of *tetO*/TetR-GFP markers (green circles) at the indicated positions in metaphase. **c**, Separation of *tetO*/TetR-GFP markers after insertion of two short model genes downstream of the first border genes on both sides of pericentromere IV, in wild-type and reoriented strains. In **b**, **c**, for each biological replicate (data points,  $n=3$ ), 200 cells were scored. Data show means of three replicates, with error bars representing  $\pm$ s.e.m.; unpaired two-tailed *t*-test; \* $P<0.05$ ; NS,  $P>0.05$ . **d**, Hi-C maps of pericentromere IV ( $n=1$ ).

reoriented chromosome IV showed a delay in, and reduced frequency of, sister kinetochore biorientation (Fig. 4a), consistent with collapse of the V-shape structure indicated by Hi-C analysis. Live-cell imaging confirmed that the biorientation delay was specific to the reoriented chromosome IV, as chromosome III (*CEN3*-GFP) in this strain bioriented with similar efficiency to *CEN4*-GFP in wild-type cells (Fig. 4b). Inefficient sister kinetochore biorientation of the reoriented chromosome might lead to greater reliance on the Aurora B (Ipl1)-dependent error-correction process<sup>30,31</sup>. In the temperature-sensitive *ipl1-321* background under semipermissive conditions, after a single cell cycle, the reoriented chromosome IV strain showed a modest decrease in G1-phase cells that had inherited a single *CEN4*-GFP focus and an accumulation of cells in mitosis (Fig. 4c). Notably, reoriented chromosome IV *ipl1-321* cells, grown at the permissive temperature, showed a pronounced loss of viability after microtubule depolymerization (Fig. 4d). Therefore, convergent genes at borders structure pericentromeres to enable efficient sister kinetochore biorientation and proficient error correction, which is essential for cellular fitness.

We find that targeted cohesin loading at centromeres, and trapping between convergent genes at borders, folds the pericentromere into a multilooped structure (Fig. 4e). This conformation is probably the product of loop extrusion anchored on each side of the centromere, with borders acting to restrict loop size. Consequently, each centromere



**Fig. 4 | Gene orientation affects biorientation efficiency and cell viability.** **a**, Biorientation of sister kinetochores following spindle reassembly. Shown are percentages of cells with separated GFP foci ( $n=200$  at each time point); mean of three biological replicates  $\pm$ s.e.m.; unpaired two-tailed *t*-test; \* $P<0.05$  (P value refers to wild-type inside versus reoriented inside). **b**, Biorientation time (from separation of spindle pole bodies until splitting of centromere-linked GFP foci) in live cells. Centre lines, medians; box limits, second and third quartiles; whiskers, first and fourth quartiles for 120 cells split equally across two biological replicates; two-sided Mann–Whitney test; \* $P<0.05$ . **c**, Categorization of 500 cells on the basis of morphology and GFP foci number following partial *ipl1-321* inactivation. Data show means of three biological replicates  $\pm$ s.e.m.; unpaired two-tailed *t*-test; \* $P<0.05$ . **d**, Cell viability following nocodazole treatment after plating approximately 1,000 cells at each time point. Data are means of three biological replicates  $\pm$ s.e.m.; unpaired two-tailed *t*-test; \* $P<0.05$  (P values refer to wild-type *ipl1-321* versus reoriented *ipl1-321*). **e**, Model of pericentromere structure. Cohesin loaded at kinetochores extrudes a chromatin loop on either side of the centromere until halted by convergent genes at pericentromere borders. When biorientation extends pericentromeric chromatin outwards, intrasister-chromatid cohesin at the base of loops is passively removed from chromosomes, while intersister-chromatid cohesin is trapped at borders, converting centromere-flanking cis-loops to a V-shaped structure.

is isolated from its two flanking pericentromeric regions, providing structural integrity to support the establishment of sister kinetochore biorientation. The resultant pulling forces extend pericentromeric chromatin outwards, until cohesin stalling by convergent transcription at borders prevents further unzipping of the sister chromatids. In the absence of either convergent genes (in the reoriented strain) or efficient cohesin loading at centromeres (in *chlΔ* cells), borders are unable to provide the robust cohesion needed to resist pulling forces at metaphase, and further unzipping occurs.

The suggestion that cohesin makes intrasister-chromatid linkages between two sides of the pericentromere<sup>19</sup> (Extended Data Fig. 6c) is difficult to reconcile with the strong isolation of these regions in the absence of tension (Fig. 2a), or with the observation that cohesin is passively removed within the pericentromere when tension is applied (Fig. 1a). Instead, we favour the idea that while some pericentromeric cohesin entraps sister chromatids to provide cohesion, other cohesin molecules make intrasister-chromatid interactions on either side of the centromere to extrude chromatid loops. While spindle forces will pull chromatin through the intersister-chromatid-entrapping cohesin until they are halted by the transcriptional machinery at borders, intrasister-chromatid loop-extruding cohesin will be evicted from the chromosomes, consistent with passive removal (Fig. 4e).

We have shown here that targeted cohesin loading collaborates with the linear organization of genes to fold a chromosomal domain into a structure that is competent for chromosome segregation. Noncoding transcription and enrichment of cohesin are features of centromeric regions in many organisms, suggesting that general principles may underlie their structure<sup>32</sup>. Potentially, the linear order of transcriptional units throughout a genome has evolved in such a way as to broadly influence its function by locally controlling its architecture.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2244-6>.

1. Rowley, M. J. & Corces, V. G. Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* **19**, 789–800 (2018).
2. Ng, T. M., Waples, W. G., Lavoie, B. D. & Biggins, S. Pericentromeric sister chromatid cohesion promotes kinetochore biorientation. *Mol. Biol. Cell* **20**, 3818–3827 (2009).
3. Weber, S. A. et al. The kinetochore is an enhancer of pericentric cohesin binding. *PLoS Biol.* **2**, E260 (2004).
4. Fernius, J. et al. Cohesin-dependent association of Scc2/4 with the centromere initiates pericentromeric cohesion establishment. *Curr. Biol.* **23**, 599–606 (2013).
5. Fernius, J. & Marston, A. L. Establishment of cohesion at the pericentromere by the Ctf19 kinetochore subcomplex and the replication fork-associated factor, Csm3. *PLoS Genet.* **5**, e1000629 (2009).
6. Eckert, C. A., Gravdahl, D. J. & Megee, P. C. The enhancement of pericentromeric cohesin association by conserved kinetochore components promotes high-fidelity chromosome segregation and is sensitive to microtubule-based tension. *Genes Dev.* **21**, 278–291 (2007).
7. Ocampo-Hafalla, M. T., Katou, Y., Shirahige, K. & Uhlmann, F. Displacement and re-accumulation of centromeric cohesin during transient pre-anaphase centromere splitting. *Chromosoma* **116**, 531–544 (2007).
8. Nerusheva, O. O., Galander, S., Fernius, J., Kelly, D. & Marston, A. L. Tension-dependent removal of pericentromeric shugoshin is an indicator of sister chromosome biorientation. *Genes Dev.* **28**, 1291–1309 (2014).
9. Lopez-Serra, L., Lengronne, A., Borges, V., Kelly, G. & Uhlmann, F. Budding yeast Wapl controls sister chromatid cohesion maintenance and chromosome condensation. *Curr. Biol.* **23**, 64–69 (2013).
10. Hinshaw, S. M., Makrantoni, V., Kerr, A., Marston, A. L. & Harrison, S. C. Structural evidence for Scc4-dependent localization of cohesin loading. *eLife* **4**, e06057 (2015).
11. Hinshaw, S. M., Makrantoni, V., Harrison, S. C. & Marston, A. L. The kinetochore receptor for the cohesin loading complex. *Cell* **171**, 72–84 (2017).
12. Lengronne, A. et al. Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* **430**, 573–578 (2004).
13. Hu, B. et al. Biological chromodynamics: a general method for measuring protein occupancy across the genome by calibrating ChIP-seq. *Nucleic Acids Res.* **43**, e132 (2015).
14. Verzijlbergen, K. F. et al. Shugoshin biases chromosomes for biorientation through condensin recruitment to the pericentromere. *eLife* **3**, e01374 (2014).
15. Peplowska, K., Wallek, A. U. & Storchová, Z. Sgo1 regulates both condensin and Ipl1/Aurora B to promote chromosome biorientation. *PLoS Genet.* **10**, e1004411 (2014).
16. He, X., Asthana, S. & Sorger, P. K. Transient sister chromatid separation and elastic deformation of chromosomes during mitosis in budding yeast. *Cell* **101**, 763–775 (2000).
17. Goshima, G. & Yanagida, M. Establishing biorientation occurs with precocious separation of the sister kinetochores, but not the arms, in the early spindle of budding yeast. *Cell* **100**, 619–633 (2000).
18. Tanaka, T., Fuchs, J., Loidl, J. & Nasmyth, K. Cohesin ensures bipolar attachment of microtubules to sister centromeres and resists their precocious separation. *Nat. Cell Biol.* **2**, 492–499 (2000).
19. Yeh, E. et al. Pericentric chromatin is organized into an intramolecular loop in mitosis. *Curr. Biol.* **18**, 81–90 (2008).
20. Schalbetter, S. A. et al. SMC complexes differentially compact mitotic chromosomes according to genomic context. *Nat. Cell Biol.* **19**, 1071–1080 (2017).
21. Lazar-Stefanita, L. et al. Cohesins and condensins orchestrate the 4D dynamics of yeast chromosomes during the cell cycle. *EMBO J.* **36**, 2684–2697 (2017).
22. Ganji, M. et al. Real-time imaging of DNA loop extrusion by condensin. *Science* **360**, 102–105 (2018).
23. Davidson, I. F. et al. DNA loop extrusion by human cohesin. *Science* **366**, 1338–1345 (2019).
24. Kim, Y., Shi, Z., Zhang, H., Finkelstein, I. J. & Yu, H. Human cohesin compacts DNA by loop extrusion. *Science* **366**, 1345–1349 (2019).
25. Vian, L. et al. The energetics and physiological impact of cohesin extrusion. *Cell* **173**, 1165–1178 (2018).
26. Fudenberg, G. et al. Formation of chromosomal domains by loop extrusion. *Cell Rep.* **15**, 2038–2049 (2016).
27. Bausch, C. et al. Transcription alters chromosomal locations of cohesin in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **27**, 8522–8532 (2007).
28. Ocampo-Hafalla, M., Muñoz, S., Samora, C. P. & Uhlmann, F. Evidence for cohesin sliding along budding yeast chromosomes. *Open Biol.* **6**, 150178 (2016).
29. Bresson, S., Tuck, A., Staneva, D. & Tollervey, D. Nuclear RNA decay pathways aid rapid remodeling of gene expression in yeast. *Mol. Cell* **65**, 787–800 (2017).
30. Biggins, S. & Murray, A. W. The budding yeast protein kinase Ipl1/Aurora allows the absence of tension to activate the spindle checkpoint. *Genes Dev.* **15**, 3118–3129 (2001).
31. Tanaka, T. U. et al. Evidence that the Ipl1-Sli15 (Aurora kinase-INCENP) complex promotes chromosome bi-orientation by altering kinetochore-spindle pole connections. *Cell* **108**, 317–329 (2002).
32. Perea-Resa, C. & Blower, M. D. Centromere biology: transcription goes on stage. *Mol. Cell. Biol.* **38**, e00263–18 (2018).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

# Article

## Methods

No statistical methods were used to predetermine sample size.

### Yeast strains and plasmids

All yeast strains were derivatives of W303 and are listed in Supplementary Table 1. Plasmids generated here are listed in Supplementary Table 2. For calibrated ChIP-seq we used *Schizosaccharomyces pombe* strain spAM635 (h-*rad21-6HA::KanMX6*), where 6HA refers to 6× haemagglutinin tag. The yeast strain carrying chromosome III with an ectopic centromere was described previously<sup>3</sup>. To visualize chromosomal loci, we integrated *tetO* sequences at defined sites on chromosomes I, III and IV after cloning of the appropriate region into *pRS306(tetOx224)* (Supplementary Table 2). We inserted *URA3* between convergent gene pairs using a polymerase chain reaction (PCR)-directed approach. To reorient potential border genes on chromosome IV, we cloned the gene cassette including its promoter into a plasmid (Supplementary Table 2), upstream of *KanMX*, which was flanked by *LoxP* sites. Plasmids were used as templates for PCR, the products of which were used for yeast transformation, resulting in replacement of a gene and its promoter with the same gene in the reversed orientation, together with a *LoxP-KanMX6-LoxP* downstream marker. Insertion in the desired orientation was confirmed by PCR. The *KanMX6* marker was then excised by Cre-mediated recombination. Plasmids for rescue constructs were assembled using five-fragment Gibson assembly, and the resulting *pURA3::ABlx2-V5::TRP1* and *pURA3::PYLx2-FLAG::HISMX6* constructs were inserted at chromosome IV pericentromere borders by a PCR-directed approach.

### Growth conditions

Cells carrying *pMET-CDC20* were arrested in metaphase in the presence and absence of tension as described previously<sup>8</sup>. In brief, cultures were arrested in G1 in synthetic medium lacking methionine (SC-/Met/D) with alpha factor ( $5\text{ }\mu\text{g ml}^{-1}$ ) for 1.5 h, before readding alpha factor to  $2.5\text{ }\mu\text{g ml}^{-1}$  and shaking for a further 1.5 h. Cells were washed with rich medium lacking glucose (YEP) and released into rich medium containing  $8\text{ }\mu\text{M}$  methionine (YPDA/Met). Methionine was readded at  $4\text{ }\mu\text{M}$  every hour. To achieve a metaphase arrest in the absence of microtubules (no tension), cells were released from G1 into YPDA/Met medium containing  $15\text{ }\mu\text{g ml}^{-1}$  nocodazole and  $30\text{ }\mu\text{g ml}^{-1}$  benomyl. Nocodazole was readded at  $7.5\text{ }\mu\text{g ml}^{-1}$  every hour. For both tension and no tension (nocodazole) conditions, cells were harvested 2 h after release from G1. For biorientation assays, cells were arrested in the absence of tension as above; after 2 h, nocodazole was washed out by filtering with rich medium lacking glucose, before cultures were released into YPDA+Met to allow spindles to reform while maintaining the metaphase arrest. Samples were taken at 20-min intervals and scored blind. To arrest cells lacking the *pMET-CDC20* construct in metaphase in the absence of spindle tension, cycling cells (optical density at 600 nm ( $\text{OD}_{600}$ ) = 0.2) were treated with  $15\text{ }\mu\text{g ml}^{-1}$  nocodazole and  $30\text{ }\mu\text{g ml}^{-1}$  benomyl; every hour,  $7.5\text{ }\mu\text{g ml}^{-1}$  nocodazole was added and cells were harvested after a total of 3 h. For partial *ipl1-321* inactivation, cultures were arrested in G1 at room temperature with alpha factor ( $5\text{ }\mu\text{g ml}^{-1}$ ) for 1.5 h, before readding alpha factor to  $2.5\text{ }\mu\text{g ml}^{-1}$  and shaking for a further 1.5 h. Cells were washed with rich medium lacking glucose (YEP), released into rich medium (YPDA) prewarmed to 32 °C and samples were collected after 2 h. Alpha factor was readded to block cells in the next G1 phase when small buds started to appear. For viability assays, cultures in exponential phase were diluted to  $\text{OD}_{600} = 0.2$ , and for each condition approximately 1,000 cells were plated over 6 YPDA plates before (0 h) and 2 h or 4 h after the addition of  $15\text{ }\mu\text{g ml}^{-1}$  nocodazole. Nocodazole was readded at  $7.5\text{ }\mu\text{g ml}^{-1}$  every 90 min. Cells on plates were grown at 25 °C for 2 days, and then the numbers of colonies were scored. Viability was calculated as a percentage of the colony count at 0 h.

### ChIP-qPCR, ChIP-seq and data analysis

ChIP-quantitative PCR (qPCR) and ChIP-seq were carried out as described<sup>14</sup>, except that cells were fixed for 30 min, and that DNA from purified chromatin was recovered using a PCR purification kit (Promega). Sequencing libraries were generated using standard methods and samples were sequenced on a MiniSeq instrument (Illumina), with the exception of the data shown in Extended Data Figs. 3, 4c–f, for which libraries were prepared and sequenced by the EMBL Genomics Core Facility (Heidelberg, Germany). ChIP-seq data used to generate Extended Data Fig. 4c, d were published previously<sup>14</sup>. Scripts, data files and workflows used to analyse the data and prepare the ChIP-seq figures can be found on the GitHub repository at <https://github.com/AlastairKerr>. For the strains in which the centromere was repositioned, or the *URA3* cassette was inserted at borders, or the gene orientation at pericentromere borders was reversed, the corresponding reference genome sequence was assembled in silico and the appropriate reference was used to map sequencing reads for each strain. Plots showing averages of all centromeres were generated using Seqplots<sup>33</sup>. Read counts were normalized to reads per million mapped (RPM) and the ratio of ChIP reads to input was calculated. The mean or the median value was determined for all 16 chromosomes per 50-base-pair window and its  $\log_2$  value is graphed. Mean values are shown for the  $\pm 3\text{-kb}$  plots; the  $\pm 25\text{-kb}$  plots use median values. Reference genome coordinates used to centre Seqplots are given in Supplementary Table 3.

To allow quantitative comparison between different conditions, we calibrated all ChIP-seq analyses, with the exception of the data shown in Extended Data Figs. 3, 4c–f, with an internal reference by modifying the procedure described in ref.<sup>13</sup>. Instead of *C. glabrata*, we used *S. pombe* carrying Rad21-6HA as the calibration genome. In brief, for each immunoprecipitation, we grew 100 ml of *S. pombe* cells in YES to  $\text{OD}_{595} = 0.25\text{--}0.3$ , and fixed them by adding a 1/10 volume of 11% formaldehyde in diluent (0.143 M NaCl, 1.43 mM EDTA, 71.43 mM HEPES-KOH) with gentle agitation for 30 min. Cell pellets were washed twice with 10 ml cold TBS (20 mM Tris-HCl, pH 7.5, 150 mM NaCl) and once with 10 ml cold FA lysis buffer (100 mM Hepes-KOH, pH 7.5, 300 mM NaCl, 2 mM EDTA, 2% Triton X-100, 0.2% sodium deoxycholate)/0.1% SDS, frozen in liquid nitrogen and stored at -80 °C. *S. pombe* cell pellets were resuspended in 400 µl of cold 1× FA lysis buffer/0.5% SDS containing 1× complete protease-inhibitor cocktail (Roche) and 1 mM phenylmethanesulfonyl fluoride (PMSF), and mixed with thawed *S. cerevisiae* pellet (approximately 100 ml of cells at  $\text{OD}_{600} = 0.4$ ). ChIP and sequencing were performed as described above. Calculation of occupancy ratios and data analysis were performed as described<sup>13</sup>. The number of reads at each position was normalized to the total number of reads for each sample, multiplied by the occupancy ratio and shown in the Integrated Genome Viewer from the Broad Institute. For ChIP-seq, data from a representative experiment are shown, with  $n$  referring to the number of biological replicates performed overall. Data were comparable in all cases. Cohesin ChIP-seq in wild-type and reoriented strains (Fig. 3a) was performed in three biological replicates, using different epitopes for immunoprecipitation and different strain genotypes. Primers used for qPCR analysis are given in Supplementary Table 4.

### Microscopy

Cells were fixed in formaldehyde for visualization of TetR-GFP and Spc42-tdTomato foci. Yeast were mounted onto a glass slide and imaged on a Zeiss Axio Imager Z1 equipped with a  $\times 100\text{ }\alpha$  Plan Fluar/1.45 NA (oil) objective lens. Images were recorded using a Photometrics Evolve electron-multiplying charge-coupled device (EMCCD) camera (Photometrics, Tucson, USA) controlled using MicroManager 1.4 acquisition software (US National Institutes of Health). The fluorescent intensity and distance between the GFP foci were measured using a custom ImageJ plugin that can be found on the GitHub repository at <a href="https://github.com/dkelly604/CellClicker\_.Live-cell imaging was performed</a>

on a Zeiss Axio Observer Z1 (Zeiss UK, Cambridge) equipped with a Hamamatsu Flash 4 sCMOS camera, a Prior motorized stage and Zen 2.3 acquisition software. Cells were imaged at 25 °C using CellASIC ONIX microfluidics plates, with images taken at 15-min intervals.

### RNA isolation and RNA-seq

Cell pellets were lysed by bead-beating in RLT buffer (Qiagen) and RNA was isolated using the RNeasy mini kit (Qiagen) according to the manufacturer's instructions, except that on-column DNA digestion was performed using the Qiagen DNase digestion kit. RNA concentration was determined by nanodrop. To synthesize complementary DNAs for qRT-PCR, we incubated 12 ng purified total RNA (diluted in HyClone dH<sub>2</sub>O) and 10 mM oligo(dT)15 primer (Roche) or 1.5 mM gene-specific reverse primer at 65 °C for 10 min before placement on ice to denature RNA. Subsequently, we added 4 µl 5× Transcriptor reverse-transcription reaction buffer (Roche), 0.5 µl RNase OUT (Fisher), 1 µM deoxynucleoside triphosphate (dNTP) and 0.5 µl Transcriptor reverse transcriptase plus Hyclone dH<sub>2</sub>O to a volume of 20 µl, and incubated the samples at 55 °C for 3 h before heat-inactivating the reverse transcriptase at 85 °C for 5 min. RNA was depleted of ribosomal RNA and libraries were prepared for sequencing by Genecore, EMBL. Sequencing was also performed by Genecore on an Illumina Next Seq 500 with a read-l length of 75, multiplexed with a pool size of 4.

### Hi-C library preparation and data analysis

Our Hi-C protocol was modified from refs.<sup>20,34,35</sup>. Cells were cultured, fixed and lysed as described<sup>34</sup>. In brief, 200 ml of cells (at an OD<sub>600</sub> of roughly 0.6) carrying *pMET-CDC20* were arrested in metaphase at 25 °C in the presence and absence of tension as above, and fixed with 3% formaldehyde for 20 min at 25 °C at 250 r.p.m.; the reaction was quenched for 5 min by adding 0.35 M glycine (final concentration). Cells were washed with cold water, resuspended in 5 ml 1× NEBuffer 2 and frozen in liquid nitrogen. Lysates were prepared by grinding the frozen pellet in a chilled mortar with a pestle for 15 min, and one-tenth of the initial pellet weight (roughly 0.5 g) was taken for further processing. Digestion with a restriction enzyme (*DpnII*), filling-in, ligation, crosslink reversal, DNA concentration and purification, and biotin removal were carried out as described<sup>35</sup>. DNA was then fragmented on a Bioruptor Plus sonication device (Diagenode) for a total of 2×30 cycles, with 30 s on/off, at high setting. Following DNA end repair and A-tailing using T4 DNA polymerase, T4 polynucleotide kinase and Klenow fragment DNA polymerase I (as in ref.<sup>35</sup>), Hi-C libraries were fractionated using Ampure XP beads as described<sup>34</sup>. Biotin pull-down, adaptor ligation (NextFlex, Bioo Scientific) and sequencing (EMBL Core Genomics Facility, Heidelberg, Germany) were carried out as in ref.<sup>20</sup>. Hi-C read numbers are given in Supplementary Table 5.

For Hi-C data analysis, Fastq reads were aligned to the sacCer3 or the 'reoriented' reference genome using HiC-Pro v2.11.1 (ref.<sup>36</sup>) bowtie2 v2.3.4.1 (–very-sensitive –L 30-score-min L, -0.6, -0.2-end-to-end-reorder), removing singleton, multi hit and duplicated reads. Read pairs were assigned to restriction fragment (*DpnII*) and invalid pairs filtered out. Valid interaction pairs were converted into the .cool contact matrix format using the cooler library, and matrices balanced using iterative correction down to 1-kb resolution. Multiresolution cool files were uploaded onto a local HiGlass<sup>37</sup> server for visualization; 'cooler show' was also used to generate individual plots for each chromosome. White stripes on plots represent regions in which data were lost stochastically during mapping owing to stringency settings filtering out reads. In the case of the reoriented chromosome, the presence of two to three small LoxP 'scars' are likely to affect mapping in these regions and may account for the observed data loss. To generate pile-ups at

centromeres/pericentromeric borders, we used the cooltools library; cool matrixes were binned at 1-kb resolutions. Plots were created around the midpoints of centromeres with 10/25/100-kb flanks on each side, or around the midpoint of borders with 25-kb flanks, showing the log<sub>10</sub> mean interaction frequency with a colour map similar to HiGlass 'fall'. All centromere/pericentromere annotations were duplicated in the forward and reverse strand orientations to create an image that is mirror symmetrical. Ratios of pile-ups between samples were created in a similar fashion, plotting the log<sub>2</sub> difference between samples in the 'coolwarm' colour map: that is, A/B; red signifies increased contacts in A relative to B, and blue signifies decreased contacts in B relative to A. Scripts are available at [https://github.com/danrobertson87/Paldi\\_2019](https://github.com/danrobertson87/Paldi_2019).

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

### Data availability

Sequencing datasets are available at the Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>), accession number GSE104135. Source data for Figs. 1, 3, 4 and Extended Data Figs. 2, 5, 9, 10 are provided with the paper.

### Code availability

The ImageJ plugin for measuring the fluorescent intensity and distance between GFP foci can be found on the GitHub repository at [https://github.com/dkelly604/CellClicker\\_](https://github.com/dkelly604/CellClicker_). Scripts for Hi-C data analysis are available at [https://github.com/danrobertson87/Paldi\\_2019](https://github.com/danrobertson87/Paldi_2019).

33. Stempor, P. & Ahringer, J. Seqplots—interactive software for exploratory data analyses, pattern discovery and visualization in genomics. *Wellcome Open Res.* **1**, 14 (2014).
34. Belton, J.-M. & Dekker, J. Hi-C in budding yeast. *Cell Spring Harv. Protoc.* <https://doi.org/10.1101/pdb.prot085209> (2015).
35. Schalbetter, S. A., Fudenberg, G., Baxter, J., Pollard, K. S. & Neale, M. J. Principles of meiotic chromosome assembly revealed in *S. cerevisiae*. *Nat. Commun.* **10**, 4795 (2018).
36. Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 259 (2015).
37. Kerpedjiev, P. et al. HiGlass: web-based visual exploration and analysis of genome interaction maps. *Genome Biol.* **19**, 125 (2018).

**Acknowledgements** We thank B. Baying and V. Beres (Genecore, EMBL) for next-generation sequencing and library preparation; the Wellcome Centre for Cell Biology bioinformatics core for computational support; the Centre Optical Imaging Laboratory for support with microscopy; and W. Borek for help with statistical analysis. We thank R. Allshire, S. Bresson and D. Tollervey for helpful discussions; P. Megee for yeast strains; and W. Borek, S. Hinshaw, V. Makrantioti and P. Romé for comments on the manuscript. This study was funded by a Wellcome Senior Research Fellowship (grant 107827 to A.L.M., B.A. and F.P.); a Wellcome PhD studentship (109091 to F.P.); core funding for the Wellcome Centre for Cell Biology (203149 to A.L.M., F.P., B.A., D.R., A.K. and D.A.K.); a European Research Council (ERC) Consolidator Award (311336 to M.J.N. and S.A.S.); and a Wellcome Trust Investigator Award (200843 to M.J.N. and S.A.S.).

**Author contributions** A.L.M., F.P. and B.A. conceived the study. F.P. performed Hi-C, ChIP-seq, microscopy and viability experiments. B.A. performed ChIP-seq and ChIP-qPCR experiments. F.P., B.A. and A.L.M. generated reagents and analysed data. D.R., A.K. and S.A.S. performed bioinformatics analyses. S.A.S. developed protocols. D.A.K. wrote custom code for image analysis. F.P., A.L.M., S.A.S., J.B. and M.J.N. interpreted Hi-C data. A.L.M. supervised the study. A.L.M. and F.P. wrote the paper, with input from all authors.

**Competing interests** The authors declare no competing interests.

### Additional information

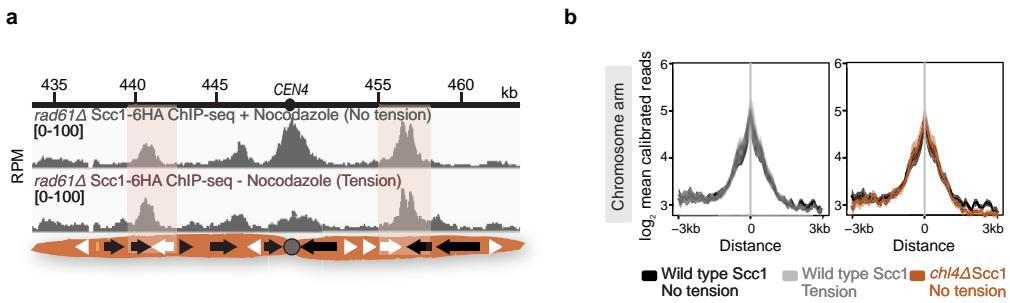
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2244-6>.

**Correspondence and requests for materials** should be addressed to A.L.M.

**Peer review information** *Nature* thanks Zhijun Duan, Katsuhiko Shirahige and Hongtao Yu for their contribution to the peer review of this work.

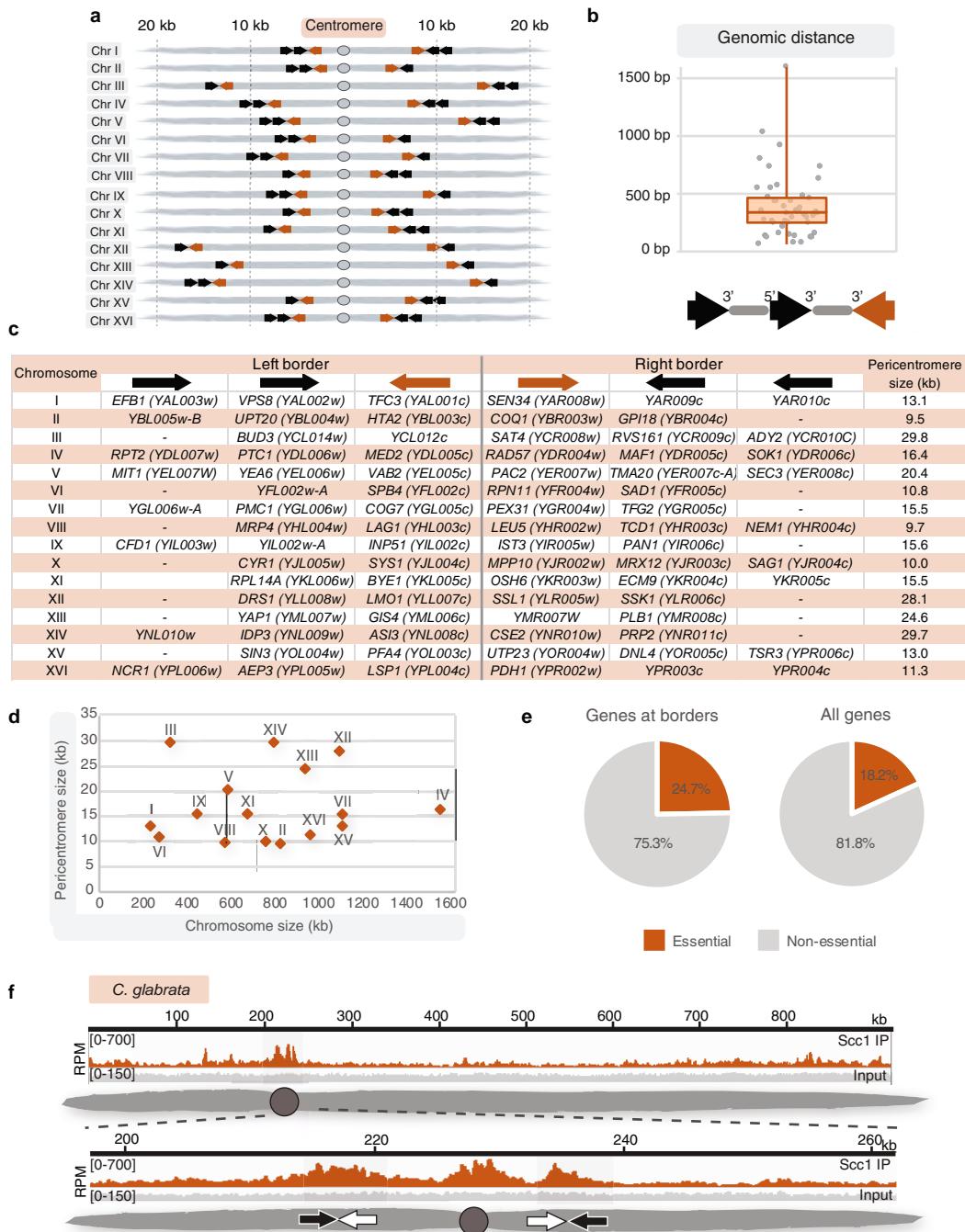
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

# Article



**Extended Data Fig. 1 | Tension-dependent cohesin removal at metaphase is restricted to the pericentromere and occurs independently of Wpl1/Rad61.** **a**, Scc1-6HA-calibrated ChIP-seq profiles for the pericentromeric region of chromosome IV are shown for *rad61 $\Delta$*  cells arrested in metaphase, in the absence and presence of spindle tension ( $n=1$ ). **b**, Mean calibrated ChIP

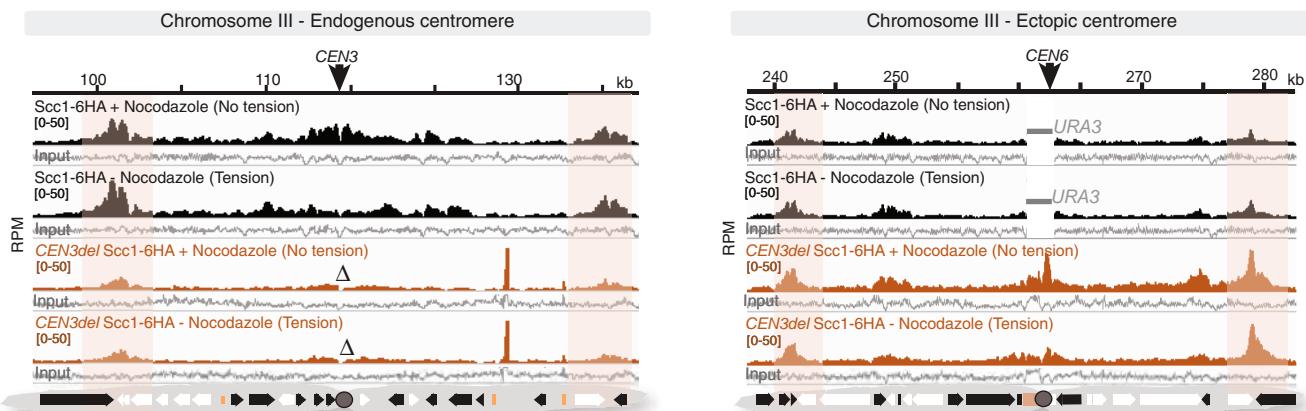
reads (solid lines), standard errors (dark shading) and 95% confidence intervals (light shading) at a pericentromere-proximal cohesin site on each chromosome arm ( $n=32$  sites), for wild-type and *chl4 $\Delta$*  cells, in either the presence or the absence of tension.



**Extended Data Fig. 2 | Overview of border gene organization and pericentromere size.** **a**, Diagram showing the positions of convergent gene pairs flanking centromeres. Grey ovals represent the centromeres; convergent gene pairs at borders are indicated by arrows. **b**, Genomic distances (thick grey lines) between the 3' ends of convergent genes (black and orange arrows) at borders, as well as the 3' and 5' ends of the two genes transcribed towards centromeres (black arrows) ( $n=49$ ). Centre line, median; box limits, second and third quartiles; whiskers, first and fourth quartiles. **c**, Table of convergent

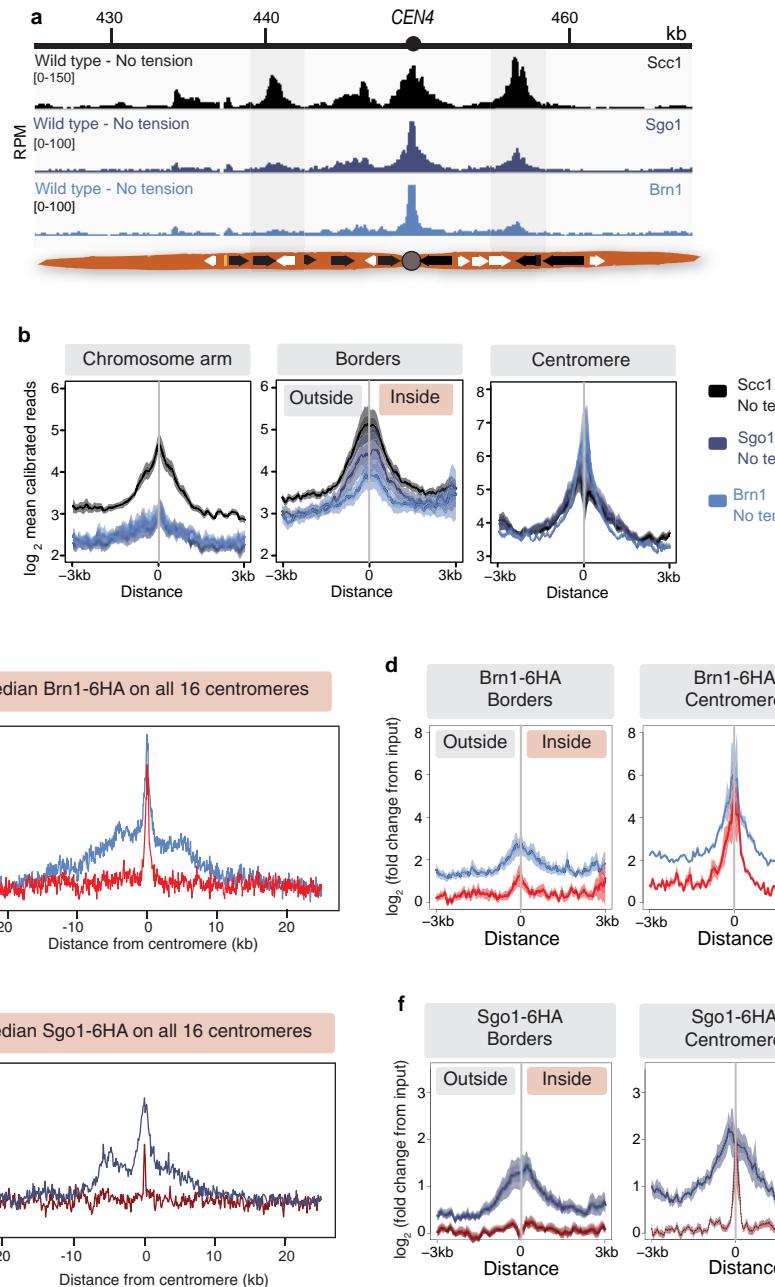
genes identified at pericentromere borders for each chromosome, along with the corresponding pericentromere size. A border was defined as the innermost cohesin peak near the centromere that persisted in the presence of tension. **d**, Pericentromere sizes determined in **c** are plotted against chromosome size. **e**, Percentages of genes essential for growth on rich glucose media among genes at borders and among all genes. **f**, Scc1 ChIP-seq in asynchronous *C. glabrata* cells, showing chromosome F from ref.<sup>13</sup>.

# Article



**Extended Data Fig. 3 | An ectopic centromere establishes new borders at convergent genes on a chromosome arm.** Shown are cohesin (Scc1) ChIP-seq profiles for the region surrounding the endogenous centromere on chromosome III (left), and for a roughly 50-kb region of chromosome III

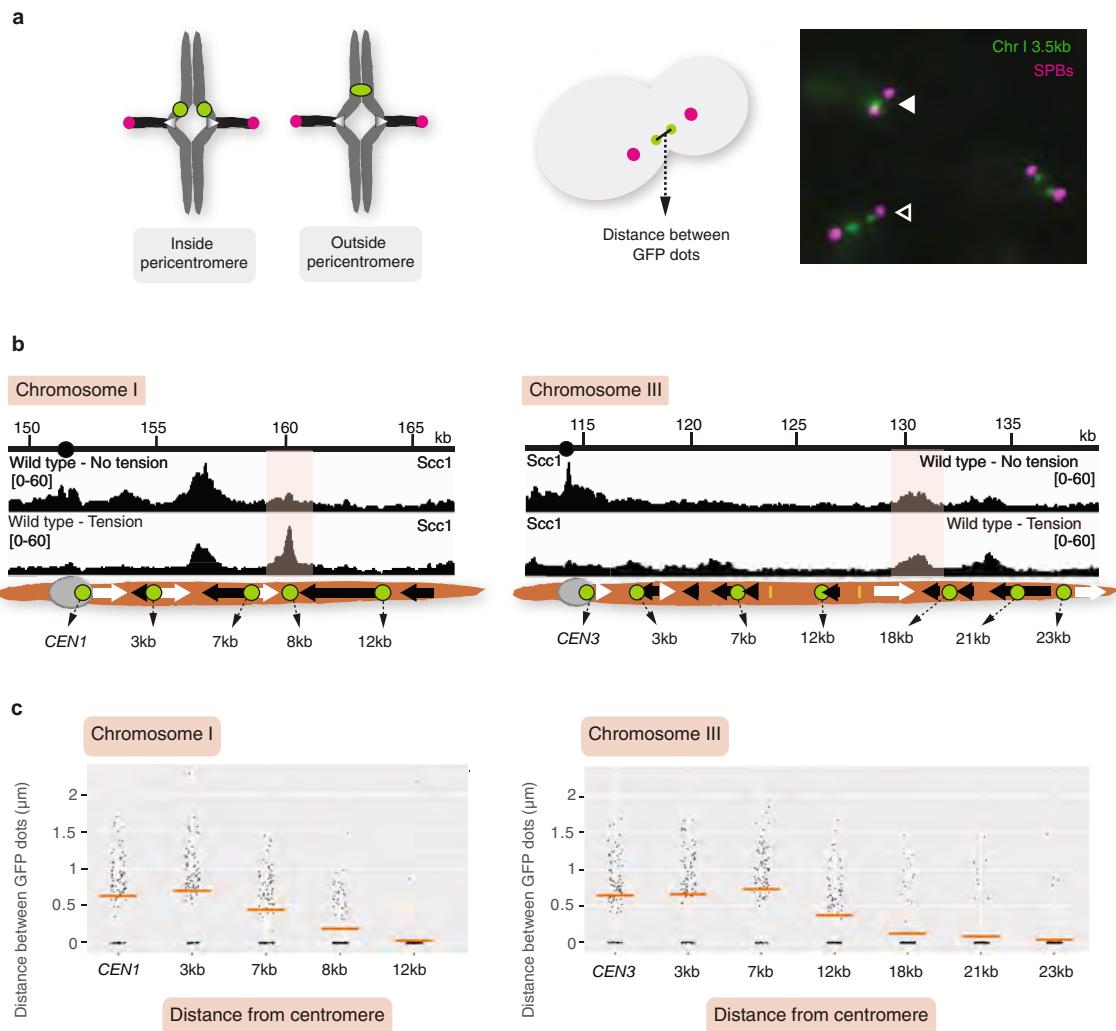
surrounding the ectopic centromeric arm site (right) ( $n=1$ ). Regions of tension-insensitive cohesin peaks at convergent sites flanking the endogenous and ectopic centromeres are highlighted.



**Extended Data Fig. 4 | Shugoshin and condensin localize to pericentromere borders.** **a**, Representative cohesin (Scc1), shugoshin (Sgo1) and condensin (Brn1) enrichment in metaphase-arrested cells in the presence of nocodazole in the pericentromeric region of chromosome IV ( $n=2$ ; immunoprecipitation was performed using proteins tagged with different epitopes in the two biological replicates). **b**, Plots show mean calibrated ChIP reads (solid lines), standard errors (dark shading) and 95% confidence intervals (light shading) at all 32 borders and 16 centromeres. For comparison, similar plots for the next convergent gene site on each chromosome arm are shown. **c, d**, Condensin associates with pericentromere borders in a Sgo1-dependent manner in cells arrested in metaphase in the absence of tension. **c**, These ChIP-seq data have

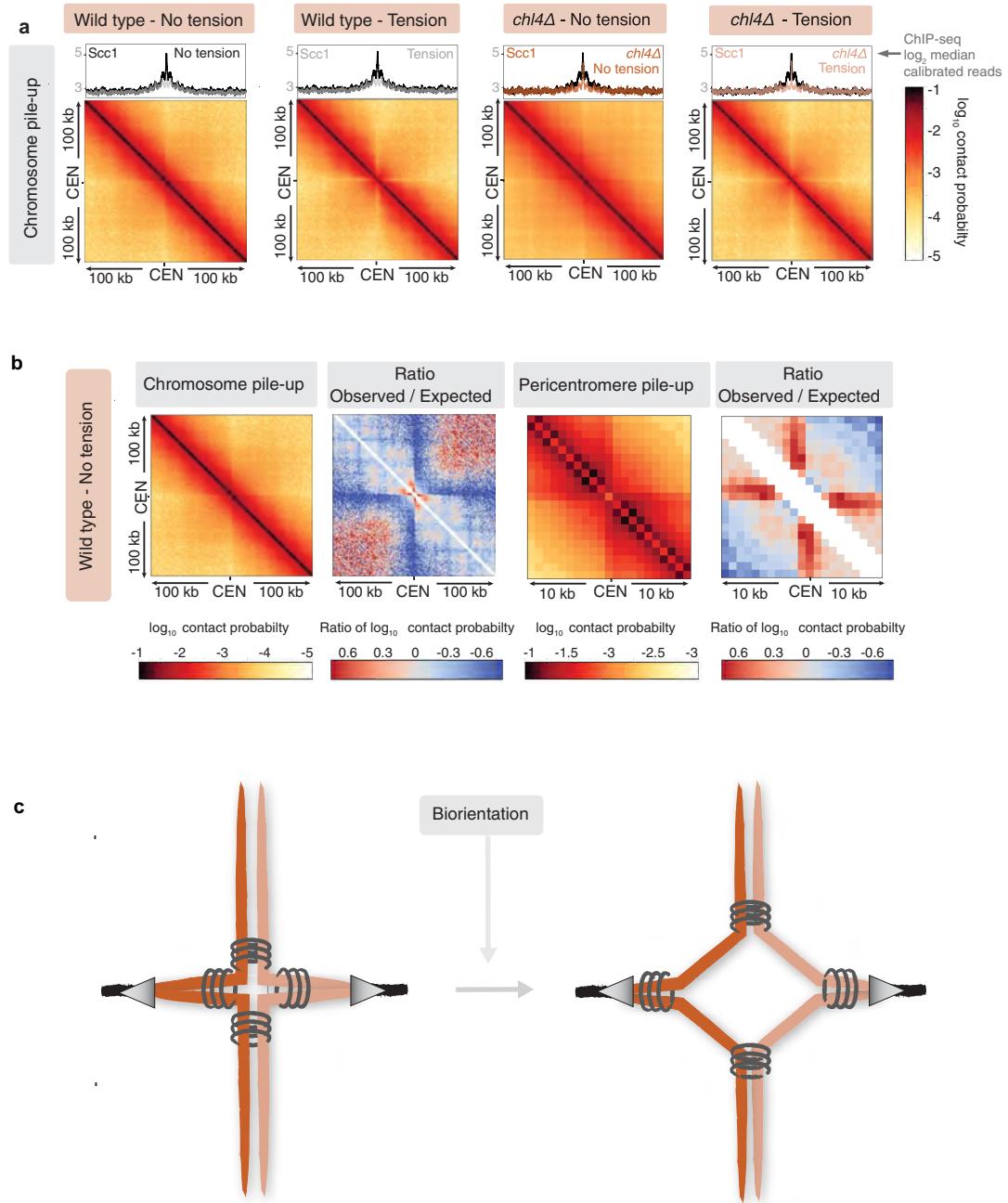
been published previously<sup>14</sup>, and show the median condensin (Brn1) signal across a 50-kb region surrounding all 16 centromeres. **d**, The mean Brn1 signal centred around all 32 borders (left) or 16 centromeres (right). Sgo1 is removed from the borders, but not from core centromeres, in response to spindle tension (solid lines show medians; dark shading shows standard errors; light shading shows 95% confidence intervals). **e**, Median Sgo1 enrichment by ChIP-seq, plotted over a 50-kb region surrounding all 16 centromeres in metaphase-arrested cells in the presence or absence of tension. **f**, Mean Sgo1 signal centred around all 32 borders (left) or 16 centromeres (right) (solid lines, medians; dark shading, standard errors; light shading, 95% confidence intervals).

# Article



**Extended Data Fig. 5 | Pericentromere borders resist sister-chromatid separation under tension.** **a**, Assay to measure separation of loci on sister chromatids in metaphase-arrested cells. Cells carry *tetO*/TetR-GFP foci (green) integrated at various positions and Spc42-tdTomato foci (pink) to mark spindle-pole bodies (SPBs); the cells are arrested in metaphase by Cdc20 depletion. The diagram at the left shows the expected separation of GFP foci inside and outside pericentromere loci; this distance is measured as described

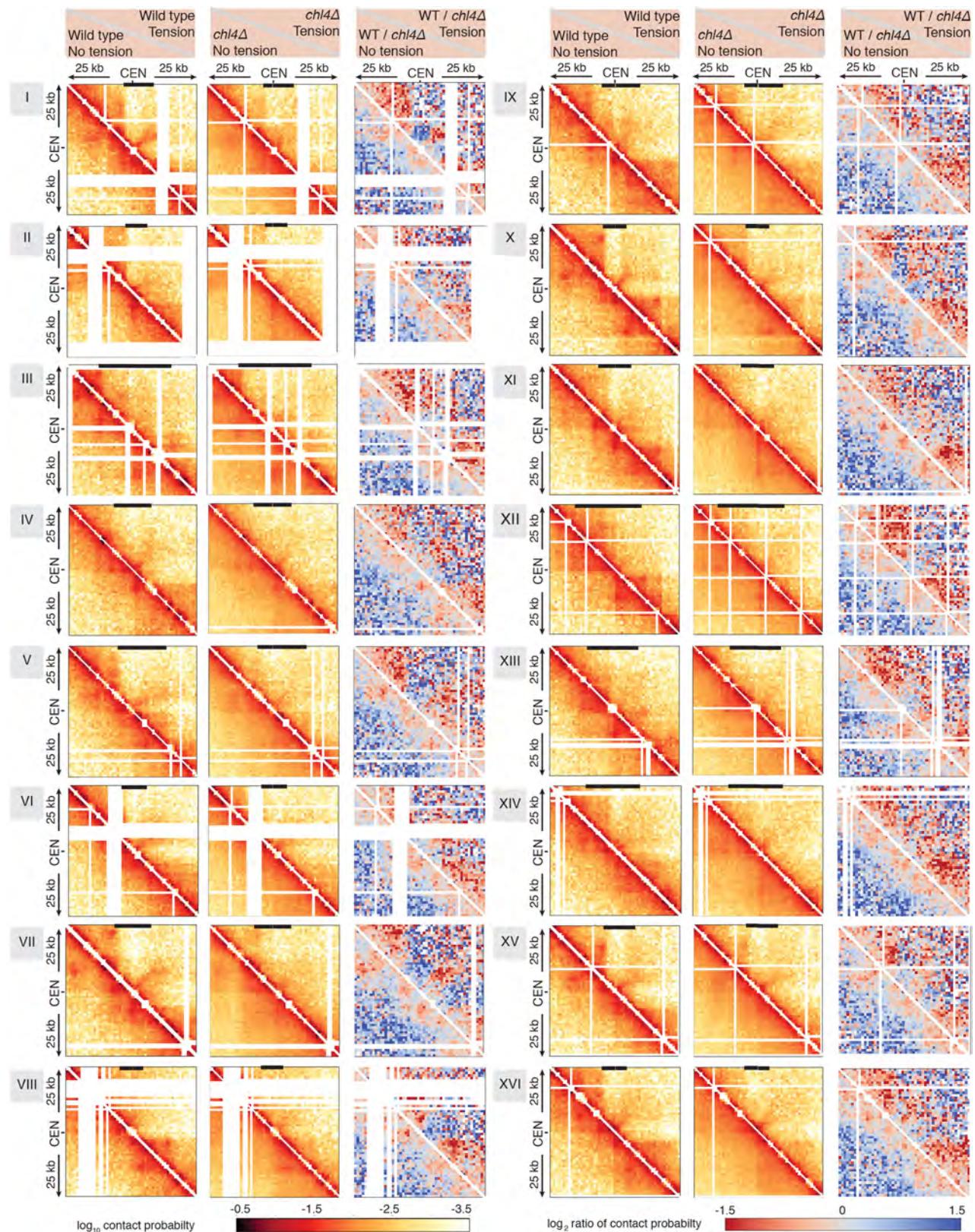
in the Methods. A representative image is shown on the right ( $n=3$  biological replicates). Filled and open arrowheads mark cells with a single GFP focus or split foci, respectively. **b**, Positions of GFP foci and corresponding calibrated Scc1-6HA ChIP-seq profiles ( $n=3$ ) for chromosomes I and III. **c**, Cells carrying *tetO*/TetR-GFP foci integrated at various positions were arrested in metaphase and the distance between GFP dots was measured in 100 cells. Horizontal lines indicate means.



**Extended Data Fig. 6 | Cis-looping and alternative model for pericentromeric chromosome conformation.** **a**, Pile-ups (bin size 1 kb) of *cis* contacts located in the 100 kb around all 16 centromeres for wild-type and *chl4Δ* cells in the absence or presence of spindle tension. **b**, Pile-ups of *cis* contacts in the 100 kb around centromeres (left); ratio of expected/observed signals (second panel); pericentromere pile-up (third panel, 10 kb surrounding centromeres); and its ratio of expected/observed signals (right), for wild-type

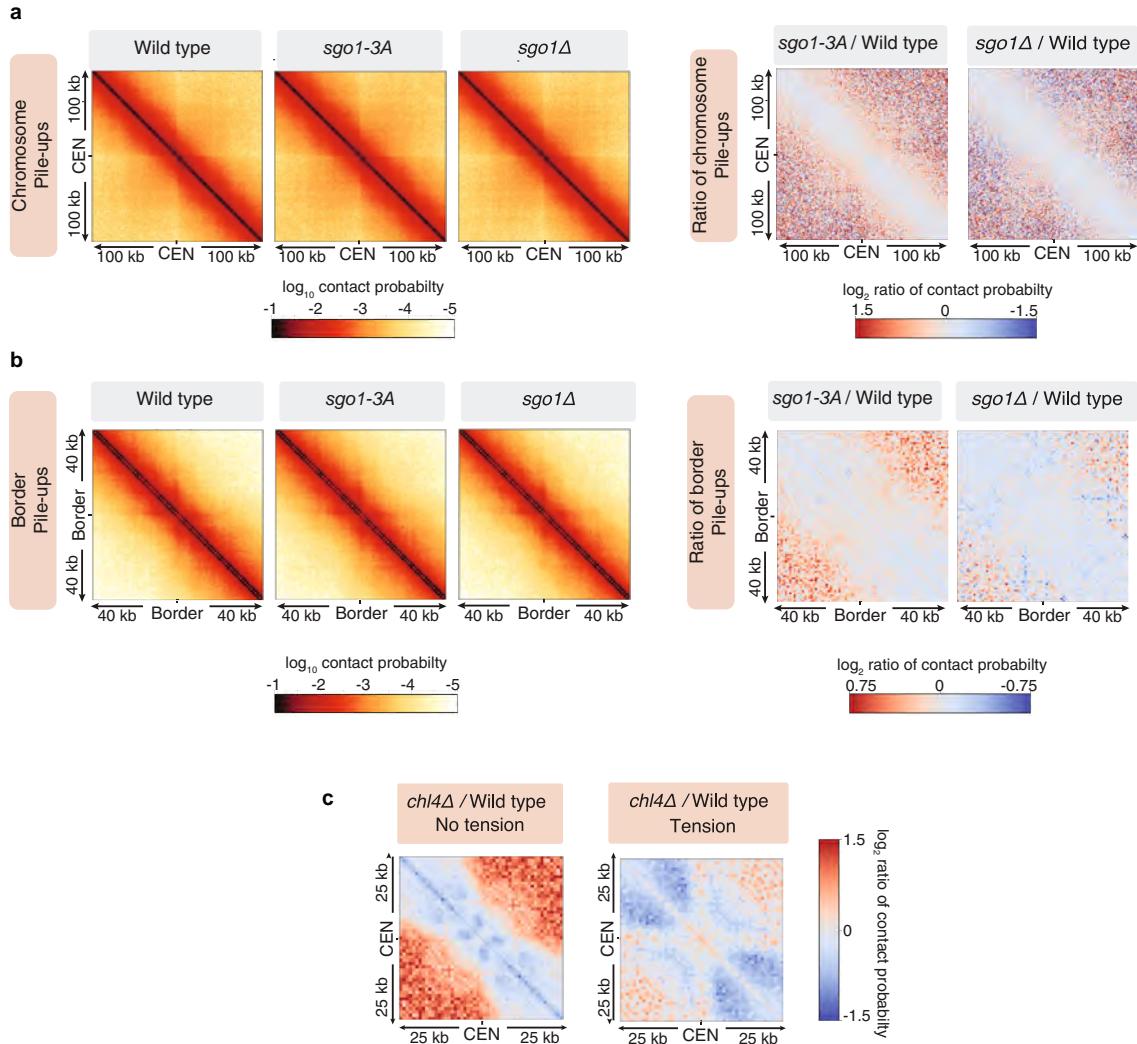
cells in the absence of spindle tension ( $n=16$ ). **c**, Model proposed previously<sup>19</sup> showing the formation of intramolecular loops by centromere-flanking pericentromeric chromatin, with cohesin bridging the two sides of the pericentromere flanking the centromere; meanwhile chromosome arms are paired intermolecularly between sister chromatids, resulting in a cruciform chromosome conformation.

# Article



**Extended Data Fig. 7 | Changes in pericentromere structure on individual chromosomes in response to tension and in the absence of pericentromeric cohesin.** Shown are Hi-C contact maps (1-kb bin) over a 50-kb region surrounding all centromeres in wild-type cells (WT; left panels for each chromosome) without tension (bottom half of heatmap) and with tension (top

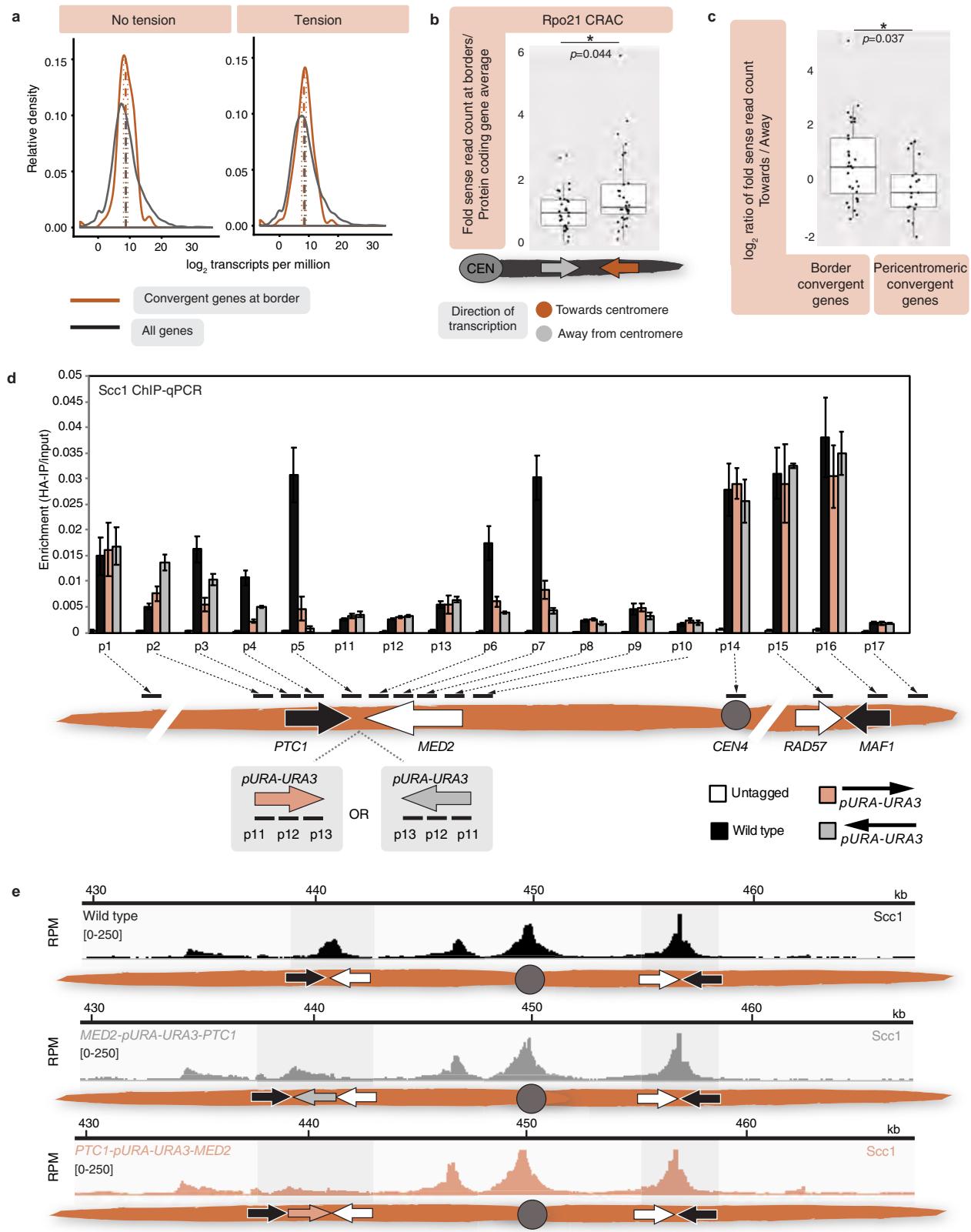
half of heatmap), and in *chl4Δ* cells (middle panels) without tension (bottom half) and with tension (top half) ( $n=1$ ). Right,  $\log_2$  ratios between wild-type and *chl4Δ* cells, without tension (bottom half) and with tension (top half). The extent of the pericentromere for each chromosome is marked by the black bars on top of the contact maps.



**Extended Data Fig. 8 | Pericentromere structure depends on pericentromeric cohesin rather than condensin.** Hi-C analysis of *sgo1-3A* and *sgo1Δ* metaphase-arrested cells in the absence of tension reveals similar patterns to wild type. Data for wild type are reproduced from Extended Data Fig. 6 for comparison. **a**, Pile-ups (bin size 1 kb) of *cis* contacts surrounding all 16 centromeres in the absence of spindle tension for the indicated strains (left

three panels), and  $\log_2$  ratio maps between wild type and *sgo1-3A* or *sgo1Δ* (right two panels), detect little change. **b**, Pile-ups (bin size 1 kb) and  $\log_2$  ratio maps of *cis* contacts surrounding all 32 borders. **c**, The  $\log_2$  ratio between 25-kb pile-ups centred on the centromere in wild-type and *chl4Δ* cells, in the absence (left) and presence (right) of tension ( $n=16$ ).

## Article

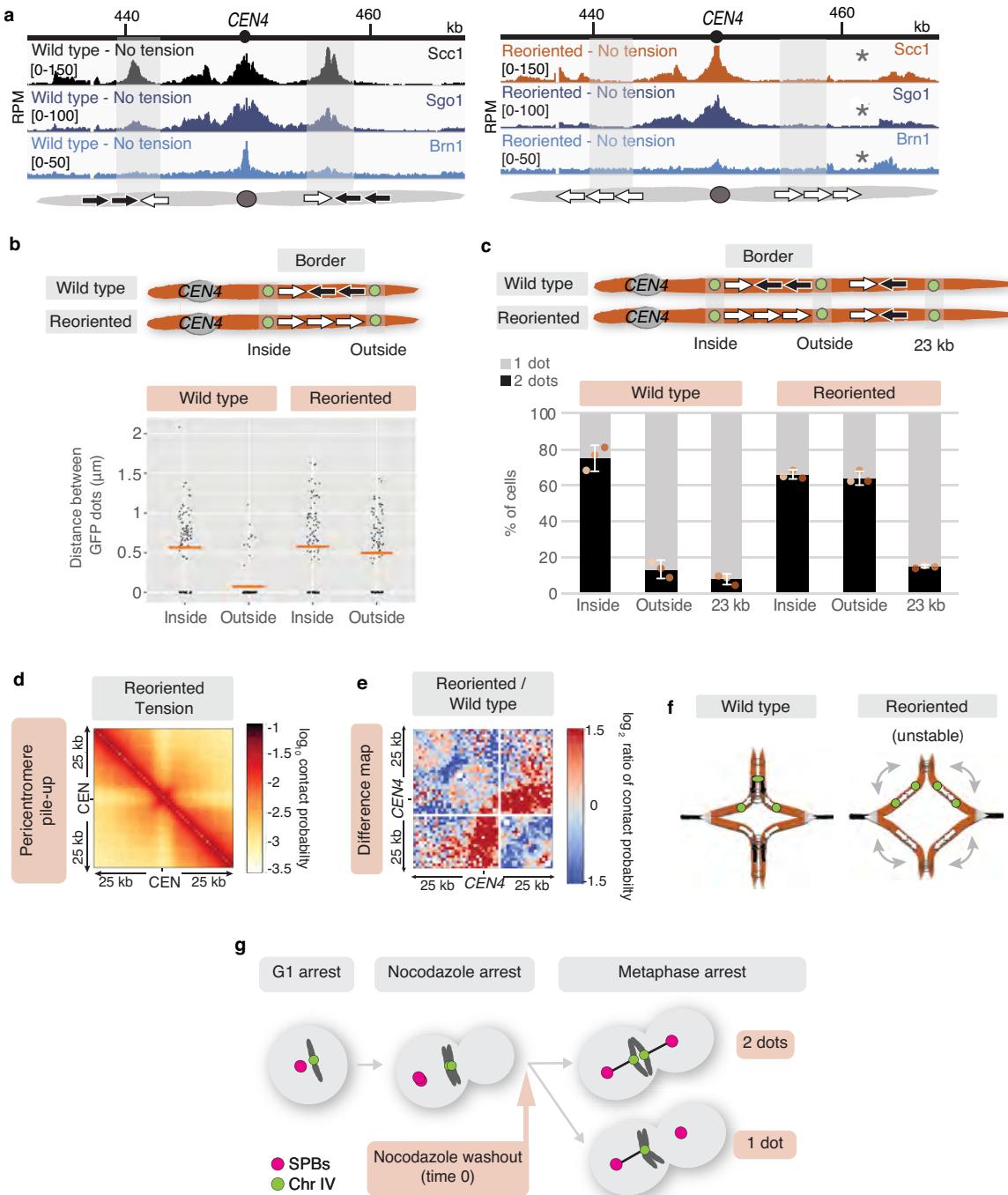


**Extended Data Fig. 9** | See next page for caption.

**Extended Data Fig. 9 | Transcription at pericentromere borders influences cohesin position.** **a**, Genes at pericentromere borders are moderately transcribed on average. Shown is the relative RNA density for convergent border gene pairs compared with all genes under conditions of no tension or tension ( $n=1$ ). RNA-seq analysis of wild-type cells arrested in metaphase in the presence or absence of tension. Dashed lines indicate means; dotted lines mark 95% confidence intervals. **b**, Box plot showing transcription levels of genes at pericentromere borders based on RNA polymerase II (Rpo21) cross-linking and analysis of cDNA (CRAC) from ref.<sup>29</sup>. Rpo21 CRAC sense read counts of genes at borders were normalized to the protein-coding-gene average, and genes at pericentromere borders were grouped by their relative orientation to centromeres. Data points correspond to means of three biological repeats.

Centre lines, medians; box limits, second and third quartiles; whiskers, first and fourth quartiles (non-normal distribution, Shapiro-Wilk; \* $P<0.05$ , two-sided Mann-Whitney test). **c**, Box plot showing relative transcription levels of genes transcribed towards and away from centromeres, at pericentromere borders and at non-border convergent genes inside pericentromeres, as in **b**. **d, e**, Insertion of a *URA3* cassette between a convergent gene pair shifts the localization of cohesin in the direction of transcription. *URA3* was integrated in either orientation between the convergent gene pairs at the left pericentromere border on chromosome IV, and cohesin (Scc1) ChIP-qPCR ( $n=3$ ; bars show means  $\pm$  s.e.m.) using primers at the indicated positions (**d**) and ChIP-seq ( $n=1$ ) (**e**) were performed.

# Article



**Extended Data Fig. 10 | Gene reorientation at borders affects pericentromeric protein localization, sister-chromatid separation and chromosome conformation.** **a**, Cohesin (*Scc1*), shugoshin (*Sgo1*) and condensin (*Brn1*) enrichment in metaphase-arrested cells in the presence of nocodazole in the pericentromeric region of wild-type and reoriented chromosome IV is shown ( $n=2$ ; immunoprecipitation was performed using proteins tagged with different epitopes in the two replicates). Asterisks indicate new peaks in reoriented strain. **b**, The distance between GFP foci does not change following gene reorientation. We measured 100 cells; horizontal lines indicate means. **c**, The region of separation upon gene reorientation does not extend beyond the next convergent gene pair. Strains with *tetO* arrays integrated at the indicated positions were arrested in metaphase, and the

percentage of cells with two GFP foci was determined. For each biological replicate (data points;  $n=3$ ), 200 cells were scored; data are means  $\pm$  s.e.m. **d**, Pile-up of *cis* contacts across all 16 pericentromeres in the reoriented strain, in the presence of spindle tension. **e**, log<sub>2</sub> ratio map of Hi-C signal in pericentromere IV in wild-type and reoriented strains ( $n=1$ ). **f**, Model for pericentromere expansion and disorganization in the absence of convergent genes. **g**, Diagram showing biorientation of sister kinetochores following spindle repolymerization. Cells carrying chromosomal GFP labels, Spc42-tdTTomato and pMET-CDC20 were released from G1 phase into a metaphase arrest by depleting Cdc20 in the presence of nocodazole. Nocodazole was washed out while maintaining metaphase arrest and the percentage of cells with two GFP foci was scored over time.

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

## Software and code

Policy information about [availability of computer code](#)

## Data collection

MicroManager 1.4 acquisition software  
Zen Pro 2.3 acquisition software

## Data analysis

HiC-Pro v2.11.1  
bowtie2 v2.3.4.1  
IGV 2.4.16  
Image J  
ImageJ plugin to measure the fluorescent intensity and distance between the GFP foci [https://github.com/dkelly604/CellClicker\\_Scripts](https://github.com/dkelly604/CellClicker_Scripts) for Hi-C data analysis [https://github.com/danrobertson87/Paldi\\_2019](https://github.com/danrobertson87/Paldi_2019)  
Scripts for SeqPlots <https://github.com/AlastairKerr>  
samtools v1.1  
MultiQC v1.7

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

*ImageJ plugin to measure the fluorescent intensity and distance between the GFP foci can be found on the github repository [https://github.com/dkelly604/CellClicker\\_Scripts](https://github.com/dkelly604/CellClicker_Scripts). Scripts for Hi-C data analysis are available at [https://github.com/danrobertson87/Paldi\\_2019](https://github.com/danrobertson87/Paldi_2019), scripts to produce SeqPlots are available at <https://github.com/AlastairKerr>*

[github.com/AlastairKerr](https://github.com/AlastairKerr). Raw and processed sequencing data can be found on GEO under the accession GSE104135. Figures that have associated raw data: Fig 1a-c, e; Fig 2a-e; 3a, d; Extended Data Fig 1a, b; Extended Data Fig 3, Extended Data Fig 4a-f; Extended Data Fig 5b; Extended Data Fig 6a, b; Extended Data Fig 7; Extended Data Fig 8a-c; Extended Data Fig 9a, e; Extended Data Fig 10a, d, e.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For the Hi-C and ChIP-Seq experiments, cultures containing several million cells were used to determine the genome average. For microscopy experiments with fixed cells, at least 200 cells were scored for each timepoint in each biological repeat. For live cell imaging at least 60 cells were analyzed for each biological repeat. For viability assays, 1000 colonies were plated. Sample sizes are provided in the figure legend. These sample sizes are typical of similar studies and were selected to ensure robust and statistically significant comparisons.
Data exclusions	No data was excluded
Replication	At least three replicates of each time course microscopy experiment and viability assay were performed and the average of three repeats is shown. Hi-C and most ChIP-Seq experiments are a single replicate for each condition, however, chromosomal arms show the same pattern in all conditions, indicating the reproducibility and all 16 pericentromeres in each case show the same pattern. All replicates are included where adequate arrest conditions were met during quality control.
Randomization	Not applicable for genetic comparisons because yeast strains used were isogenic clones with test strains having only the indicated mutations compared to wild type. All cells analyzed in population and single cell analyses were derived from the same verified clone. In experiments where drug treated (nocodazole) and untreated (DMSO) cells were compared, cultures were grown from the same yeast isolate cultured in the same flask and split, or in parallel cultures which were treated identically, prior to nocodazole or DMSO addition.
Blinding	Microscopy for the nocodazole wash-out, partial ip1-321 inactivation, live-cell imaging and rescue experiments was scored blind after a colleague set up randomised cultures. In other experiments flasks were labelled with strain numbers but no other identifying information. GFP foci were scored manually in Extended Data Fig. 5 and 10 and after using the custom-written macro with similar results.

## Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).
Research sample	State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.
Sampling strategy	Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.
Data collection	Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.
Timing	Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.
Data exclusions	If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.
Non-participation	State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.
Randomization	If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.

# Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

## Study description

*Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.*

## Research sample

*Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.*

## Sampling strategy

*Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*

## Data collection

*Describe the data collection procedure, including who recorded the data and how.*

## Timing and spatial scale

*Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken*

## Data exclusions

*If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*

## Reproducibility

*Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.*

## Randomization

*Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.*

## Blinding

*Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Did the study involve field work?  Yes  No

## Field work, collection and transport

### Field conditions

*Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).*

### Location

*State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).*

### Access and import/export

*Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).*

### Disturbance

*Describe any disturbance caused by the study and how it was minimized.*

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

- |                                     |   |
|-------------------------------------|---|
| n/a                                 | Involved in the study                                     |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Antibodies            |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology                    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms      |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants      |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                    |

### Methods

- |                                     |   |
|-------------------------------------|---|
| n/a                                 | Involved in the study                           |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> ChIP-seq    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

## Antibodies

Antibodies used	Monoclonal ANTI-FLAG® M2 antibody produced in mouse (F1804, Sigma Aldrich) LOT 087K5001 - 10 µg per ChIP sample Monoclonal anti-HA 12CA5 antibody produced in mouse (11583816001 Roche) LOT 14307900 - 6 µg per ChIP sample
Validation	Both antibodies have been extensively used for ChIP in our laboratory and antibodies have been validated using no tag controls, see in Hinshaw et al. 2017 (doi: 10.1016/j.cell.2017.08.017).

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	All <i>Saccharomyces cerevisiae</i> strains are derivatives of w303. Detailed genotypes are listed in Extended Data Table 1. AMy1105 - Fernius et al. 2009 AMy1145 - Fernius et al. 2009 AMy2508 - Fernius et al. 2009 AMy3950 - Fernius et al. 2009 AMy6389 - Nerusheva et al. 2014 AMy6471 - Lab strain collection AMy6884 - Lab strain collection AMy7217 - Lab strain collection AMy8955 - Nerusheva et al. 2014 AMy14126 - This study AMy16144 - This study AMy16541 - This study AMy16721 - This study AMy22078 - This study AMy22900 - This study AMy22936 - This study AMy23081 - This study AMy23082 - This study AMy23125 - This study AMy23185 - This study AMy25236 - This study AMy25297 - This study AMy25298 - This study AMy25299 - This study AMy25379 - This study AMy25409 - This study AMy25764 - This study AMy26822 - This study AMy26964 - This study AMy26965 - This study AMy26966 - This study AMy27213 - This study AMy27214 - This study AMy27215 - This study AMy27216 - This study AMy27936 - This study AMy28477 - This study AMy28478 - This study AMy28726 - This study AMy28787 - This study AMy28788 - This study AMy28790 - This study AMy28791 - This study AMy28792 - This study AMy28793 - This study AMy28794 - This study AMy28878 - This study
Authentication	All genetic modifications were originally created by transformation and combined using genetic crosses. Strains were verified by the presence or absence of marker genes allowing growth on selective medium and/or by PCR to determine the presence of the desired genetic alteration and/or by DNA sequencing and/or western blotting to confirm the presence of epitope tags, as appropriate.
Mycoplasma contamination	Not applicable - Yeast strains were not tested for mycoplasma contamination, but all cultures were observed by microscopy to be free of bacterial contamination.
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	No commonly misidentified lines were used. This study used the yeast <i>Saccharomyces cerevisiae</i> .

## Palaeontology

Specimen provenance	<i>Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).</i>
Specimen deposition	<i>Indicate where the specimens have been deposited to permit free access by other researchers.</i>
Dating methods	<i>If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.</i>

Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

## Animals and other organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	<i>For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.</i>
Wild animals	<i>Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.</i>
Field-collected samples	<i>For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.</i>
Ethics oversight	<i>Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.</i>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	<i>Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural &amp; social sciences study design questions and have nothing to add here, write "See above."</i>
Recruitment	<i>Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.</i>
Ethics oversight	<i>Identify the organization(s) that approved the study protocol.</i>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	<i>Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.</i>
Study protocol	<i>Note where the full trial protocol can be accessed OR if not available, explain why.</i>
Data collection	<i>Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.</i>
Outcomes	<i>Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.</i>

## ChIP-seq

### Data deposition

- Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE104135">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE104135</a>
	<i>May remain private before publication.</i>

## Files in database submission

GSM2787669\_6390\_Sgo1-6HA\_WT+tension\_Input.bw  
 GSM2787670\_6390\_Sgo1-6HA\_WT+tension\_ChIP.bw  
 GSM2787671\_6390\_Sgo1-6HA\_WT-tension\_Input.bw  
 GSM2787672\_6390\_Sgo1-6HA\_WT-tension\_ChIP.bw  
 GSM2789271\_WTInput+ten.bw  
 GSM2789271\_WTInput+ten\_ref\_CEN3del.bw  
 GSM2789272\_WTChip+ten.bw  
 GSM2789272\_WTChip+ten\_ref\_CEN3del.bw  
 GSM2789273\_WTInput-ten.bw  
 GSM2789273\_WTInput-ten\_ref\_CEN3del.bw  
 GSM2789274\_WTChip-ten.bw  
 GSM2789274\_WTChip-ten\_ref\_CEN3del.bw  
 GSM2789279\_AM16144\_Scc1-6HA\_CEN3del\_tension\_input.bw  
 GSM2789279\_AM16144\_Scc1-6HA\_CEN3del\_tension\_input\_ref\_CEN3del.bw  
 GSM2789281\_AM16144\_Scc1-6HA\_CEN3del\_no\_tension\_input.bw  
 GSM2789281\_AM16144\_Scc1-6HA\_CEN3del\_no\_tension\_input\_ref\_CEN3del.bw  
 GSM2789283\_AM16144\_Scc1-6HA\_CEN3del\_tension\_IP.bw  
 GSM2789283\_AM16144\_Scc1-6HA\_CEN3del\_tension\_IP\_ref\_CEN3del.bw  
 GSM2789285\_AM16144\_Scc1-6HA\_CEN3del\_no\_tension\_IP.bw  
 GSM2789285\_AM16144\_Scc1-6HA\_CEN3del\_no\_tension\_IP\_ref\_CEN3del.bw  
 GSM3667525\_26966-T-input\_S1.sacCer3.only.rpm.bw  
 GSM3667526\_26966-NT-input\_S2.sacCer3.only.rpm.bw  
 GSM3667527\_26966-T-IP\_S3.sacCer3.only.rpm.bw  
 GSM3667527\_26966-NT-IP\_S4.sacCer3.only.rpm.calibrated.bw  
 GSM3667528\_26966-NT-IP\_S4.sacCer3.only.rpm.bw  
 GSM3667528\_26966-NT-IP\_S4.sacCer3.only.rpm.calibrated.bw  
 GSM3667553\_22078\_input\_S1\_plus\_tension.sacCer3.only.rpm.bw  
 GSM3667554\_22078\_input\_S3\_no\_tension.sacCer3.only.rpm.bw  
 GSM3667555\_22078\_IP\_S2\_plus\_tension.sacCer3.only.calibrated.bw  
 GSM3667555\_22078\_IP\_S2\_plus\_tension.sacCer3.only.rpm.bw  
 GSM3667556\_22078\_IP\_S4\_no\_tension.sacCer3.only.calibrated.bw  
 GSM3667556\_22078\_IP\_S4\_no\_tension.sacCer3.only.rpm.bw  
 GSM3667557\_25298-input\_S1\_sgo1.sacCer3.only.rpm.bw  
 GSM3667558\_25298-IP\_S2\_sgo1.sacCer3.only.calibrated.bw  
 GSM3667558\_25298-IP\_S2\_sgo1.sacCer3.only.rpm.bw  
 GSM3667559\_25299-input\_S1\_brn1.sacCer3.only.rpm.bw  
 GSM3667560\_25299-IP\_S2\_brn1.sacCer3.only.calibrated.bw  
 GSM3668244\_25409-input\_S3\_sgo1.sacCer3.only.rpm.bw  
 GSM3668245\_25409-IP\_S4\_sgo1.sacCer3.only.calibrated.bw  
 GSM3668245\_25409-IP\_S4\_sgo1.sacCer3.only.rpm.bw  
 GSM3668246\_25379-input\_S3\_brn1.sacCer3.only.rpm.bw  
 GSM3668247\_25379-IP\_S4\_brn1.sacCer3.only.calibrated.bw  
 GSM3668247\_25379-IP\_S4\_brn1.sacCer3.only.rpm.bw  
 GSM3668252\_1105-T-input\_S1.sacCer3.rpm.bw  
 GSM3668253\_1105-NT-input\_S2.sacCer3.rpm.bw  
 GSM3668254\_1105-T-IP\_S3.sacCer3.calibrated.bw  
 GSM3668254\_1105-T-IP\_S3.sacCer3.rpm.bw  
 GSM3668255\_1105-NT-IP\_S4.sacCer3.calibrated.bw  
 GSM3668255\_1105-NT-IP\_S4.sacCer3.rpm.bw  
 GSM3668256\_3950-T-input\_S1.sacCer3.rpm.bw  
 GSM3668257\_3950-NT-input\_S2.sacCer3.rpm.bw  
 GSM3668258\_3950-T-IP\_S3.sacCer3.calibrated.bw  
 GSM3668258\_3950-T-IP\_S3.sacCer3.rpm.bw  
 GSM3668259\_3950-NT-IP\_S4.sacCer3.calibrated.bw  
 GSM3668259\_3950-NT-IP\_S4.sacCer3.rpm.bw  
 GSM4139633 Scc1-6HA PTC1-pURA3-URA3-MED2 Input  
 GSM4139634 Scc1-6HA PTC1-pURA3-URA3-MED2 IP  
 GSM4139635 Scc1-6HA MED2-pURA3-URA3-PTC1 Input  
 GSM4139636 Scc1-6HA MED2-pURA3-URA3-PTC1 IP  
 GSM4139637 Sgo1-6HA No tension Input  
 GSM4139638 Sgo1-6HA No tension IP  
 GSM4139639 Brn1-6HA No tension Input  
 GSM4139640 Brn1-6HA No tension IP  
 GSM4139641 1145 Scc1-6HA No tension Input  
 GSM4139642 1145 Scc1-6HA No tension IP

Genome browser session  
(e.g. [UCSC](#))

<https://software.broadinstitute.org/software/igv/>

## Methodology

## Replicates

BA5: 6390\_Sgo1-6HA\_WT+tension\_Input - 1 replicate  
 BA6: 6390\_Sgo1-6HA\_WT+tension\_ChIP - 1 replicate

BA7: 6390\_Sgo1-6HA\_WT-tension\_Input - 1 replicate  
 BA8: 6390\_Sgo1-6HA\_WT-tension\_ChIP - 1 replicate  
 WTInput+ten - 1 replicate  
 WTChip+ten - 1 replicate  
 WTInput-ten - 1 replicate  
 WTChip-ten - 1 replicate  
 AM16144\_Scc1-6HA\_CEN3del\_tension\_input - 1 replicate  
 AM16144\_Scc1-6HA\_CEN3del\_no\_tension\_input - 1 replicate  
 AM16144\_Scc1-6HA\_CEN3del\_tension\_IP - 1 replicate  
 AM16144\_Scc1-6HA\_CEN3del\_no\_tension\_IP - 1 replicate  
 Scc1-6HA rad61Δ Tension input - 1 replicate  
 Scc1-6HA rad61Δ No tension input - 1 replicate  
 Scc1-6HA rad61Δ Tension IP - 1 replicate  
 Scc1-6HA rad61Δ No tension IP - 1 replicate  
 Scc1-6HA Reoriented Tension Input - 1 replicate  
 Scc1-6HA Reoriented No tension Input - 1 replicate  
 Scc1-6HA Reoriented Tension IP - 1 replicate  
 Scc1-6HA Reoriented No tension IP - 1 replicate  
 Sgo1-6HIS-3FLAG Reoriented No tension Input - 1 replicate  
 Sgo1-6HIS-3FLAG Reoriented No tension IP - 1 replicate  
 Brn1-6HIS-3FLAG Reoriented No tension Input - 1 replicate  
 Brn1-6HIS-3FLAG Reoriented No tension IP - 1 replicate  
 Sgo1-6HIS-3FLAG No tension Input - 1 replicate  
 Sgo1-6HIS-3FLAG No tension IP - 1 replicate  
 Brn1-6HIS-3FLAG No tension Input - 1 replicate  
 Brn1-6HIS-3FLAG No tension IP - 1 replicate  
 Scc1-6HA WT Tension input - 2 biological replicates  
 Scc1-6HA WT No tension input - 2 biological replicates  
 Scc1-6HA WT Tension IP - 2 biological replicates  
 Scc1-6HA WT No tension IP- 2 biological replicates  
 Scc1-6HA chl4Δ Tension input - 2 biological replicates  
 Scc1-6HA chl4Δ No tension input - 2 biological replicates  
 Scc1-6HA chl4Δ Tension IP - 2 biological replicates  
 Scc1-6HA chl4Δ No tension IP - 2 biological replicates  
 Sgo1-6HA No tension Input - 1 replicate  
 Sgo1-6HA No tension IP - 1 replicate  
 Brn1-6HA No tension Input - 1 replicate  
 Brn1-6HA No tension IP - 1 replicate  
 1145 Scc1-6HA WT No tension input - 1 replicate  
 1145 Scc1-6HA WT No tension IP - 1 replicate  
 Scc1-6HA PTC1-pURA3-URA3-MED2 No tension input - 1 replicate  
 Scc1-6HA PTC1-pURA3-URA3-MED2 No tension IP - 1 replicate  
 Scc1-6HA MED2-pURA3-URA3-PTC1 No tension input - 1 replicate  
 Scc1-6HA MED2-pURA3-URA3-PTC1 No tension IP - 1 replicate

## Sequencing depth

Scc1-6HA rad61Δ Tension input Total - 6.2 M, Unique - 5.0 M, 76 bp - Paired-end  
 Scc1-6HA rad61Δ No tension input Total - 5 M, Unique - 4.2 M, 76 bp - Paired-end  
 Scc1-6HA rad61Δ Tension IP Total - 30 M, Unique - 11.4 M, 76 bp - Paired-end  
 Scc1-6HA rad61Δ No tension IP Total - 27.4 M, Unique - 11.9 M, 76 bp - Paired-end  
 Scc1-6HA Reoriented Tension Input Total - 5.8 M, Unique - 5.0 M, 76 bp - Paired-end  
 Scc1-6HA Reoriented No tension Input Total - 5.4 M, Unique - 4.7 M, 76 bp - Paired-end  
 Scc1-6HA Reoriented Tension IP Total - 29.8 M, Unique - 18.2 M, 76 bp - Paired-end  
 Scc1-6HA Reoriented No tension IP Total - 28.8 M, Unique - 13.9 M, 76 bp - Paired-end  
 Sgo1-6HIS-3FLAG Reoriented No tension Input Total - 4 M, Unique - 3.3 M, 76 bp - Paired-end  
 Sgo1-6HIS-3FLAG Reoriented No tension IP Total - 29.6 M, Unique - 15.0 M, 76 bp - Paired-end  
 Brn1-6HIS-3FLAG Reoriented No tension Input Total - 0.4 M, Unique - 0.4 M, 76 bp - Paired-end  
 Brn1-6HIS-3FLAG Reoriented No tension IP Total - 4 M, Unique - 3.3 M, 76 bp - Paired-end  
 Sgo1-6HIS-3FLAG No tension Input Total - 5.2 M, Unique - 4.2 M, 76 bp - Paired-end  
 Sgo1-6HIS-3FLAG No tension IP Total - 28.8 M, Unique - 13.9 M, 76 bp - Paired-end  
 Brn1-6HIS-3FLAG No tension Input Total - 4.8 M, Unique - 3.9 M, 76 bp - Paired-end  
 Brn1-6HIS-3FLAG No tension IP Total - 30 M, Unique - 15.1 M, 76 bp - Paired-end  
 Scc1-6HA WT Tension input Total - 5.2 M, Unique - 4.2 M, 76 bp - Paired-end  
 Scc1-6HA WT No tension input Total - 5.4 M, Unique - 4.3 M, 76 bp - Paired-end  
 Scc1-6HA WT Tension IP Total - 27.8 M, Unique - 17.1 M, 76 bp - Paired-end  
 Scc1-6HA WT No tension IP Total - 27 M, Unique - 14.7 M, 76 bp - Paired-end  
 Scc1-6HA chl4Δ Tension input Total - 5.2 M, Unique - 4.2 M, 76 bp - Paired-end  
 Scc1-6HA chl4Δ No tension input Total - 4.2 M, Unique - 3.4 M, 76 bp - Paired-end  
 Scc1-6HA chl4Δ Tension IP Total - 24 M, Unique - 12.3 M, 76 bp - Paired-end  
 Scc1-6HA chl4Δ No tension IP Total - 24 M, Unique - 11.5 M, 76 bp - Paired-end  
 BA5: 6390\_Sgo1-6HA\_WT+tension\_Input Total - 17.4 M, Unique - 10.8 M, 51 bp - Single-end  
 BA6: 6390\_Sgo1-6HA\_WT+tension\_ChIP Total - 14.0 M, Unique - 7.3 M, 51 bp - Single-end  
 BA7: 6390\_Sgo1-6HA\_WT-tension\_Input Total - 16.8 M, Unique - 9.6 M, 51 bp - Single-end  
 BA8: 6390\_Sgo1-6HA\_WT-tension\_ChIP Total - 12.2 M, Unique - 6.8 M, 51 bp - Single-end  
 WTInput+ten Total - 16.2 M, Unique - 9.7 M, 50 bp - Single-end  
 WTChip+ten Total - 26.7 M, Unique - 11.1 M, 50 bp - Single-end

WTInput-ten Total - 21.3 M, Unique - 11.6 M, 50 bp - Single-end  
 WTChip-ten Total - 24.3 M, Unique - 10.8 M, 50 bp - Single-end  
 AM16144\_Scc1-6HA\_CEN3del\_tension\_input Total - 17.1 M, Unique - 8.4 M, 51 bp - Single-end  
 AM16144\_Scc1-6HA\_CEN3del\_no\_tension\_input Total - 15.3 M, Unique - 8.0 M, 51 bp - Single-end  
 AM16144\_Scc1-6HA\_CEN3del\_tension\_IP Total - 15.4 M, Unique - 9.4 M, 51 bp - Single-end  
 AM16144\_Scc1-6HA\_CEN3del\_no\_tension\_IP Total - 14.6 M, Unique - 8.9 M, 51 bp - Single-end  
 Sgo1-6HA No tension Input Total - 3.6 M, Unique - 3.1 M, 76 bp - Paired-end  
 Sgo1-6HA No tension IP Total - 5.2 M, Unique - 3.6 M, 76 bp - Paired-end  
 Brn1-6HA No tension Input Total - 3.2 M, Unique - 2.8 M, 76 bp - Paired-end  
 Brn1-6HA No tension IP Total - 7 M, Unique - 3.9 M, 76 bp - Paired-end  
 1145\_Scc1-6HA\_WT\_No\_tension\_input Total - 4.2 M, Unique - 3.5 M, 76 bp - Paired-end  
 1145\_Scc1-6HA\_WT\_No\_tension\_IP Total - 11.8 M, Unique - 6.0 M, 76 bp - Paired-end  
 Scc1-6HA\_PTC1-pURA3-URA3-MED2\_No\_tension\_input Total - 3.6 M, Unique - 3.0 M, 76 bp - Paired-end  
 Scc1-6HA\_PTC1-pURA3-URA3-MED2\_No\_tension\_IP Total - 6.4 M, Unique - 4.0 M, 76 bp - Paired-end  
 Scc1-6HA\_MED2-pURA3-URA3-PTC1\_No\_tension\_input Total - 4.2 M, Unique - 3.6 M, 76 bp - Paired-end  
 Scc1-6HA\_MED2-pURA3-URA3-PTC1\_No\_tension\_IP Total - 6.6 M, Unique - 4.7 M, 76 bp - Paired-end

**Antibodies**

Monoclonal ANTI-FLAG® M2 antibody produced in mouse (F1804, Sigma Aldrich)  
 Monoclonal anti-HA 12CA5 antibody produced in mouse (11583816001 Roche)

**Peak calling parameters**

N/A

**Data quality**

N/A

**Software**

Basecalls performed using CASAVA  
 Reads were trimmed with cutadapt (MAPQ 10, min. length 65)  
 To obtain reads mapping only to sacCer3;  
 ChIP-Seq reads were first mapped with MiniMap2 (-ax sr) to pombe  
 Unmapped pombe reads were filtered using samtools  
 Unmapped pombe reads were then mapped to sacCer3  
 Unmapped reads were filtered using samtools  
 rDNA regions were removed from chrXII (saturated with reads)  
 chrMito was excluded using bedtools  
 To obtain reads mapping only to Pombe;  
 Reads were also mapped to sacCer3  
 Unmapped sacCer3 reads were filtered using samtools  
 Unmapped sacCer3 reads were then mapped to pombe  
 Unmapped reads were filtered using samtools  
 chrMito was excluded using bedtools  
 Genome\_build: sacCer3 + s.pombe ASM294v2.22  
 Supplementary\_files\_format\_and\_content: bigwigs were created using bedtools genomeCoverageBed & wigToBigWig (RPM normalization), To obtain calibrated bigWigs; samtools flagstat was used to count reads mapping to sacCer3 and pombe only for each sample the occupancy ratio (OR) value was calculated as  $W_c * I_p / W_x * I_p$  ( $W$ =Input;  $I$ =ChIP;  $c$ =calibration genome (pombe);  $x$ =experimental genome (sacCer3)), calibrated sacCer3 chIP bigwigs were created using bedtools genomeCoverageBed & wigToBigWig

## Flow Cytometry

**Plots**

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

**Methodology****Sample preparation**

Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.

**Instrument**

Identify the instrument used for data collection, specifying make and model number.

**Software**

Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.

**Cell population abundance**

Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.

**Gating strategy**

*Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

## Magnetic resonance imaging

### Experimental design

**Design type**

*Indicate task or resting state; event-related or block design.*

**Design specifications**

*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*

**Behavioral performance measures**

*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

### Acquisition

**Imaging type(s)**

*Specify: functional, structural, diffusion, perfusion.*

**Field strength**

*Specify in Tesla*

**Sequence & imaging parameters**

*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*

**Area of acquisition**

*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

**Diffusion MRI**

Used

Not used

### Preprocessing

**Preprocessing software**

*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*

**Normalization**

*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*

**Normalization template**

*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI1305, ICBM152) OR indicate that the data were not normalized.*

**Noise and artifact removal**

*Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*

**Volume censoring**

*Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

### Statistical modeling & inference

**Model type and settings**

*Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*

**Effect(s) tested**

*Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

**Specify type of analysis:**  Whole brain  ROI-based  Both**Statistic type for inference  
(See Eklund et al. 2016)**

*Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*

**Correction**

*Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

### Models & analysis

**n/a Involved in the study**

- Functional and/or effective connectivity
- Graph analysis
- Multivariate modeling or predictive analysis

**Functional and/or effective connectivity**

*Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*

**Graph analysis**

*Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*

**Multivariate modeling and predictive analysis**

*Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*

## Article

# Multilayered mechanisms ensure that short chromosomes recombine in meiosis

<https://doi.org/10.1038/s41586-020-2248-2>

Received: 7 January 2019

Accepted: 12 March 2020

Published online: 06 May 2020

 Check for updates

Hajime Murakami<sup>1</sup>✉, Isabel Lam<sup>1,2,5</sup>, Pei-Ching Huang<sup>1,3</sup>, Jacquelyn Song<sup>1</sup>, Megan van Overbeek<sup>1,6</sup> & Scott Keeney<sup>1,2,3,4</sup>✉

In most species, homologous chromosomes must recombine in order to segregate accurately during meiosis<sup>1</sup>. Because small chromosomes would be at risk of missegregation if recombination were randomly distributed, the double-strand breaks (DSBs) that initiate recombination are not located arbitrarily<sup>2</sup>. How the nonrandomness of DSB distributions is controlled is not understood, although several pathways are known to regulate the timing, location and number of DSBs. Meiotic DSBs are generated by Spo11 and accessory DSB proteins, including Rec114 and Mer2, which assemble on chromosomes<sup>3–7</sup> and are nearly universal in eukaryotes<sup>8–11</sup>. Here we demonstrate how *Saccharomyces cerevisiae* integrates multiple temporally distinct pathways to regulate the binding of Rec114 and Mer2 to chromosomes, thereby controlling the duration of a DSB-competent state. The engagement of homologous chromosomes with each other regulates the dissociation of Rec114 and Mer2 later in prophase I, whereas the timing of replication and the proximity to centromeres or telomeres influence the accumulation of Rec114 and Mer2 early in prophase I. Another early mechanism enhances the binding of Rec114 and Mer2 specifically on the shortest chromosomes, and is subject to selection pressure to maintain the hyperrecombinogenic properties of these chromosomes. Thus, the karyotype of an organism and its risk of meiotic missegregation influence the shape and evolution of its recombination landscape. Our results provide a cohesive view of a multifaceted and evolutionarily constrained system that allocates DSBs to all pairs of homologous chromosomes.

Homologous chromosomes can recombine only if they have at least one DSB. Simulations in *S. cerevisiae* show that, if DSBs were distributed randomly, the shortest chromosomes (chromosome (chr) 1, 3 and 6) would be at risk of having none at all (Fig. 1a). However, from observations that chromosome size is negatively correlated with crossover density<sup>12–14</sup>, DSB density<sup>15,16</sup> and DSB-protein binding<sup>3,17</sup>, it is evident that mechanisms exist to attenuate this risk (Fig. 1b, Extended Data Fig. 1a). Short chromosomes therefore recruit more DSB proteins, presumably yielding a higher density of DSBs and crossovers; however, the mechanisms by which DSB proteins are preferentially recruited to short chromosomes are not known.

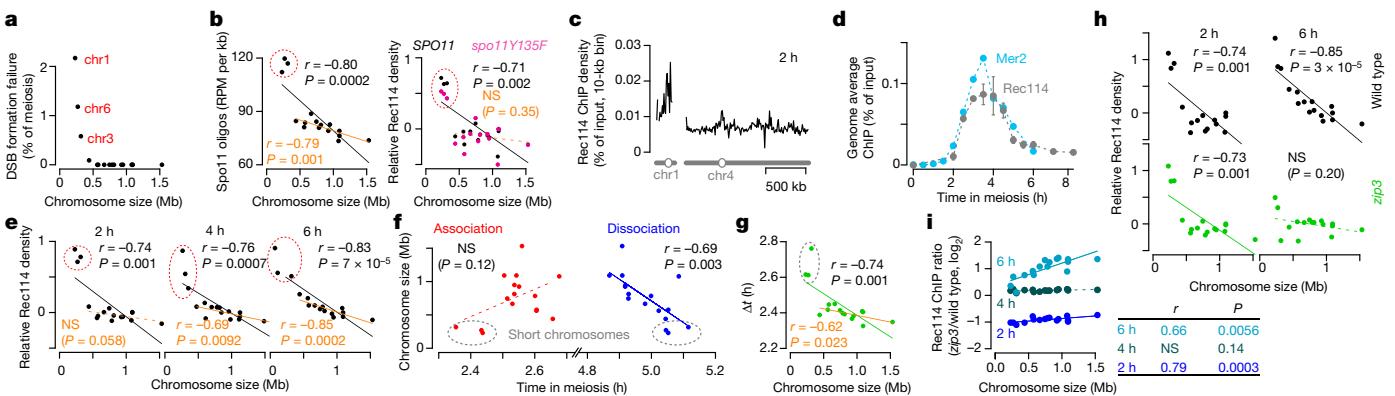
### Rec114 persists on short chromosomes

We studied published<sup>18</sup> and new meiotic time courses obtained from the analysis of myc-tagged Rec114 or Mer2 by calibrated chromatin immunoprecipitation followed by sequencing (ChIP-seq) (Fig. 1c, d, Extended Data Fig. 1b, c). Two key patterns emerged from the relative abundances of these proteins on each chromosome (Fig. 1e, Extended Data Fig. 1d, e). First, both proteins were overrepresented on short

chromosomes throughout. Second, the ChIP densities of these proteins were markedly different between early and late time points. At early times (2–3 h), the short chromosomes had notably higher relative ChIP densities, whereas the remaining chromosomes showed low protein densities that did not correlate with chromosome size. At the later times (4–6 h), the 13 largest chromosomes showed a clear negative correlation between ChIP density and size, and the three smallest chromosomes deviated less from a linear relationship.

These patterns suggested that distinct mechanisms regulate early and late chromosomal binding. To test this, we measured the protein–chromosome association and dissociation times at each protein peak. This revealed large early and late domains along chromosomes (Extended Data Fig. 2a). Nevertheless, on average, Rec114 associated earlier and dissociated later from the shortest chromosome (chr1) than from the longest chromosome (chr4) (Extended Data Fig. 2a), and this trend extended in a size-related manner across the entire complement (Fig. 1f). Notably, the patterns for each chromosome mirrored the ChIP-density patterns. The shortest chromosomes had early association times, whereas the rest showed no clear relationship. Across all chromosomes, dissociation times showed a negative

<sup>1</sup>Molecular Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>2</sup>Louis V. Gerstner, Jr., Graduate School of Biomedical Sciences, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>3</sup>Weill Graduate School of Medical Sciences, Cornell University, New York, NY, USA. <sup>4</sup>Howard Hughes Medical Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>5</sup>Present address: Ann Romney Center for Neurologic Disease, Department of Neurology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. <sup>6</sup>Present address: Caribou Biosciences, Inc., Berkeley, CA, USA. ✉e-mail: murakamh@mskcc.org; s-keeney@ski.mskcc.org



**Fig. 1 | Rec114 and Mer2 accumulate preferentially on smaller chromosomes.** **a**, Simulated rate of the failure to form DSBs on *S. cerevisiae* chromosomes, assuming a random distribution of DSBs. For each simulation, 200 DSBs were distributed randomly by chromosome size; *S. cerevisiae* makes 150–200 DSBs per meiosis<sup>15</sup>. The plot shows the percentage of simulations (10,000 total) in which the indicated chromosome failed to acquire a DSB. **b**, Dependence of DSBs (left, Spo11-oligonucleotide density in reads per million (RPM) per kb, 4 h,  $n=28$  maps), and Rec114 binding (right, ChIP-chip enrichment, 4 h,  $n=1$  culture for each strain) on chromosome size. In all figures, chr12 is represented without including its full rDNA cluster. Data are from refs. <sup>3,38</sup>. For the relative Rec114 density, the mean ChIP signal for each chromosome was normalized to the mean for chr15 and log<sub>2</sub>-transformed. All correlations reported are Pearson's  $r$ . **c**, Example Rec114 ChIP-seq for chr1 and chr4 at 2 h.

Absolute ChIP signal was calibrated by quantitative PCR (qPCR) and smoothed with a 10-kb sliding window. **d**, Time course of genome-average Rec114 and Mer2 levels. Grey points are mean  $\pm$  range of two Rec114 datasets. Cyan points are from the single Mer2 dataset. **e**, Dependence of the per-chromosome Rec114 ChIP density on chromosome size. The full time course is shown in Extended Data Fig. 1d. **f**, Per-chromosome association and dissociation times for Rec114. **g**, Per-chromosome duration of Rec114 binding. Data in **c** and **e–g** are from the *ARS<sup>+</sup>* time course ( $n=1$  culture); other time courses are in Extended Data Figs. 1, 2. **h**, **i**, Zip3-dependent homologue engagement shapes Rec114 abundance late in prophase. Graphs show per-chromosome Rec114 ChIP density (normalized to chr15) (**h**) and log-transformed fold change in the *zip3* mutant (**i**).  $n=1$  culture for each strain. In **b**, **e** and **g**, orange lines indicate correlations excluding the three short chromosomes.

correlation with size, with the shortest three fitting a global linear relationship (Fig. 1f).

For Rec114, the mean duration of binding on each chromosome correlated negatively with chromosome size, and the shortest three chromosomes again stood out with especially long durations (Fig. 1g, Extended Data Fig. 2b). These patterns were highly reproducible (Extended Data Fig. 2c, d) and Mer2 gave similar results (Extended Data Fig. 2b, d).

If the duration of Rec114 and other pro-DSB factors binding to the chromosome dictates for how long chromosomes are DSB-competent<sup>2,3,7,19</sup>, then Rec114- and Mer2-binding patterns should be indicative of DSB distributions. To test this, we measured DSBs at 4 h and 6 h by sequencing Spo11 oligonucleotides, which are quantitative byproducts of DSBs<sup>15</sup>. Spo11-oligonucleotide densities correlated negatively with chromosome size, but the short three chromosomes had substantially higher density than would be predicted for a linear relationship (Extended Data Fig. 2e). Moreover, the 13 largest chromosomes showed a better negative correlation at 6 h than at 4 h. The distribution of DSBs therefore mirrors the distributions of Rec114 and Mer2.

From these results we infer that the association and dissociation of DSB proteins with chromosomes controls the number of DSBs on all chromosomes in a manner dependent on their size, and ensures the overrepresentation of DSBs and crossovers on very short chromosomes. We next set out to elucidate the mechanisms of this temporal regulation.

## Homologue engagement dictates dissociation

The formation of DSBs is inhibited when chromosomes engage their homologues, and more effective inhibition on longer chromosomes establishes a negative correlation between DSB density and chromosome size<sup>2,7,16,20</sup>. We therefore proposed the following: the association and the dissociation of DSB proteins are controlled by distinct mechanisms; shorter chromosomes take longer to engage their partners; and homologue engagement displaces DSB proteins (Supplementary Discussion 1), thereby ensuring a negative correlation between the binding duration of Rec114 or Mer2 on the chromosome and the chromosome size. This hypothesis predicts that disrupting homologue engagement

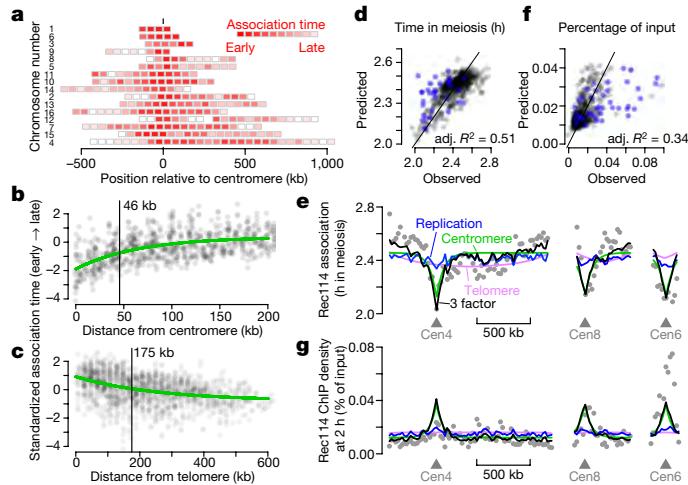
should not affect early Rec114 binding, but should cause the inappropriate retention of Rec114 on larger chromosomes.

We tested this prediction using a *zip3* mutant strain, which is defective in homologue engagement<sup>16</sup>. Compared with the wild-type *ZIP3*, modest differences were seen in the genome-wide absolute Rec114 ChIP signal at 2 h and 4 h, possibly reflecting small differences in culture timing; however, an approximately twofold increase at 6 h was consistent with the persistence of Rec114 for longer (Extended Data Fig. 3a–c). At 2 h and 4 h, Rec114 was overrepresented on the shortest chromosomes in *zip3*, similar to wild type (Fig. 1h, Extended Data Fig. 3b, d). By contrast, at 6 h, the mutant failed to establish a negative correlation between ChIP density and chromosome size (Fig. 1h) because of the preferential retention of Rec114 on larger chromosomes (Fig. 1i, Extended Data Fig. 3b, c). We conclude that feedback from homologue engagement establishes much of late-prophase DSB control by governing the dissociation of Rec114 and other proteins.

## Three pathways regulating association

Replication regulates the binding of Rec114 to chromatin<sup>18</sup>, so the early replication of short chromosomes<sup>21</sup> might explain their early association with Rec114. We tested this by deleting origins on the left arm of chr3 (chr3L) to delay replication and using a *tof1* mutation to compromise the coordination between replication and DSB formation<sup>18</sup>. We found that the overrepresentation of Rec114 or Mer2 and their early association times on short chromosomes were partially dependent on early replication (Extended Data Fig. 4a, b). We therefore propose that the coordination of replication contributes to the differences observed with chromosome size, but is not sufficient to explain the full extent of these differences; additional controls must therefore exist.

To delineate these controls, we examined subchromosomal domains. Colour-coded maps suggested early association and higher ChIP densities around centromeres<sup>22</sup>, and the converse towards telomeres (Fig. 2a, Extended Data Fig. 4c, d). Fitting trend lines to all datasets and all 32 chromosome arms quantified these effects and showed how they diminished with distance (Fig. 2b, c, Extended Data Fig. 4e, f). The telomere effect may reflect a known DSB formation delay<sup>23</sup>. Both



**Fig. 2 | The timing of replication and the proximity to centromeres and telomeres influence the timing and the level of Rec114 association.** **a**, Intra-chromosomal distributions of Rec114 association times in the *arsΔ* strain. Each block is a 50-kb bin colour-coded to the average Rec114 association times for peaks within the bin. Chromosomes are ranked by size, with the centromere at position zero. See also Extended Data Fig. 4c, d. **b, c**, Effects of the distance from the centromere (**b**) and the telomere (**c**) on the association time of DSB proteins. The two Rec114 and one Mer2 ChIP time courses were combined as described in Methods. Each point is the value from one dataset for a 20-kb bin. Darker colour indicates overlapping points. Green lines are exponential fits. Vertical bars are distances at which the effects decay to half the original value. **d, f**, Replication index and centromere and telomere effects were binned in 20-kb windows and used as explanatory variables to model the Rec114 association time (**d**,  $n = 539$  bins) or ChIP density at 2 h (**f**,  $n = 597$  bins) in the *arsΔ* strain by multiple linear regression. Each point is a bin; bins on short chromosomes are blue. **e, g**, Examples of within-chromosome patterns predicted by the multiple regression model and each of its component factors. Grey dots are average Rec114 association times (**e**) or Rec114 ChIP density at 2 h (**g**) in 20-kb bins, in *arsΔ* datasets. Blue, green and magenta lines are the component replication, centromere and telomere effects, respectively, and the black line is the prediction from the three-factor regression model.

effects were retained in *tof1* mutants but appeared weaker (Extended Data Fig. 4g), possibly because of constitutively early (centromeres) or late (telomeres) replication<sup>21</sup>.

Additional early time points in the Mer2 dataset enabled a more detailed view. Pericentromeric enrichment was detectable at 0.5 h, reached a maximum at 1.5 h, then diminished as Mer2 accumulation elsewhere balanced binding near centromeres (Extended Data Fig. 4h). By contrast, there was little telomere-proximal depletion of Mer2 up to 1 h, after which the depletion became progressively more apparent (Extended Data Fig. 4i). The centromere effect was normal in the *zip3* mutant (Extended Data Fig. 4j), which is consistent with homologue engagement influencing DSB protein binding at late but not at early stages of prophase.

Using multiple linear regression, we tested whether replication timing as well as the centromere and telomere effects might explain the observed association of DSB proteins. Regression models accounted for 37% to 51% of the variance in association timing (Fig. 2d, Extended Data Fig. 5a). Therefore, a simple three-factor model provided a reasonable fit across the whole genome (Fig. 2e). However, a poor fit was seen for the shortest chromosomes, with association observed earlier than predicted by the model (Fig. 2d, Extended Data Fig. 5a, b). Similar results were seen with the ChIP density at 2 h: models fit the genome-wide data well, but underperformed for small chromosomes by predicting less enrichment than was observed (Fig. 2f, g, Extended Data Fig. 5c, d).

Thus, for most chromosomes, association of DSB proteins early in meiotic prophase is shaped by the timing of replication and the

proximity to the centromere or telomeres. The shortest three chromosomes, however, accumulate these proteins earlier and at higher levels than these influences predict. We therefore proposed that small chromosomes have an additional feature(s) boosting their ability to compete, at early stages, for binding to a limited pool of DSB proteins. If this feature is intrinsic to the DNA sequence, two predictions arise: first, segments from short chromosomes should retain this boost when fused to a longer chromosome; second, making an artificially small chromosome by bisecting a larger one should not be sufficient to establish a similar effect.

## Boosting short chromosomes

To test whether chr1 (230 kb) intrinsically boosts Rec114 binding, we artificially lengthened it by reciprocal translocation with chr4 (1.5 Mb) (Fig. 3a, Extended Data Fig. 5e, f). We asked whether the chr1 portions of derivative chromosomes (der(1) at 532 kb and der(4) at 1.2 Mb) still behaved like a short chromosome, as predicted, or whether they now behaved like a longer chromosome, as previous studies of crossing-over might predict<sup>14</sup>.

At 2 h, segments from chr1 retained Rec114 overrepresentation (Fig. 3b), yielding sharp ChIP transitions between chr4- and chr1-derived sequences (Fig. 3c). Moreover, a three-factor regression model again predicted lower levels of Rec114 on chr1-derived sequences than were observed (Extended Data Fig. 5g). These results support the hypothesis that chr1 has an intrinsic feature(s) promoting preferential early Rec114 association. This feature acts *in cis*, and independently of chromosome size.

At later timepoints (4 h and 6 h), translocated chr1 segments still showed Rec114 overrepresentation, but to a lesser extent than was seen for native chr1 (Fig. 3b, Extended Data Fig. 5h, i). A lower Rec114 abundance matches our conclusion that feedback from homologue engagement dominates late patterns: because homologue engagement is linked to chromosome size, chr1-derived segments should conform to their long-chromosome context. The remaining overrepresentation relative to naturally long chromosomes may be a remnant of high early enrichment.

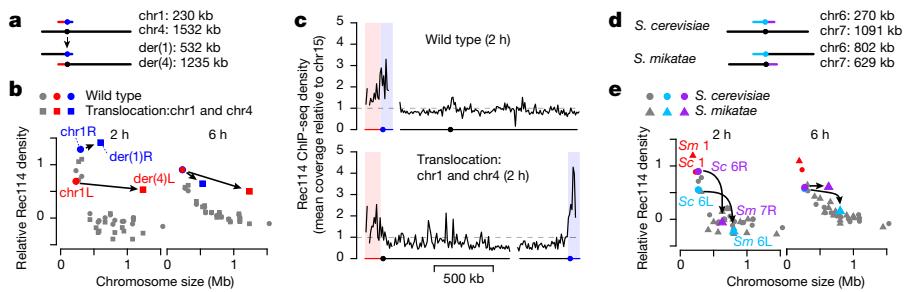
A converse experiment created an artificial short chromosome (177 kb) by translocation between two medium-size chromosomes. As predicted, this chromosome did not behave like a natural short chromosome in early prophase, when effects of the boost should be apparent, but did so later, when homologue engagement dominates (Extended Data Fig. 6a–f, Supplementary Discussion 2).

## Selective pressure maintains the boost

Most *Saccharomyces* species have the same three short chromosomes<sup>24</sup>. This conservation suggests that mechanisms mitigating the risk of meiotic nondisjunction are shared—implying, in turn, that evolutionary selection maintains the hyperrecombinogenic properties of small chromosomes.

*Saccharomyces mikatae* is a natural choice for studying the hallmarks of such selection. In other species, chr6 is the second shortest chromosome, but in *S. mikatae* the regions syntenic to ancestral chr6 (hereafter, syn6) are on longer chromosomes<sup>24,25</sup> (Fig. 3d). We previously showed that the DSB densities of syn6 match the chromosomal context, but inferred that density is tied to chromosome size and not to DNA sequence<sup>26</sup>. To revisit this conclusion, we reasoned that a lack of selective pressure to maintain a DSB-protein-binding boost on syn6 during post-translocation generations would lead to absence of the boost in extant *S. mikatae* strains. Chr1 should retain the boost because it is still small.

As predicted, chr1 in *S. mikatae* showed a strong overrepresentation of Rec114 at 2 h, but Rec114 binding was lower on syn6 in *S. mikatae* than on chr6 in *S. cerevisiae*. Indeed, syn6 segments were indistinguishable



**Fig. 3 | The Rec114 boost is intrinsic to short chromosomes.** **a**, Targeted translocation between chr1 and chr4 (to scale). **b**, Chr1-derived sequences retain high-level Rec114 binding when in a large-chromosome context. Per-chromosome Rec114 ChIP densities normalized to chr15 are shown at 2 h and 6 h. **c**, Rec114 profiles for wild-type and translocated chromosomes. ChIP-

seq data were normalized relative to chr15 (dashed line) and smoothed with a 10-kb sliding window. **d, e**, Selective pressure has maintained the short chromosome boost on chr6. Shown are a to-scale schematic of syntenic regions on different-sized chromosomes in *S. cerevisiae* (*Sc*) and *S. mikatae* (*Sm*) (**d**) and per-chromosome Rec114 ChIP densities (**e**).

from larger chromosomes (Fig. 3e, Extended Data Fig. 6g). A three-factor regression model fit the data from syn6 well, but still underperformed for chr1 (Extended Data Fig. 6h). At later timepoints (4 h and 6 h), the syn6 segments had relative Rec114 densities that were in line with the sizes of their chromosomes (Fig. 3e, Extended Data Fig. 6i, j).

### The boost requires axis proteins

Meiotic chromosomes form axial structures that anchor chromatin loops<sup>27</sup>. Because DSB proteins assemble on axes, arranging a DNA segment as short loops on a long axis is proposed to yield higher DSB density than if the segments are arrayed as long loops on a short axis<sup>28–30</sup>. The analysis of published data<sup>31</sup> showed that chr3 has an approximately 1.2-fold larger axis:DNA ratio (3.89 μm per Mb on average) than chr4 or chr15 (3.22 μm per Mb) (Extended Data Fig. 7a), although the densities of preferred DSB-protein-binding sites were not different between small and large chromosomes (Extended Data Fig. 7b). These findings are consistent with the loop-axis structure contributing to the intrinsic potential to form DSBs.

The axis proteins Hop1, Red1 and Rec8 promote normal chromatin association of Rec114 and Mer2, and Hop1 and Red1 are overrepresented on short chromosomes; this suggests that enrichment of axis proteins contributes to high DSB density<sup>3,17</sup>. To test this hypothesis, we assessed the ChIP density of Rec114 and the formation of DSBs in axis mutants.

Both *hop1* and *red1* single mutants eliminated the overrepresentation of Rec114 on small chromosomes at early timepoints (2 h) and ablated the negative correlation of Rec114 binding with chromosome size at all times (Extended Data Fig. 7c). Consistent with a previous study<sup>3</sup>, both mutations greatly decreased levels of Rec114 (as assessed by ChIP) across the genome; however, unexpectedly, we found that Rec114 binding was substantially higher in a *hop1 red1* double mutant, without rescue of either the DSB defects or Rec114 spatial patterns (Fig. 4a, Extended Data Fig. 7d–g, Supplementary Discussion 3). Even with this more robust ChIP signal, short chromosomes still lacked Rec114 overrepresentation at 2 h, and only a weak size dependence emerged at later timepoints (Fig. 4b, Extended Data Fig. 7c).

The absence of Rec8 did not eliminate the relative enrichment of Rec114 on short chromosomes<sup>17</sup> (Extended Data Fig. 7c), although total Rec114 binding was greatly reduced (Fig. 4a). A *hop1 red1 rec8* triple mutant behaved like *hop1 red1*, showing improved Rec114 recruitment but no short-chromosome boost (Fig. 4a, b). Rec8—which is enriched at centromeres<sup>22,32</sup>—was required for preferential Rec114 binding near centromeres, but Hop1 and Red1 were not (Extended Data Fig. 7h).

Spo11-oligonucleotide maps demonstrated the functional importance of the loss of Hop1 and Red1: mutants lacked a higher density of DSBs on short chromosomes (Fig. 4c). Instead, the observation of a greater number of DSBs with increased chromosome size suggested the complete loss of size-dependent control. These findings implicate the chromosome axis as a platform that supports DSB regulation.

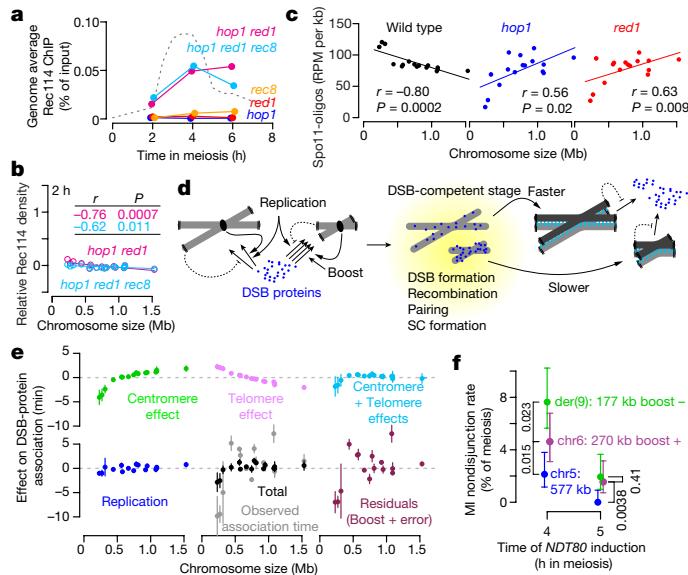
### Safeguarding chromosome segregation

Early in prophase, three mechanisms—dependent on replication timing, the distance to the centromere and the distance to the telomere—govern the spatial and temporal patterns of Rec114 and Mer2 binding to all chromosomes, and an additional mechanism(s) boosts binding on the smallest chromosomes (Fig. 4d). Each mechanism differs in magnitude and contributes differently to the dependence on chromosome size (Fig. 4e, Supplementary Discussion 4). Another pathway—homologue engagement—primarily regulates the dissociation of DSB proteins from the chromosome (Fig. 4d, Extended Data Fig. 7i, Supplementary Discussion 4). Within the context of homologue engagement, a tendency for chromosome end-adjacent regions to prolong DSB formation<sup>33</sup> contributes modestly when compared with other feature(s) such as size-dependent pairing kinetics (Extended Data Figs. 8, 9, Supplementary Discussion 5, 6).

We propose that these pathways co-operate to govern the amount of DSB protein that binds to a chromosome and the duration of a DSB-permissive state, thus shaping the ‘DSB potential’ within each chromosome and for each chromosome as a whole (Supplementary Discussion 7). The early pathways proactively establish when chromosomal segments become competent to make DSBs, so they modulate the population-average DSB probability. By contrast, feedback from homologue engagement reacts to a favourable outcome, so it not only affects the population-average DSB probability, but also ensures a low failure rate by allowing DSB formation to continue until success is achieved. Regulatory circuits involving Tel1<sup>ATM</sup> and Mec1<sup>ATR</sup> further ensure success<sup>2,7,34–37</sup>.

The purpose of this pathway integration is presumably to ensure that every chromosome has the opportunity to pair and recombine. To test this idea, we disrupted reactive pathways by inducing premature exit from prophase I with an exogenously controlled Ndt80 transcription factor (Supplementary Discussion 8). As predicted, early Ndt80 induction increased the frequency of meioses in which the artificial short chromosome (der(9)) was non-exchanged and suffered meiosis I nondisjunction (Extended Data Fig. 10a–c). Accurate segregation of a naturally short chromosome (chr6) was more sensitive to premature prophase exit than that of a mid-sized chromosome (chr5), and der(9)—which lacks the boost—was more sensitive still (Fig. 4f, Extended Data Fig. 10d, e).

Targeted boosting of DSB protein binding—perhaps by organizing short loops on long axes—may be a versatile strategy through which to mitigate the risk of meiotic missegregation caused by karyotypic constraints, for example the restriction of sex-chromosome recombination to the pseudoautosomal region in mammalian males<sup>28</sup> (Extended Data Fig. 10f, Supplementary Discussion 9). The multilayered control exhibited by budding yeast provides a paradigm for how cells solve the challenge of ensuring recombination on every chromosome, no matter how small.



**Fig. 4 | The role of axis proteins in the short chromosome boost and an integrated view of DSB control.** **a**, Genome-average Rec114 levels. The dashed line indicates the wild-type profile from Fig. 1d. **b**, The dependence of Rec114 binding on chromosome size requires Hop1 and Red1. Although correlations with size are statistically significant, slopes are negligible compared to that of the wild type (Fig. 1e). **c**, The preferential formation of DSBs on small chromosomes is lost in *hop1* and *red1* mutants. Samples were collected at 4 h in meiosis.  $n=28$  (wild type), 2 (*hop1* and *red1*) independent maps. **d**, Schematic depicting the multiple pathways that govern DSB protein association and dissociation from long compared with short chromosomes. **e**, Size dependence of each early pathway and their integration. Net effects for each chromosome (relative to mean association time) were estimated from the three-factor models in Fig. 2d and Extended Data Fig. 5a (points and vertical bars are means  $\pm$  s.d. of the three datasets). The top graphs show centromere and telomere effects and their combination. The bottom left graph shows the effect of replication timing (the high s.d. observed for chr3 is from *arsΔ*). The bottom middle graph merges all three pathways and compares this total to observed association time, and the bottom right graph depicts what is not explained by these three pathways, inferred as the short chromosome boost. See Supplementary Discussion 4 for further details. **f**, Small chromosomes are at greater risk of missegregation when DSB control is dysregulated. Chromosome missegregation was scored using spore-autonomous fluorescent reporters<sup>39</sup> (Extended Data Fig. 10d, e) in a strain with *NDT80* under inducible control. Premature *NDT80* expression (4 h,  $n=563$  tetrads) caused more missegregation for smaller chromosomes, especially for der(9), which lacks the small-chromosome boost. Segregation became more accurate when *NDT80* was induced later (5 h,  $n=516$  tetrads) to allow more time for DSBs to continue to accumulate; because of feedback tied to homologue engagement, a longer time is predicted to provide greater benefit to smaller chromosomes, as was observed. Error bars represent 95% confidence intervals. *P* values are from one-sided Fisher's exact test.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2248-2>.

- Hunter, N. Meiotic recombination: the essence of heredity. *Cold Spring Harb. Perspect. Biol.* **7**, a016618 (2015).
- Keeney, S., Lange, J. & Mohibullah, N. Self-organization of meiotic recombination initiation: general principles and molecular pathways. *Annu. Rev. Genet.* **48**, 187–214 (2014).
- Panizza, S. et al. Spo11-accessory proteins link double-strand break sites to the chromosome axis in early meiotic recombination. *Cell* **146**, 372–383 (2011).

- Arora, C., Kee, K., Maleki, S. & Keeney, S. Antiviral protein Ski8 is a direct partner of Spo11 in meiotic DNA break formation, independent of its cytoplasmic role in RNA metabolism. *Mol. Cell* **13**, 549–559 (2004).
- Maleki, S., Neale, M. J., Arora, C., Henderson, K. A. & Keeney, S. Interactions between Mei4, Rec114, and other proteins required for meiotic DNA double-strand break formation in *Saccharomyces cerevisiae*. *Chromosoma* **116**, 471–486 (2007).
- Li, J., Hooker, G. W. & Roeder, G. S. *Saccharomyces cerevisiae* Mer2, Mei4 and Rec114 form a complex required for meiotic double-strand break formation. *Genetics* **173**, 1969–1981 (2006).
- Carballo, J. A. et al. Budding yeast ATM/ATR control meiotic double-strand break (DSB) levels by down-regulating Rec114, an essential component of the DSB-machinery. *PLoS Genet.* **9**, e1003545 (2013).
- Kumar, R., Bourbon, H. M. & de Massy, B. Functional conservation of Mei4 for meiotic DNA double-strand break formation from yeasts to mice. *Genes Dev.* **24**, 1266–1280 (2010).
- Robert, T. et al. The TopoVIB-like protein family is required for meiotic DNA double-strand break formation. *Science* **351**, 943–949 (2016).
- Stanzione, M. et al. Meiotic DNA break formation requires the unsynapsed chromosome axis-binding protein IHO1 (CCDC36) in mice. *Nat. Cell Biol.* **18**, 1208–1220 (2016).
- Tessé, S. et al. Asy2/Mer2: an evolutionarily conserved mediator of meiotic recombination, pairing, and global chromosome compaction. *Genes Dev.* **31**, 1880–1893 (2017).
- Mancera, E., Bourgon, R., Brozzi, A., Huber, W. & Steinmetz, L. M. High-resolution mapping of meiotic crossovers and non-crossover in yeast. *Nature* **454**, 479–485 (2008).
- Chen, S. Y. et al. Global analysis of the meiotic crossover landscape. *Dev. Cell* **15**, 401–415 (2008).
- Kaback, D. B., Guacci, V., Barber, D. & Mahon, J. W. Chromosome size-dependent control of meiotic recombination. *Science* **256**, 228–232 (1992).
- Pan, J. et al. A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. *Cell* **144**, 719–731 (2011).
- Thacker, D., Mohibullah, N., Zhu, X. & Keeney, S. Homologue engagement controls meiotic DNA break number and distribution. *Nature* **510**, 241–246 (2014).
- Sun, X. et al. Transcription dynamically patterns the meiotic chromosome-axis interface. *eLife* **4**, (2015).
- Murakami, H. & Keeney, S. Temporospatial coordination of meiotic DNA replication and recombination via DDK recruitment to replisomes. *Cell* **158**, 861–873 (2014).
- Wojtasz, L. et al. Mouse HORMAD1 and HORMAD2, two conserved meiotic chromosomal proteins, are depleted from synapsed chromosome axes with the help of TRIP13 AAA-ATPase. *PLoS Genet.* **5**, e1000702 (2009).
- Kauppi, L. et al. Numerical constraints and feedback control of double-strand breaks in mouse meiosis. *Genes Dev.* **27**, 873–886 (2013).
- Blitzblau, H. G., Chan, C. S., Hochwagen, A. & Bell, S. P. Separation of DNA replication from the assembly of break-competent meiotic chromosomes. *PLoS Genet.* **8**, e1002643 (2012).
- Kugou, K. et al. Rec8 guides canonical Spo11 distribution along yeast meiotic chromosomes. *Mol. Biol. Cell* **20**, 3064–3076 (2009).
- Borde, V., Goldman, A. S. & Lichten, M. Direct coupling between meiotic DNA replication and recombination initiation. *Science* **290**, 806–809 (2000).
- Fischer, G., James, S. A., Roberts, I. N., Oliver, S. G. & Louis, E. J. Chromosomal evolution in *Saccharomyces*. *Nature* **405**, 451–454 (2000).
- Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E. S. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**, 241–254 (2003).
- Lam, I. & Keeney, S. Nonparadoxical evolutionary stability of the recombination initiation landscape in yeast. *Science* **350**, 932–937 (2015).
- Kleckner, N. Chiasma formation: chromatin/axis interplay and the role(s) of the synaptonemal complex. *Chromosoma* **115**, 175–194 (2006).
- Acquaviva, L. et al. Ensuring meiotic DNA break formation in the mouse pseudoautosomal region. *Nature* (in the press).
- Kauppi, L. et al. Distinct properties of the XY pseudoautosomal region crucial for male meiosis. *Science* **331**, 916–920 (2011).
- Kleckner, N., Storlazzi, A. & Zickler, D. Coordinate variation in meiotic pachytene SC length and total crossover/chiasma frequency under conditions of constant DNA length. *Trends Genet.* **19**, 623–628 (2003).
- Zhang, L. et al. Topoisomerase II mediates meiotic crossover interference. *Nature* **511**, 551–556 (2014).
- Blat, Y., Protacio, R. U., Hunter, N. & Kleckner, N. Physical and functional interactions among basic chromosome organizational features govern early steps of meiotic chiasma formation. *Cell* **111**, 791–802 (2002).
- Subramanian, V. V. et al. Persistent DNA-break potential near telomeres increases initiation of meiotic recombination on short chromosomes. *Nat. Commun.* **10**, 970 (2019).
- Lange, J. et al. ATM controls meiotic double-strand-break formation. *Nature* **479**, 237–240 (2011).
- Zhang, L., Kim, K. P., Kleckner, N. E. & Storlazzi, A. Meiotic double-strand breaks occur once per pair of (sister) chromatids and, via Mec1/ATR and Tel1/ATM, once per quartet of chromatids. *Proc. Natl. Acad. Sci. USA* **108**, 20036–20041 (2011).
- Gray, S., Allison, R. M., Garcia, V., Goldman, A. S. & Neale, M. J. Positive regulation of meiotic DNA double-strand break formation by activation of the DNA damage checkpoint kinase Mec1(ATR). *Open Biol.* **3**, 130019 (2013).
- Cooper, T. J., Wardell, K., Garcia, V. & Neale, M. J. Homeostatic regulation of meiotic DSB formation by ATM/ATR. *Exp. Cell Res.* **329**, 124–131 (2014).
- Mohibullah, N. & Keeney, S. Numerical and spatial patterning of yeast meiotic DNA breaks by Tel1. *Genome Res.* **27**, 278–288 (2017).
- Thacker, D., Lam, I., Knop, M. & Keeney, S. Exploiting spore-autonomous fluorescent protein expression to quantify meiotic chromosome behaviors in *Saccharomyces cerevisiae*. *Genetics* **189**, 423–439 (2011).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

# Article

## Methods

### Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

### Yeast strain and plasmid construction

#### Myc-epitope tagging of Rec114 and Mer2 in *S. cerevisiae* and *S. mikatae*.

To perform ChIP-seq experiments in *S. cerevisiae*, Rec114 or Mer2 were C-terminally tagged with 8 and 5 copies of the Myc-epitope marked with the *hphMX4* cassette and the *URA3* gene, respectively (*REC114-Myc* and *MER2-Myc*), as described previously<sup>18,40</sup>. Rec114 in *S. mikatae* was C-terminally tagged with the Myc-epitope with the same construct used to tag Rec114 in *S. cerevisiae*. Epitope-tagged Rec114 in *S. mikatae* was checked by western blotting and was found to be functional, as the *REC114-Myc* strain showed good spore viability (98%, 16 tetrads).

**Targeting reciprocal translocation between chr1 and chr4.** Reciprocal translocation between chr1 and chr4 was targeted as described in Extended Data Fig. 5e. The *TRP1* gene with a 3' portion of the *Kluyveromyces lactis URA3* gene (*KIURA3*) was amplified from plasmid pWJ716<sup>41</sup> using primers TL#1AF and TL#1AR (Supplementary Table 1), which each contain 50 nt from the terminator region of *SWC3* in the left arm of chr1. The *HIS3* gene with a 5' portion of *KIURA3* was amplified from pWJ1077<sup>41</sup> using primers TL#1BF and TL#1BR (Supplementary Table 1), which each contain 50 nt from the terminator region of *SLX5* in the left arm of chr4. Each amplified DNA fragment was transformed into *MATα* and *MATα* haploid yeast (*ura3, trp1, his3*), respectively. After verifying transformants by PCR and sequencing, *MATα* and *MATα* transformants were mated. Because the two *KIURA3* segments share an identical sequence (448 bp), homologous recombination between these regions would produce uracil prototrophy along with reciprocal translocation between chr1 and chr4. We sporulated the diploid and screened for *Ura<sup>+</sup>* haploids by spreading spores on SC-ura plates. A *Ura<sup>+</sup>* haploid was verified by pulsed-field gel electrophoresis followed by Southern blotting using probes hybridizing to both ends of chr1 and chr4 generated by primers listed in Supplementary Table 1 (see Extended Data Fig. 5f for an example). We confirmed that native chr1 and chr4 had disappeared and derivative chromosomes der(1) and der(4) of the expected size had appeared. The verified haploid was crossed with a *REC114-Myc* haploid which retains native chr1 and chr4 to isolate both mating types with der(1), der(4) and *REC114-Myc*. These haploids were verified by Southern blotting again and mated to obtain a diploid with the homozygous translocation in the *REC114-Myc* background.

**Targeting reciprocal translocation between chr8 and chr9.** Reciprocal translocation between chr8 and chr9 was targeted by CRISPR–Cas9 as described in Extended Data Fig. 6a. Two guide RNA sequences were cloned into pCRCT (*URA3*, iCas9, 2 μm origin of replication)<sup>42</sup> to target cleavages in the downstream regions of *YHL012w* (80468: chr8, left arm) and *URM1* (342918, chr9, left arm). The plasmid was cotransformed with 100-bp recombination donor fragments that have translocated sequences into a *MATα REC114-Myc* haploid. *Ura<sup>+</sup>* transformants were first screened on SC-ura plates and then checked for translocation by PCR with primer pairs flanking the two junctions. Positive transformants were mated with a wild-type *MATα* haploid. The resulting diploid turned out to be homozygous for the translocated chromosomes probably because of recombination induced by Cas9 cleavages using der(8) and der(9) as template. The sizes of the translocated chromosomes in the above haploids and diploids were confirmed by pulsed-field gel electrophoresis followed by Southern blotting using probes hybridizing to both ends of chr8 and chr9 generated by primers listed in Supplementary Table 1 (Extended Data Fig. 6b). Diploids that had lost the plasmid were selected on 5-FOA plates and subjected to sporulation

followed by tetrad dissection to isolate *MATα* and *MATα* haploids with der(8), der(9), and *REC114-Myc*. These haploids were mated and the resulting diploid was used for further experiments.

**Axis mutants.** The *red1* and *hop1* deletions were made by replacing the respective coding sequences with the hygromycin B drug resistance cassette (*hphMX4*) amplified from plasmid pMJ696 (identical to pAG32<sup>43</sup>). Yeasts were transformed using standard lithium acetate methods. Gene disruption was verified by PCR and Southern blotting. The *SPO11-Flag* construct (*SPO11-6His-3FLAG-loxP-kanMX-loxP*) was provided by K. Ohta, University of Tokyo<sup>22</sup>. All axis mutants (*hop1*, *red1*, and *rec8* single mutants, the *hop1 red1* double mutant, and the *hop1 red1 rec8* triple mutant) and *zip3* mutants in the *REC114-Myc* background were created by multiple crossing followed by tetrad dissection.

**Plasmids for spore-autonomous fluorescent markers.** Plasmids pSK1269 (*P<sub>YKLO50c</sub>-GFP\*-KanMX*) and pSK1271 (*P<sub>YKLO50c</sub>-CFP-natMX*) were constructed by subcloning EcoRI fragments from pSK726 and pSK692<sup>39</sup> into EcoRI sites on pFA6a-KanMX<sup>44</sup> and pMJ695 (identical to pAG25<sup>43</sup>), respectively. Fluorescent markers and drug-resistant cassettes are tandem orientation. For the integration of RFP into the right arm of chr9, a 247-bp sequence within the downstream of YPS6 open reading frame was amplified using the primers (inf9RightRFPF and inf9Right-RFPR) listed in Supplementary Table 1 and cloned at the Tth111I site in pSK691 (*P<sub>YKLO50c</sub>-RFP-LEU2*) using In-Fusion HD (TaKaRa), yielding an integrative plasmid (pSK1320).

**Construction of strains with inducible *NDT80* and spore-autonomous fluorescent markers.** To introduce the inducible *NDT80* system, a *GAL4-ER P<sub>GAL</sub>-NDT80* strain (from A. Amon, MIT)<sup>45</sup> was crossed with our strain with der(9). The resulting haploids from tetrad dissection were transformed with spore-autonomous fluorescent markers with 50-bp homology sequence amplified from pSK1269, pSK1271 or pSK691<sup>39</sup>. Transformants were checked by PCR designed at two junctions in the following combinations (integration marker and locus: primers to amplify marker; plasmid; primers to check integration): RFP at *MSH4* downstream (cen6): TF\_RFP\_cen6F/R; pSK691; cen6\_RFP\_check1F/cen1\_RFP\_check1R and cen1\_RFP\_check2F/cen6\_RFP\_check2R. GFP at *NAS2* downstream (translocation junction, cen9): TF\_GFP\_cen9v3F/R; pSK1269; cen9\_GFP\_check3F/cen9\_GFP\_check1v2R and cen9\_GFP\_check2F/cen9\_GFP\_check3R. CFP at *GIM4* downstream (cen5): TF\_CFP\_cen5F/R; pSK1271; cen5\_CFP\_check1F/R and cen5\_CFP\_check2F/R. CFP at *YHL048w* downstream (chr8L): TF\_CFP\_chr8LF/R; pSK1271; chr8L\_CFP\_check1F/R and chr8L\_CFP\_check2F/R. RFP integration at *YPS6* (chr9R) was achieved by integrating pSK1320 linearized with NdeI and transformants were checked by PCR using chr9R\_RFP\_check1F/R and chr9R\_RFP\_check2F/R. Strains with appropriate marker configuration (Extended Data Fig. 10, Supplementary Table 2) were created by crossing the above transformants followed by tetrad dissection.

### Yeast growth conditions

Studies were performed using *S. cerevisiae* SK1 and *S. mikatae* IFO1815 strain backgrounds; strains are listed in Supplementary Table 2. Synchronous meiotic cultures were obtained using the SPS pre-growth method<sup>46</sup>. In brief, saturated overnight cultures in 4 ml YPD (1% yeast extract, 2% peptone, 2% dextrose) were used to inoculate 25 ml of SPS (0.5% yeast extract, 1% peptone, 0.67% yeast nitrogen base without amino acids, 1% potassium acetate, 0.05 M potassium biphthalate, pH 5.5, 0.002% antifoam 204 (Sigma)) to a density of  $5 \times 10^6$  cells per ml and cultured at 30 °C at 250 rpm for 7 h. Cells were then inoculated into an appropriate volume (900 ml for ChIP-seq experiments with 12 time points or 300 ml for experiments with 3 time points) of fresh SPS at a density of  $3 \times 10^5$  cells per ml and cultured at 30 °C at 250 rpm for 12–16 h until the density reached  $3–4 \times 10^7$  cells per ml. Cells were collected by filtration, washed with water, then resuspended at  $4 \times 10^7$

cells per ml in an appropriate volume (610 ml for 12 time points or 200 ml for 3 time points) of potassium acetate sporulation medium (SPM; 2% potassium acetate, 0.001% polypropylene glycol) supplemented with 0.32% amino acid complementation medium (1.5% lysine, 2% histidine, 2% arginine, 1% leucine, 0.2% uracyl, 1% tryptophan). Cultures were shaken at 250 rpm at 30 °C and 50 ml samples for ChIP-seq were collected at the desired times after transfer to SPM. For cultures with 12 time points, we collected samples as follows: 0, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 7, 8 h for Rec114 ChIP in *TOF1* background; 0, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 6, 7 h for Rec114 ChIP in *tof1* background; 0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 6 h for Mer2 ChIP. For all cultures with three time points, cells were collected at 2, 4 and 6 h.

For the Spo11-oligonucleotide mapping experiments, synchronous meiotic cultures of *S. cerevisiae* SK1 were prepared as described previously<sup>47</sup>. In brief, cells from a saturated overnight yeast extract peptone dextrose culture were used to inoculate a 14-h pre-sporulation culture in YPA (1% yeast extract, 2% peptone, 1% potassium acetate) supplemented with 0.001% antifoam 204 and grown at 30 °C (starting optical density ( $OD_{600}$ ) of 0.2). Cells were collected and resuspended in 2% potassium acetate, 0.2× supplements (2 µg ml<sup>-1</sup> adenine, 2 µg ml<sup>-1</sup> histidine, 6 µg ml<sup>-1</sup> leucine, 2 µg ml<sup>-1</sup> tryptophan and 2 µg ml<sup>-1</sup> uracil), 0.001% antifoam 204 at  $OD_{600} = 6.0$ , and were then incubated in a 30 °C shaker to induce sporulation.

To assess culture synchrony, meiotic division profiles were obtained by collecting aliquots at various times from synchronous meiotic cultures, fixing in 50% (v/v) ethanol, and staining with 4',6-diamidino-2-phenylindole (DAPI; 0.05 µg ml<sup>-1</sup>). Mono-, bi- and tetranucleate cells were scored by fluorescence microscopy.

#### Chromatin immunoprecipitation for Mer2-Myc and Rec114-Myc

We performed ChIP experiments as described previously<sup>18</sup>, with modifications in cell disruption and chromatin fragmentation. Cells were disrupted by vigorous shaking at 6.5 m s<sup>-1</sup>, 10 times for 1 min each, using a FastPrep24 (MP Biomedicals). Chromatin in the whole-cell extracts was sheared by sonication with 'M' intensity, 30 s ON/30 s OFF for 3 × 15 mins in a Bioruptor Sonication System UCD200 (Diagenode) in 15 ml polystyrene conical tubes. The insoluble fraction (cell debris) was removed by centrifugation at 21,130 g, 5 min, 4 °C. Whole-cell extracts were further sonicated under the same conditions 3–5 times to yield an average DNA size of less than 500 bp.

For qPCR, we used eight and ten primer pairs for *S. cerevisiae* 12-time-point and 3-time-point datasets, respectively. For *S. mikatae* we used ten primer sets. All primer sets are listed in Supplementary Table 1. qPCR was performed using the LightCycler 480 SYBR Green I Master (Roche) according to the manufacturer's recommendations. All measurements of ChIP samples were expressed relative to the standard (dilution series of corresponding input samples).

#### Spo11-oligonucleotide mapping

For Spo11-oligonucleotide mapping, at least 600 ml of sporulation culture was collected 4 h after transfer to sporulation media. Because of the severe DSB defect in *red1* and *hop1*, Spo11 oligonucleotides from multiple (4–5) cultures of independent colonies were pooled to generate each Spo11-oligonucleotide map. The wild-type Spo11-oligonucleotide map in Fig. 4c was from a previous study<sup>38</sup>.

Spo11-oligonucleotide mapping in *red1* and *hop1* mutants was performed essentially as described previously<sup>26</sup>, with modifications to purify enough Spo11 oligonucleotides from *red1* and *hop1* strains. For example, Spo11 oligonucleotides from independent cultures were pooled after eluting from the first immunoprecipitation and at the last step of oligonucleotide purification (after proteinase K treatment and ethanol precipitation of the oligonucleotides). When pooling Spo11-oligonucleotide complexes from five cultures after the first immunoprecipitation step, the total volume of the second immunoprecipitation was increased to 4 ml, and 500 µl of Dynabeads Protein

G slurry were pre-bound to 100 µl of 1 mg ml<sup>-1</sup> anti-Flag antibody (as opposed to 125 µl of Dynabeads Protein G slurry pre-bound to 25 µl 1 mg ml<sup>-1</sup> anti-Flag antibody, and a second IP volume of 800 µl). Purified Spo11 oligonucleotides were quantified and used for library preparation as described previously<sup>16</sup>.

Sequencing (Illumina HiSeq 2500, 2 × 75 bp paired-end reads) was performed in the MSKCC Integrated Genomics Operation. Clipping of library adapters and mapping of reads was performed by the Bioinformatics Core Facility (MSKCC) using a custom pipeline as described<sup>15,16,26,38,48</sup>. Reads were mapped to the sacCer2 genome assembly of type strain S288C from SGD (*Saccharomyces* Genome Database).

#### ChIP-seq data processing: scaling and masking

ChIP-seq experiments were performed as described<sup>18</sup>. DNA from ChIP and input samples (same samples as used for ChIP-qPCR) were further sheared by sonication to an average fragment size of around 300 bp. These were sequenced (50 bp paired-end) on the HiSeq platform (Illumina). Reads were mapped to the sacCer2 genome assembly and the *S. mikatae* genome assembly<sup>49</sup> using BWA (0.7) MEM to generate coverage maps for each time point from each strain. Each ChIP coverage map was divided by the corresponding input map for normalization. Then, to scale the ChIP-seq coverage relative to absolute ChIP efficiency, we calculated the total coverage within ±1 kb of the centre of each qPCR amplicon, plotted these as a function of the corresponding qPCR ChIP efficiency, and calculated regression lines by least squares. The resulting regression line for each time point was then used to scale the ChIP-seq coverage maps.

To remove regions with spurious mapping, we previously defined 'mask regions' in which the coverage from the 0 h sample of the wild-type *ARS*<sup>+</sup> strain was out of a fixed range (>1.5 s.d. from mean coverage) with further extension by 1 kb on either side<sup>18</sup>. These regions were censored in all input and ChIP-coverage maps from *S. cerevisiae*. Mask regions for *S. mikatae* were defined similarly when the coverage from the 2 h input sample exceeded a fixed range (mean coverage ±4 s.d., calculated between the 50–150 kb region of chr1). After the same extension, these regions were censored from *S. mikatae* coverage maps.

#### Replication index generated by ChIP input coverage maps

All masked coverage maps from input samples were binned using 5-kb windows and normalized to genomic mean coverage. For 12-time-point datasets, coverage from an 'S-phase time point' (1.5 h and 2.5 h for Rec114 ChIP *ARS*<sup>+</sup> *tof1Δ* and Rec114 ChIP *arsΔtof1Δ*, respectively; 2 h for the rest) was divided by the corresponding 'G1-phase time point' (0-h sample) to generate a 'relative coverage' map. For 3-time-point datasets, the 2-h time point map was divided by the 0-h map from the Rec114 ChIP *ARS*<sup>+</sup> dataset to generate relative coverage. For the *S. mikatae* dataset, the mean normalized 2 h map was used as relative coverage. We defined the 'replication index' as  $-\log_2(\text{relative coverage})$ . Outliers were removed from each dataset, defined as the replication index value exceeding a fixed range (mean ± 4 s.d.).

#### Estimating association and dissociation times by sequential curve fitting

The method for measuring association time is described previously<sup>18</sup>. The scaled and masked ChIP-seq coverage maps from two Rec114 ChIP-seq and Mer2 ChIP-seq datasets were smoothed using a 2010-bp Parzen (triangular) sliding window. Using the smoothed, scaled coverage map at 3.5-h time points, a total of 1,477 (Rec114 ChIP *ARS*<sup>+</sup>), 1,545 (Rec114 ChIP *arsΔ*) and 1,550 (Mer2 ChIP) peaks were called using as a threshold of 0.5× of the mean coverage of each chromosome. A ChIP temporal profile at each peak position was assembled by collecting the ChIP signals from the smoothed, scaled coverage map for each time point.

To define the empirical maximum time in the ChIP profile ( $t_{\max}$ ), Gaussian curves were fitted to ChIP signals plotted as a function of time. To create positive skew in the regression curves, times ( $t$ , in hours)

# Article

were log-transformed [ $t' = \ln(t+1)$ ]. We used an equation that is a modification of the Gaussian probability density function:

$$y = a + b \times e^{-\frac{(t'-c)^2}{d^2}}$$

where  $y$  is ChIP signal,  $a$  is the background,  $b$  is the peak height,  $c$  is the peak position, and  $d$  is the equivalent of standard deviation. We set the background parameter ( $a$ ) to the ChIP signal at 0 h, then fitted the equation to the data points by least squares to estimate the other parameters ( $b$ ,  $c$  and  $d$ ) using the “nls” function in R. The estimated parameter ( $c$ ) was transformed back to hours in meiosis ( $t_{max} = e^c - 1$ ).

Next, to estimate the association time of the DSB protein, we used this peak to fit a saturating exponential growth (logistic) curve to just the upward slope of the ChIP temporal profile (data points before  $t_{max}$ ):

$$y = a + \frac{b}{1 + e^{c(d-t)}}$$

where  $y$  is ChIP signal,  $a$  is the background,  $b$  is the maximum value,  $c$  is a shaping factor and  $d$  is the inflection point of the logistic function, respectively. We set the background and the maximum value parameters ( $a$  and  $b$ ) to the ChIP signal at 0 h and the previously estimated peak height value (parameter ( $b$ ) from skewed Gaussian fitting,  $b_{Gauss}$ ), respectively, and then fitted the equation to the data points to estimate the other parameters ( $c$  and  $d$ ) using the “nls” function in R. We used  $d$  as  $t_{association}$  where the logistic curve reaches 50% of maximum.

We also estimated the dissociation time of the DSB protein by fitting a logistic curve to the downward slope of the ChIP temporal profile (data points after  $t_{max}$ ):

$$y = a + \frac{b}{1 + e^{c(t-d)}}$$

where  $y$  is the ChIP signal,  $a$  is the background,  $b$  is the maximum value,  $c$  is a shaping factor and  $d$  is the inflection point of the logistic function, respectively. We used the “nls” function to estimate parameters ( $c$  and  $d$ ) and used  $d$  as  $t_{dissociation}$  where the logistic curve reaches 50% of maximum.

To evaluate the fitting quality for the kinetic profile at each peak, absolute distances between the data points and the fitted Gaussian curve (residuals) were summed and divided by the peak height (parameter  $b_{Gauss}$ ) from the fitted curve (normalized-total residuals,  $r_{Gauss}$ ). Total residuals from two logistic fittings were divided by  $b_{Gauss}$ , and the sum of these was defined as  $r_{logistic}$ . We excluded poorly fitted peaks with normalized residuals exceeding 1.2 for Rec114 ChIP *arsΔ* and Mer2 ChIP datasets. We used less stringent criteria (filtering out peaks with  $r_{Gauss} > 1.6$  or  $r_{logistic} > 1.5$ ) for the Rec114 ChIP *ARS<sup>+</sup>* dataset because the overall quality of fittings was less good compared with the other two datasets. After these filtering steps, a total of 998 (Rec114 ChIP *ARS<sup>+</sup>*), 1,081 (Rec114 ChIP *arsΔ*) and 1,490 (Mer2 ChIP) peaks were processed for further analyses.

For the Rec114 association time in the two *tof1* datasets, we used previously estimated values for 957 (*tof1* *ARS<sup>+</sup>*) and 2,020 (*tof1* *arsΔ*) peaks that passed filtering<sup>18</sup>.

## Estimating centromere and telomere effects on DSB protein association and dissociation time, and ChIP density at 2 h and 6 h

Association and dissociation timing, and ChIP density (2 h and 6 h) data from the two Rec114 and one Mer2 ChIP time courses were combined as follows. To capture the intra-chromosomal features separate from inter-chromosomal differences, we averaged each dataset in 20-kb bins, standardized the values for each chromosome to a mean of 0 and variance of 1, and standardized again within a dataset before pooling all datasets together. The first standardization minimizes differences

between chromosomes and the second standardization minimizes differences between datasets. The pooled association, dissociation time or ChIP density (2 h or 6 h) data were plotted as a function of distance from the centromere or the telomere. We fitted an exponential decay model to these data points:

$$y = \frac{a}{e^{x/b}} + c$$

where  $y$  is association time, dissociation time, or ChIP density at 2 h or 6 h,  $x$  is distance from the centromere or the telomere,  $a$  is the initial value of the centromere effect or the telomere effect,  $b$  is a shaping factor and  $c$  is the intercept. We used the “nls” function in R to estimate parameters ( $a$ ,  $b$  and  $c$ ) and used the parameter ( $b$ ) to present the half distance, in which the initial value decays to half, as follows:

$$\text{Half distance} = b \times \ln(2)$$

For modelling of ChIP density at 6 h to assess the telomere effect, in order to represent DSB protein repression within 20 kb from the telomere and enrichment in the adjacent regions (about 20–100 kb) we fitted a composite model consisting of two exponential decay models:

$$y = \frac{a}{e^{x/b}} + \frac{a'}{e^{x/b'}} + c$$

where  $y$  is ChIP density at 6 h,  $x$  is the distance from the telomere,  $a$  and  $a'$  are the initial values of the repression and enrichment effects,  $b$  and  $b'$  are shaping factors and  $c$  is the intercept.

## Multiple linear regression analysis

To perform multiple regression analyses, the replication index, association time, dissociation time and ChIP density at 2 h or 6 h were averaged in 20-kb bins. Distances from the centromere and the telomere at the midpoints of the 20-kb bins were input into the centromere and telomere exponential models, the parameters of which were estimated as described in the preceding section. The same centromere and telomere models were used for multiple regression in strains with translocations and in *S. mikatae*. Regression coefficients and the standardized regression coefficients (beta) are shown along with  $t$  and  $P$  values based on the standardized coefficients in Supplementary Tables 3–5.

## Fluorescent spore assay to measure crossing over and meiosis I nondisjunction

Diploids with der(9), spore-autonomous fluorescent markers and inducible *NDT80* (SKY7023 and 7034) were sporulated using the SPS presporulation method described above. SPM culture (1 ml) was removed at the times indicated in Fig. 4f and Extended Data Fig. 10c and returned to the shaker after adding β-oestradiol (Sigma-Aldrich, final concentration 1 μM). Cells were collected 54 h after transfer to SPM and 200 (SKY7023) and more than 500 (SKY7034) tetrads per time point were scored as previously described<sup>39</sup>. We used a 40× objective lens and analysed the captured tetrad images using Fiji<sup>50</sup>. We scored only tetrads with four obvious spores. For tetrads for which the presence of fluorescent signals in spores was difficult to call positive or negative from visual inspection, we quantified the signal strength of spores within a tetrad and called a given spore positive if its signal exceeded two (RFP and CFP) or four (GFP) fold higher than a sister spore with lower signal. Tetrads with aberrant numbers of ‘positive’ spores other than two positive (SKY7023) and two or four positive (SKY7034) were excluded from further analysis.

For the strain designed to detect crossovers, nonexchange chromosomes (EO), and meiosis I (MI) nondisjunction (MINDJ) on der(9) (SKY7023) (Extended Data Fig. 10a, b), 200 tetrads per time point were scored. Tetrads with an aberrant number of positive spores other than

two positive were excluded (marker gain or loss in Supplementary Table 6). The rest of the tetrads were categorized as listed. The number of crossovers was estimated for each interval using the equation  $TT + 6NPD$ , where  $TT$  is tetratypes and  $NPD$  is nonparental ditypes. Because MI nondisjunction and double NPD events would produce tetrads with identical fluorescent signal patterns (Extended Data Fig. 10b), the number of MINDJ was corrected as: observed MINDJ – fraction( $NPD^{CFP-GFP}$ ) × fraction( $NPD^{GFP-RFP}$ ) × (total number of tetrads). Similarly, the number of EO was corrected as: observed EO + fraction( $NPD^{CFP-GFP}$ ) × fraction( $NPD^{GFP-RFP}$ ) × (total number of tetrads). On the basis of the assumption that most MINDJ events are also EO, the total EO was estimated as the sum of corrected EO and corrected MINDJ. We also independently estimated EO assuming a Poisson distribution of crossovers in the population of cells with the measured average crossover number (dashed line in Extended Data Fig. 10c). Because the two ways of estimating EO values agreed well, we conclude that these measurements are robust.

For the strain designed to detect MINDJ on der(9), chr6, and chr5 (SKY7034) (Extended Data Fig. 10d, e), more than 500 tetrads per time point were scored. Tetrads with two positive and two negative fluorescent spores were counted as MINDJ (Supplementary Table 7). We did not score tetrads showing numbers of positive spores other than four or two.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

All sequencing data have been deposited in the Gene Expression Omnibus (GEO) with the accession numbers GSE52970 (Rec114 ChIP-seq, including *tof1*), GSE84859 (Spo11 oligonucleotides in *hop1* and *red1*), GSE119786 (Mer2 ChIP-seq), GSE119787 (all Rec114 ChIP-seq data generated in this study) and GSE119689 (Spo11-oligonucleotide maps in the wild type at 4 h and 6 h).

## Code availability

Custom code for Spo11-oligonucleotide mapping has been previously published and is available online (see Methods for references).

40. Henderson, K. A., Kee, K., Maleki, S., Santini, P. A. & Keeney, S. Cyclin-dependent kinase directly regulates initiation of meiotic recombination. *Cell* **125**, 1321–1332 (2006).
41. Reid, R. J., Sunjevaric, I., Keddache, M. & Rothstein, R. Efficient PCR-based gene disruption in *Saccharomyces* strains using intergenic primers. *Yeast* **19**, 319–328 (2002).
42. Bao, Z. et al. Homology-integrated CRISPR-Cas (Hi-CRISPR) system for one-step multigene disruption in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **4**, 585–594 (2015).
43. Goldstein, A. L. & McCusker, J. H. Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* **15**, 1541–1553 (1999).
44. Wach, A., Brachat, A., Pöhlmann, R. & Philippsen, P. New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. *Yeast* **10**, 1793–1808 (1994).
45. Carlile, T. M. & Armon, A. Meiosis I is established through division-specific translational control of a cyclin. *Cel* **133**, 280–291 (2008).
46. Murakami, H., Borde, V., Nicolas, A. & Keeney, S. Gel electrophoresis assays for analyzing DNA double-strand breaks in *Saccharomyces cerevisiae* at various spatial resolutions. *Methods Mol. Biol.* **557**, 117–142 (2009).
47. Neale, M. J. & Keeney, S. End-labeling and analysis of Spo11-oligonucleotide complexes in *Saccharomyces cerevisiae*. *Methods Mol. Biol.* **557**, 183–195 (2009).
48. Zhu, X. & Keeney, S. High-resolution global analysis of the influences of Bas1 and Ino4 transcription factors on meiotic DNA break distributions in *Saccharomyces cerevisiae*. *Genetics* **201**, 525–542 (2015).
49. Scannell, D. R. et al. The awesome power of yeast evolutionary genetics: new genome sequences and strain resources for the *Saccharomyces* sensu stricto genus. *G3* **1**, 11–25 (2011).
50. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
51. Yue, J. X. et al. Contrasting evolutionary genome dynamics between domesticated and wild yeasts. *Nat. Genet.* **49**, 913–924 (2017).
52. Blat, Y. & Kleckner, N. Cohesins bind to preferential sites along yeast chromosome III, with differential regulation along arms versus the centric region. *Cell* **98**, 249–259 (1999).
53. Neale, M. J., Pan, J. & Keeney, S. Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature* **436**, 1053–1057 (2005).
54. Mimitou, E. P., Yamada, S. & Keeney, S. A global view of meiotic double-strand break end resection. *Science* **355**, 40–45 (2017).
55. Benjamin, K. R., Zhang, C., Shokat, K. M. & Herskowitz, I. Control of landmark events in meiosis by the CDK Cdc28 and the meiosis-specific kinase Ime2. *Genes Dev.* **17**, 1524–1539 (2003).
56. Picard, D. in *Nuclear Receptors: a Practical Approach* (ed. Picard, D.) 261–274 (Oxford Univ. Press, 1999).

**Acknowledgements** We thank A. Viale and N. Mohibullah of the Memorial Sloan Kettering Cancer Center (MSKCC) Integrated Genomics Operation for DNA sequencing; N. Socci at the MSKCC Bioinformatics Core Facility for mapping ChIP-seq and Spo11-oligo reads; and members of the Keeney laboratory, especially S. Yamada for advice on data analysis and L. Acquaviva for sharing unpublished information. We thank V. Subramanian, A. Hochwagen and F. Klein for discussions and for sharing unpublished information; and M. Lichten, E. Louis, K. Ohta, A. Armon, W. Zachariae, J. Matos and R. Rothstein for strains or plasmids. I.L. and M.v.O. were supported in part by National Institutes of Health (NIH) fellowships F31 GM097861 and F32 GM096692, respectively. This work was supported by NIH grants R01 GM058673 and R35 GM118092 to S.K. MSKCC core facilities are supported by NCI Cancer Center Support Grant P30 CA008748.

**Author contributions** H.M. performed ChIP-seq, generated translocation strains, performed the fluorescent spore assay and analysed the data. P.-C.H. constructed strains for the fluorescent spore assay with inducible *NDT80* and performed the assay. I.L. and M.v.O. performed Spo11-oligo mapping. J.S. performed ChIP-seq under the supervision of H.M. H.M. and S.K. conceived the project and wrote the paper. S.K. analysed data, procured funding and oversaw the research. H.M., S.K. and I.L. edited the manuscript.

**Competing interests** The authors declare no competing interests.

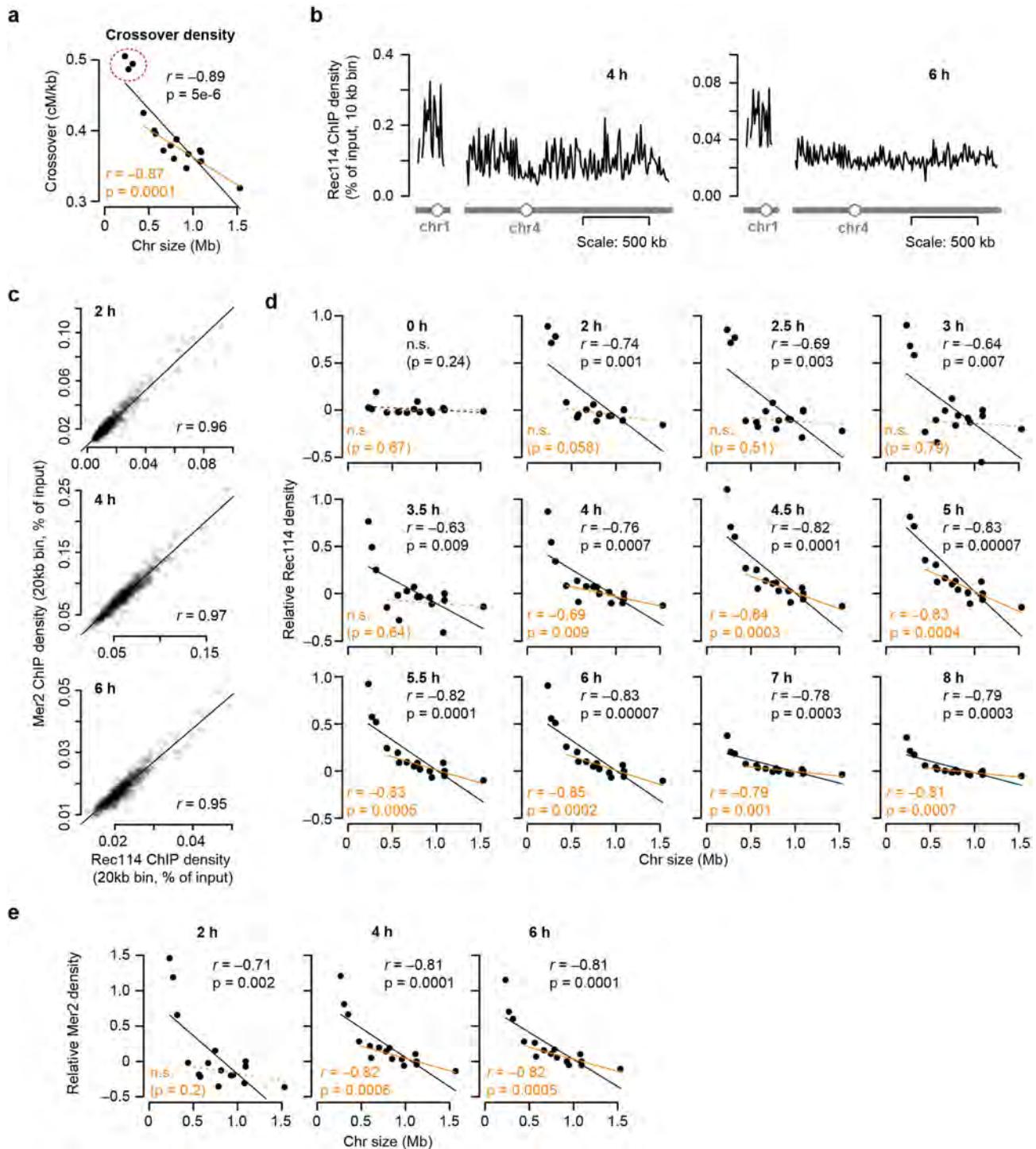
## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2248-2>.

**Correspondence and requests for materials** should be addressed to H.M. or S.K.

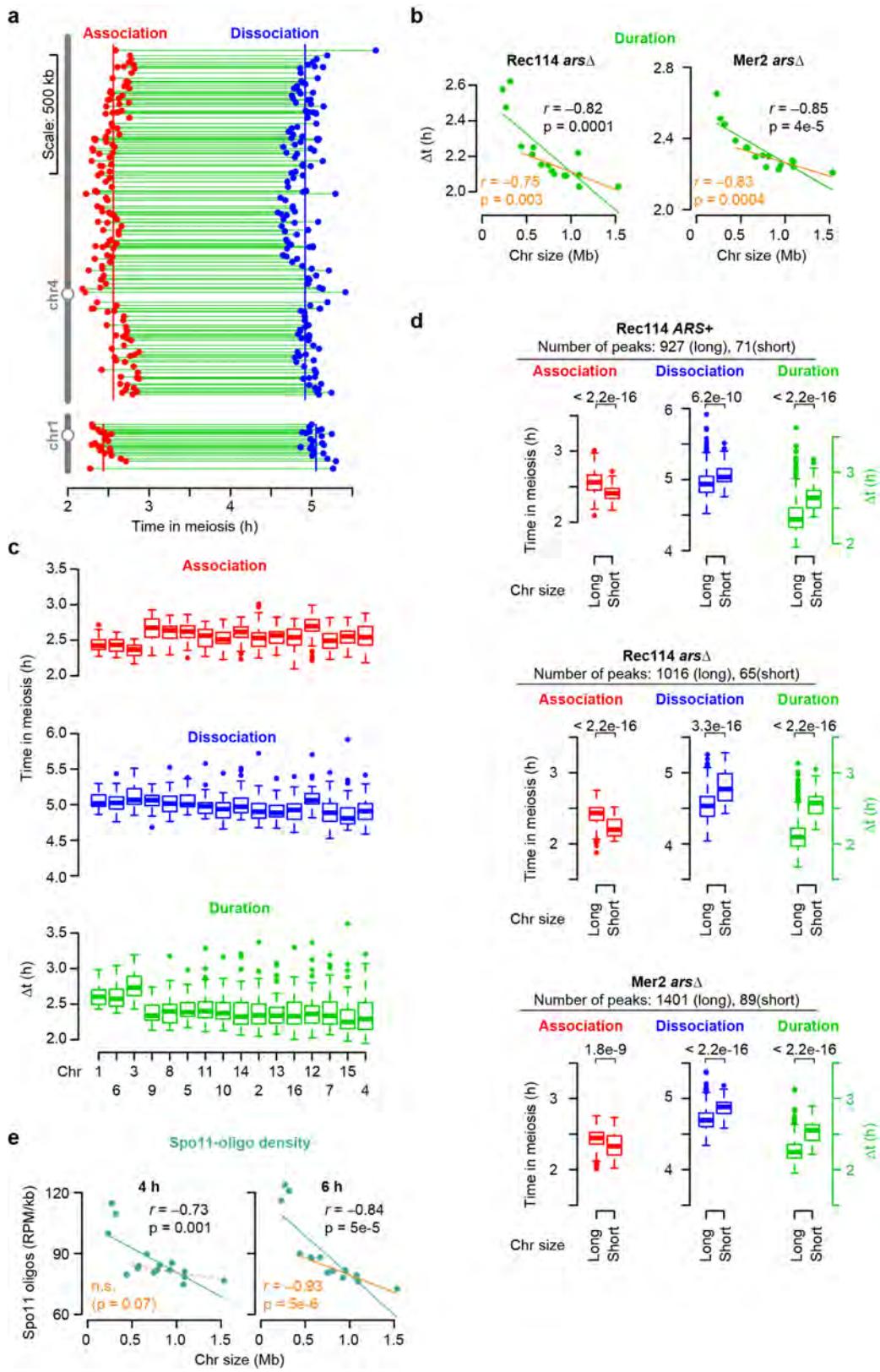
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

# Article



**Extended Data Fig. 1 | Dependence of DSB-protein binding on chromosome size.** **a**, Dependence of crossovers (centiMorgans (cM) per kb) on chromosome size. Data are from ref.<sup>12</sup>.  $n = 51$  tetrads. **b**, Example Rec114 ChIP-seq profiles for chr1 and chr4 at 4 h (left) and 6 h (right). **c**, Similarity of the ChIP-seq patterns of Rec114 and Mer2. Both strains are in an *arsΔ* background, in which all active replication origins on the left arm of chr3 are deleted. The similar profiles for

both proteins suggest similar regulation.  $n = 1$  time course for each strain. **d**, Full time course of average per-chromosome Rec114 ChIP densities (*ARS<sup>+</sup>* strain).  $n = 1$  time course. **e**, The size dependence of average per-chromosome Mer2 ChIP density changes over time, similar to that of Rec114 (compare to Fig. 1e).  $n = 1$  time course.



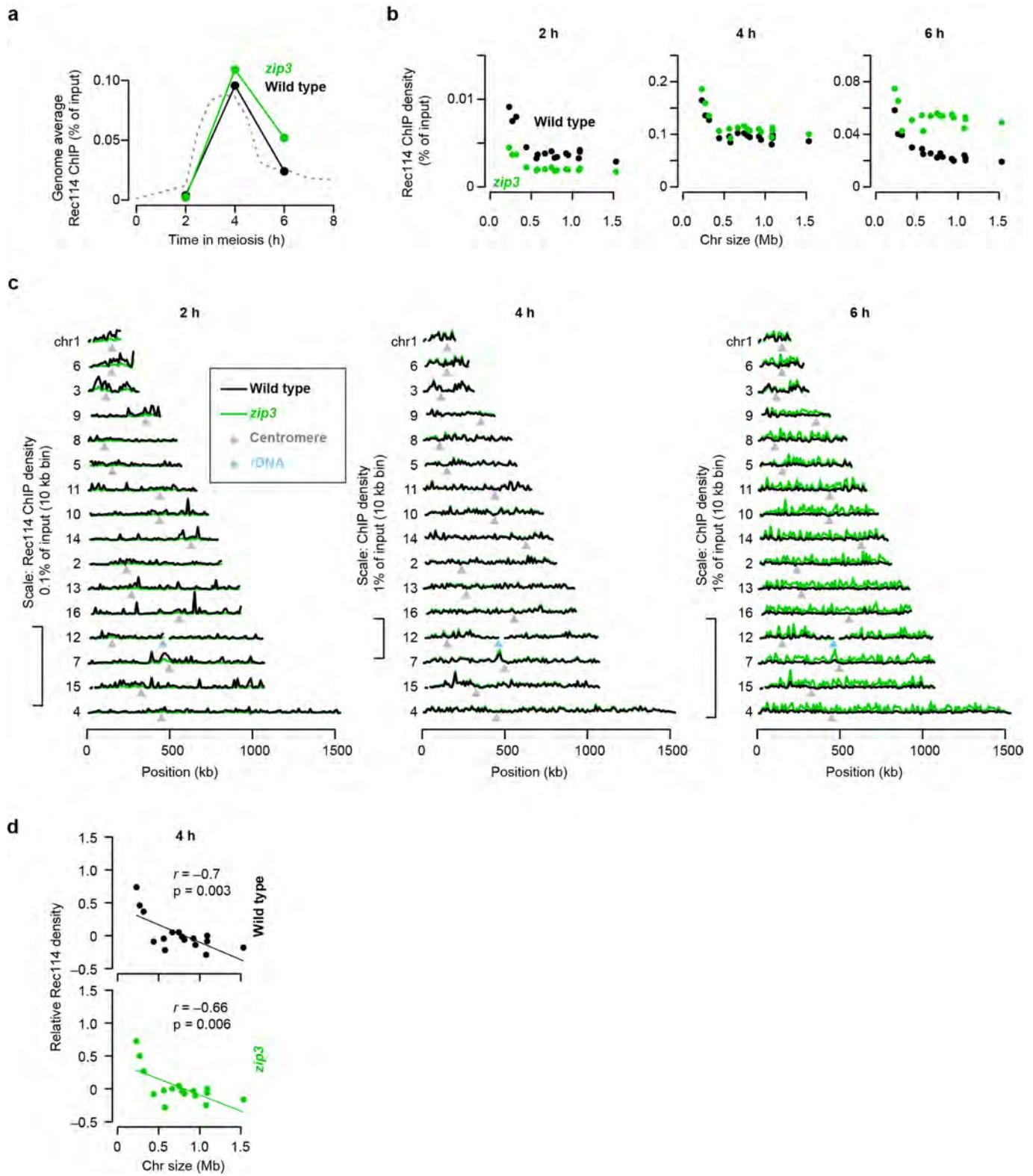
**Extended Data Fig. 2** | See next page for caption.

## Article

### Extended Data Fig. 2 | Association and dissociation times of DSB proteins.

**a**, Within-chromosome organization of Rec114 association and dissociation times (*ARS<sup>+</sup>* strain). Each point is a called Rec114 peak. Green lines indicate binding duration ( $\Delta t$ ). Vertical lines mark the per-chromosome means for association and dissociation. **b**, Per-chromosome Rec114 (left) and Mer2 (right) binding duration in *arsΔ* strains.  $n=1$  time course for each strain. **c**, Association times, dissociation times, and binding duration of Rec114 (*ARS<sup>+</sup>* strain) at Rec114 peaks, broken down by chromosome.  $n=998$  Rec114 peaks. In all box plots, thick horizontal bars denote medians, box edges mark the upper and

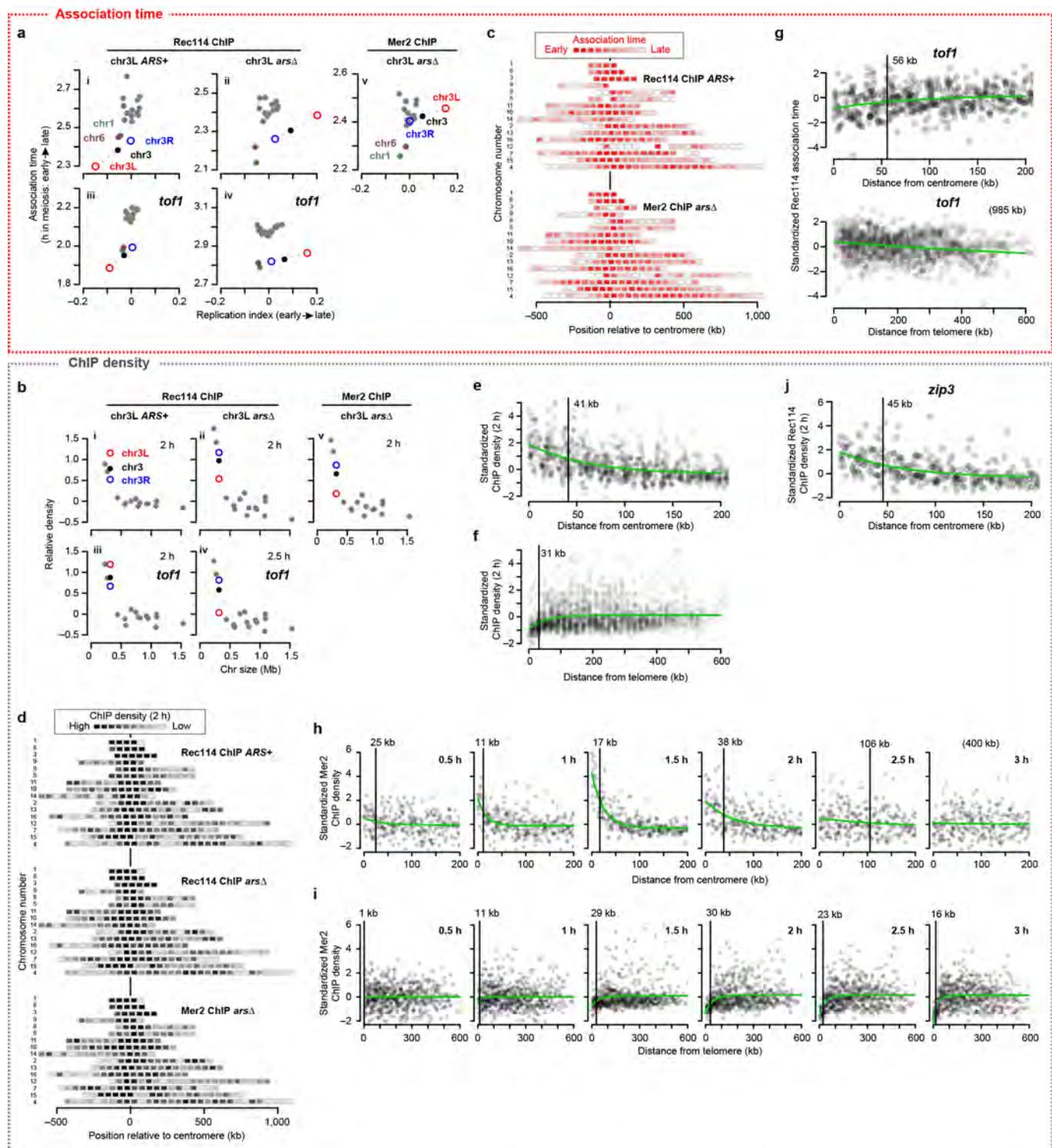
lower quartiles, whiskers indicate values within 1.5-fold of the interquartile range, and points are outliers. **d**, Rec114 and Mer2 show early association and late dissociation on short chromosomes, and thus a longer duration of binding. Values were grouped for the shortest three chromosomes and the other thirteen chromosomes ('long'). Numbers above brackets indicate  $P$  values from one-sided Wilcoxon tests. **e**, Dependence of Spo11-oligonucleotide distributions on chromosome size as a function of sampling time.  $n=2$  time courses.



**Extended Data Fig. 3 | Rec114 chromatin binding in the *zip3* mutant.** **a**, Genome-average Rec114 levels. The dashed line represents an independent wild-type profile from Fig. 1d. **b**, Effects of *zip3* mutation on absolute Rec114 ChIP densities. ChIP-seq coverage was calibrated by qPCR. Each point is the mean for a single chromosome. At 2 h, which is very early in the initial accumulation phase of Rec114 on chromatin, the *zip3* mutant has approximately twofold less Rec114 signal, affecting all chromosomes equivalently. It is possible that this decrease reflects a role for Zip3 at early timepoints, but because Zip3 is not known to have any function this early in prophase, a more likely explanation is that the *zip3* mutant culture lagged

slightly behind the wild-type culture. At 4 h, Rec114 ChIP densities were very similar between the cultures, again with little or no difference between chromosomes. However, at 6 h, the *zip3* mutant retained substantially more Rec114 signal. At least some of this increase may be attributable to delayed meiotic progression because of inhibition of Ndt80<sup>7,16</sup>, but—importantly for the purposes of this experiment—the Rec114 levels were disproportionately increased on the 13 larger chromosomes.  $n=1$  time course for each strain. **c**, Per-chromosome profiles of absolute Rec114 ChIP density. **d**, Per-chromosome Rec114 ChIP density. This is the intermediate time point for the samples shown in Fig. 1h.

# Article

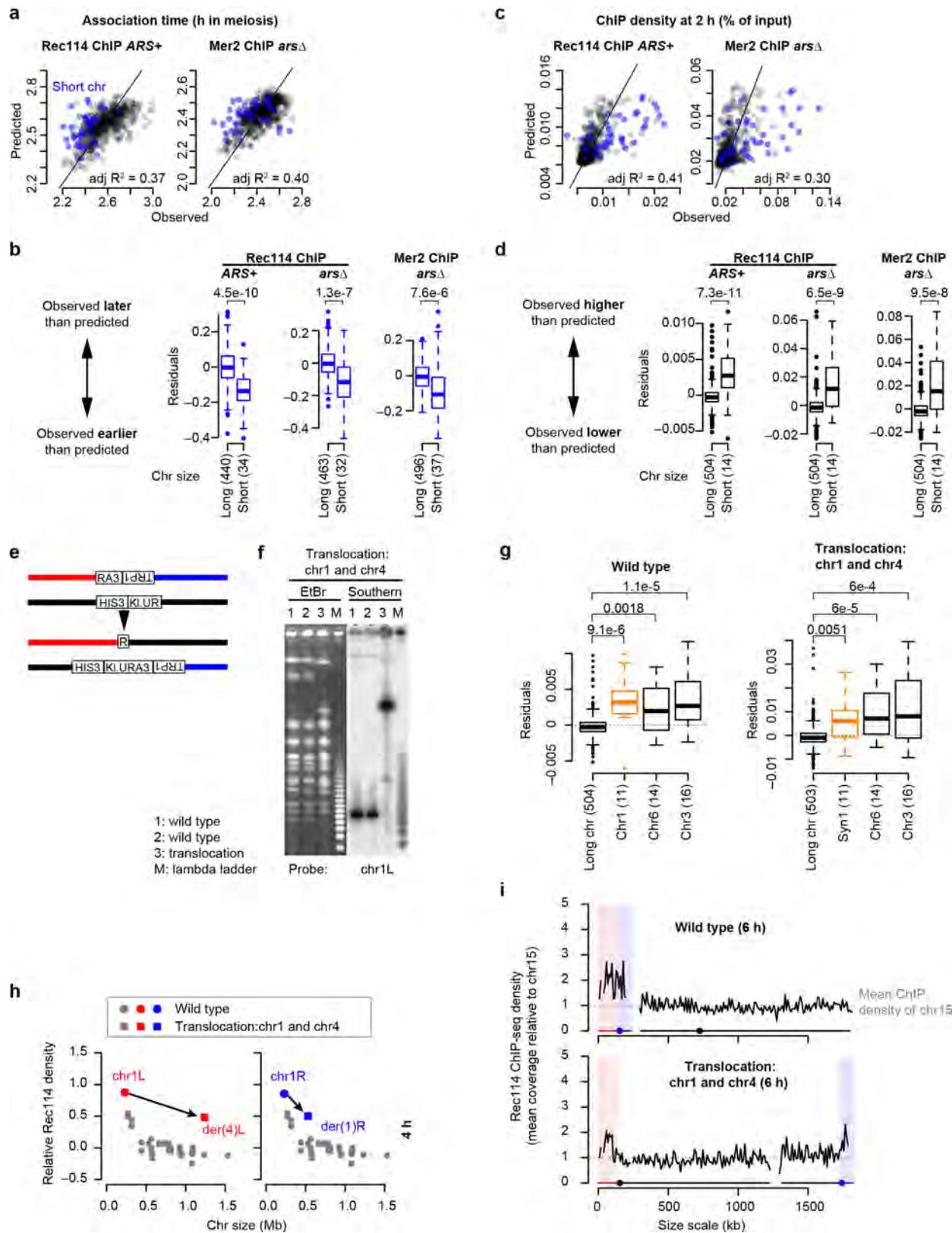


**Extended Data Fig. 4** | See next page for caption.

**Extended Data Fig. 4 | Replication timing and the centromere and telomere effects.** **a, b,** Comparison of per-chromosome Rec114 and Mer2 association time (**a**) and ChIP density at early time points (**b**, 2 and 2.5 h, normalized to chr15) with replication timing. Replication index is defined as –log<sub>2</sub>-transformed ratio of sequence coverages for ChIP input samples in S phase versus G1 phase; lower values indicate earlier replication<sup>18</sup>. The means for the left arm (chr3L, open red circle) and right arm (chr3R, open blue circle) of chr3 are plotted separately, connected to the means for the entire chromosome (solid black) by dashed lines. In wild type, chr3L showed both early replication and early Rec114 association (**a.i**). Origin deletion delayed replication and Rec114 association, but association was still on the early side compared with longer chromosomes despite extreme replication delay (**a.ii**). Mer2 was similar (**a.v**). Moreover, in *tof1*, the short three chromosomes still showed early Rec114 association (**a.iii**) and origin deletion delayed Rec114 association<sup>18</sup> less than it delayed replication (**a.iv**). ChIP densities exhibited a complementary trend: Rec114 and Mer2 overrepresentation on short chromosomes was partially dependent on early replication (**b**). Rec114 was naturally overrepresented on chr3L (**b.i**). Origin deletion caused a substantial decrease on chr3L but still left Rec114 at higher levels than on the larger chromosomes (**b.ii**) and Mer2 showed a similar pattern (**b.v**). In the *tof1* mutant, the short three chromosomes still showed high Rec114 signal (**b.iii**), and origin deletion reduced this signal on

chr3L but left it in the higher part of the longer-chromosome range (**b.iv**). The similarity with Rec114 patterns, including much later Mer2 accumulation on the late-replicating left arm of chr3, suggests that Mer2 binding to chromatin is also coordinated with replication timing. Mer2 is able to bind chromatin in the absence of Rec114<sup>3,6,40</sup>, but interaction with Rec114 (which is promoted by replication-associated Rec114 phosphorylation<sup>18</sup>) might stabilize or otherwise modify the localization of Mer2. **c, d,** Intra-chromosomal distributions of Rec114 and Mer2 association times (**c**) and ChIP density at 2 h (**d**) in the indicated strains. **c,** Average values for 50-kb bins are presented as described in the caption of Fig. 2a. **d,** Each 50-kb bin is colour-coded according to the mean ChIP density within the bin. **e, f,** Centromere (**e**) and telomere (**f**) effects on the ChIP density of DSB proteins at 2 h. The two Rec114 and one Mer2 ChIP time points were binned, standardized and pooled as in Fig. 2b, c. **g, h,** Centromere (top) and telomere (bottom) effects on Rec114 association time in *tof1* mutants. Rec114 ChIP-seq data from *ARS<sup>+</sup> tof1* and *arsΔ tof1* strains were binned, standardized and pooled as in Fig. 2b, c. The centromere and telomere effects are still apparent in the *tof1* mutants, but appear substantially weaker (particularly the telomere effect). **i, j,** Detailed time courses of Mer2 binding to chromatin near centromeres (**h**) and telomeres (**i**). **j,** The *zip3* mutation did not affect the centromere effect on Rec114 ChIP density at 2 h.

# Article

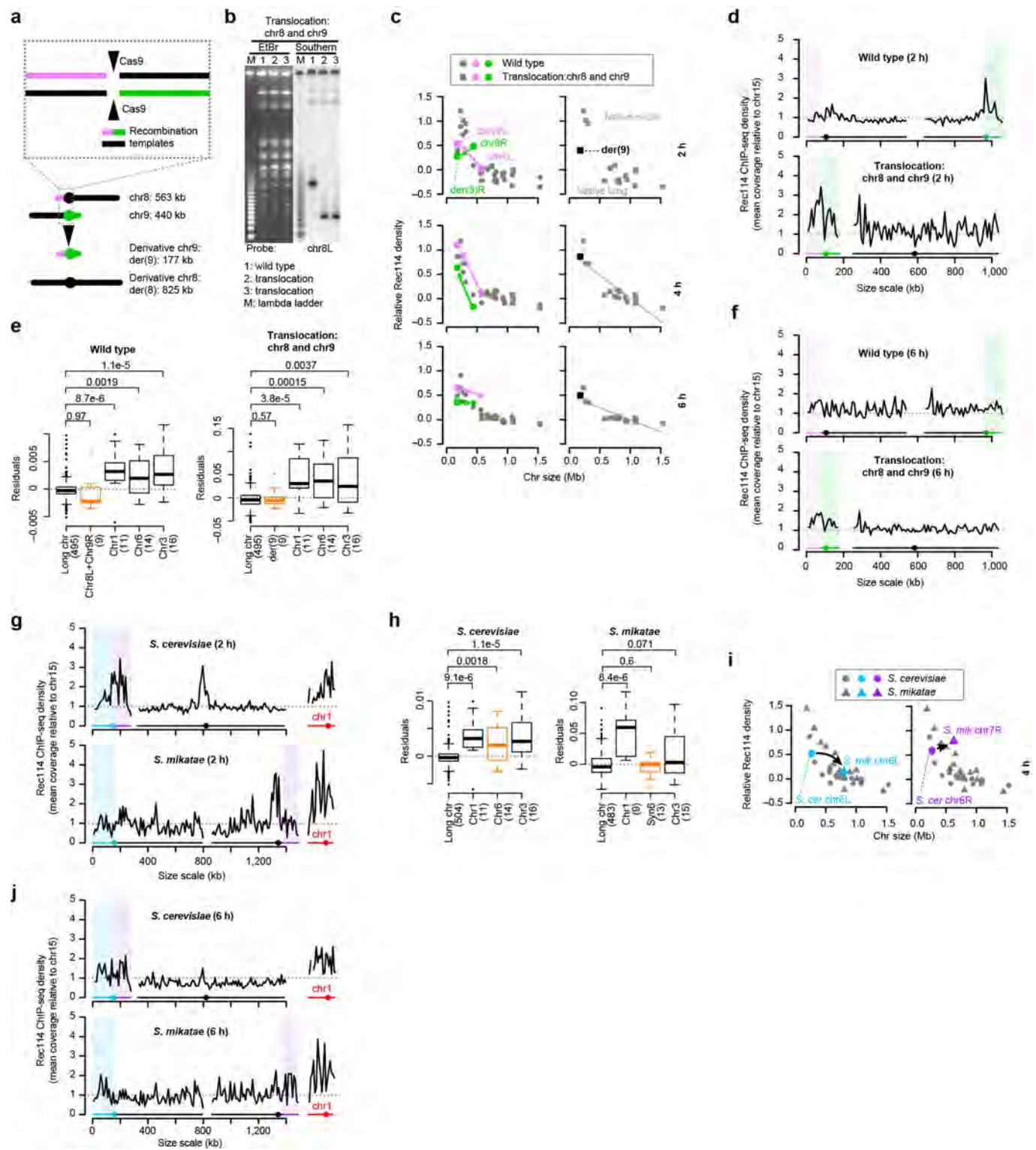


Extended Data Fig. 5 | See next page for caption.

**Extended Data Fig. 5 | DSB protein binding is boosting on smallest three chromosomes.** **a, c**, Results from multiple regression analyses using other Rec114 (*ARS*) and Mer2 (*arsΔ*) ChIP-seq datasets performed as described in Fig. 2d, f to model Rec114 ( $n=505$  bins) and Mer2 ( $n=583$  bins) association times (**a**) as well as ChIP densities at 2 h ( $n=597$  bins) (**c**). Each point is a bin; bins on short chromosomes are blue. **b, d**, Multiple regression models underperform on short chromosomes. Residuals from the three-factor models (**b**, association time; **d**, ChIP density) applied to each dataset in turn were grouped for the three shortest chromosomes compared with the others, except chr12 because of its exceptionally late association (Extended Data Fig. 2c). Numbers in parentheses indicate the number of bins.  $P$  values are from one-sided Wilcoxon tests. Box plots are as described in Extended Data Fig. 2c. **e–i**, Effects of targeted translocations on per-chromosome Rec114 binding. **e**, Strategy to target reciprocal translocation. To generate the translocation between chr1 and chr4, part of the 3' end of the *URA3* gene from *K. lactis* ('RA3') was integrated on chr1 along with the *TRP1* gene. Separately, part of the 5' end of *K. lactis URA3* ('KIUR') was integrated on chr4 along with the *HIS3* gene. The two parts of *K. lactis URA3* partially overlap, so their shared region of homology

enables reciprocal recombination between them to result in the formation of a functional *URA3* gene. **f**, Confirmation of targeted translocations. High-molecular-weight DNA was prepared from control and translocation strains and separated on pulsed-field gels and stained with ethidium bromide (EtBr). The translocations were then verified by Southern blotting using probes against the right and left ends of both chromosomes involved in the translocation ( $n=4$  different probes). A representative result using the chr1L probe is presented. For gel source data, see Supplementary Fig. 1. **g**, A three-factor multiple regression underperforms on chr1-derived sequences in both wild-type and translocation contexts. Box plots show residuals from multiple regression performed as in **d** using the 2 h data. 'Syn1' indicates sequence syntenic to chr1. Numbers in parentheses indicate number of bins.  $P$  values are from one-sided Wilcoxon tests. **h**, Chr1-derived sequences retain high-level Rec114 binding when in a large-chromosome context. Per-chromosome Rec114 ChIP densities normalized to chr15 are shown at 4 h (the intermediate time point for the samples shown in Fig. 3b). **i**, Rec114 profiles for wild-type and translocated chromosomes at 6 h. ChIP-seq data were normalized relative to chr15 and smoothed with a 10-kb sliding window.

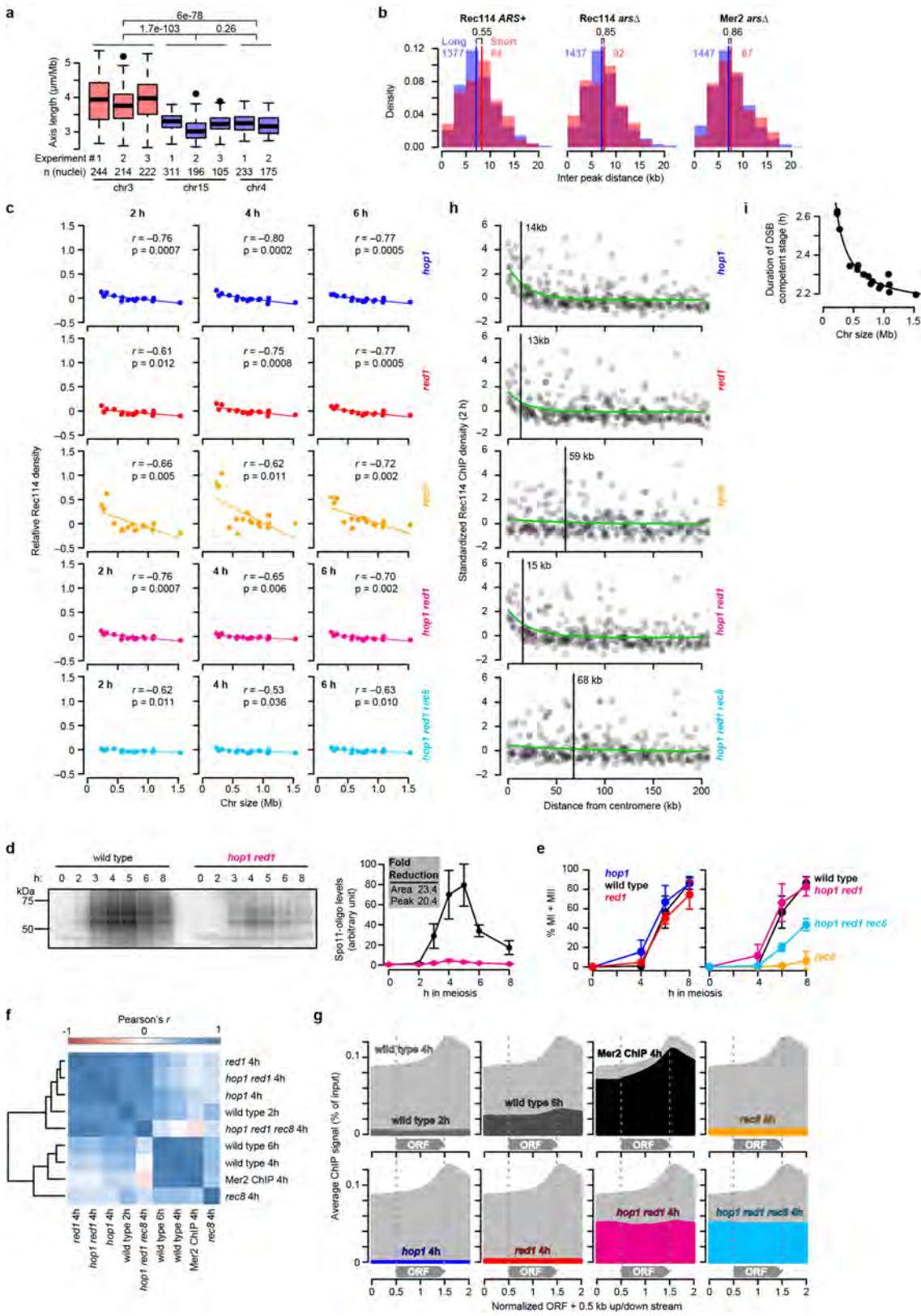
# Article



**Extended Data Fig. 6** | See next page for caption.

**Extended Data Fig. 6 | An artificially short chromosome fails to acquire a boost in Rec114 binding.** **a**, Top, strategy to target reciprocal translocation. To generate the translocation between chr8 and chr9, we introduced a plasmid expressing Cas9 and two guide RNAs that target cleavage within chr8 and chr9, respectively. This plasmid was introduced by co-transforming it along with 100-bp-long recombination templates matching the desired reciprocal recombination products. Bottom, targeted translocation between chr8 and chr9 (to scale). **b**, Confirmation of targeted translocations as described in Extended Data Fig. 5f. The translocations were verified by Southern blotting using probes against the right and left ends of both chromosomes involved in the translocation ( $n=4$  different probes). A representative result using the chr8L probe is presented. For gel source data, see Supplementary Fig. 1. **c–f**, Per-chromosome Rec114 ChIP densities (**c**), Rec114 profiles at 2 h (**d**) and 6 h (**f**), and multiple regression residuals (**e**) are shown as in Fig. 3b, c and Extended Data Fig. 5g–i. In **e**, numbers in parentheses indicate the number of bins.  $P$  values are from one-sided Wilcoxon tests. Box plots are as described in Extended Data Fig. 2c. **g–j**, High-level Rec114 binding to chr6-derived sequences is not retained in *S. mikatae*. Rec114 profiles (**g,j**), multiple regression residuals (**h**) and Rec114 ChIP density at 4 h (**i**) are shown as in Fig. 3b, c and Extended Data Fig. 5g–i. In **h**, the numbers in parentheses indicate the number of bins.  $P$  values are from one-sided Wilcoxon tests. Note that the model also tended to underperform for chr3 in *S. mikatae*, but the distribution of residuals was not statistically significant ( $P=0.071$ ).

# Article

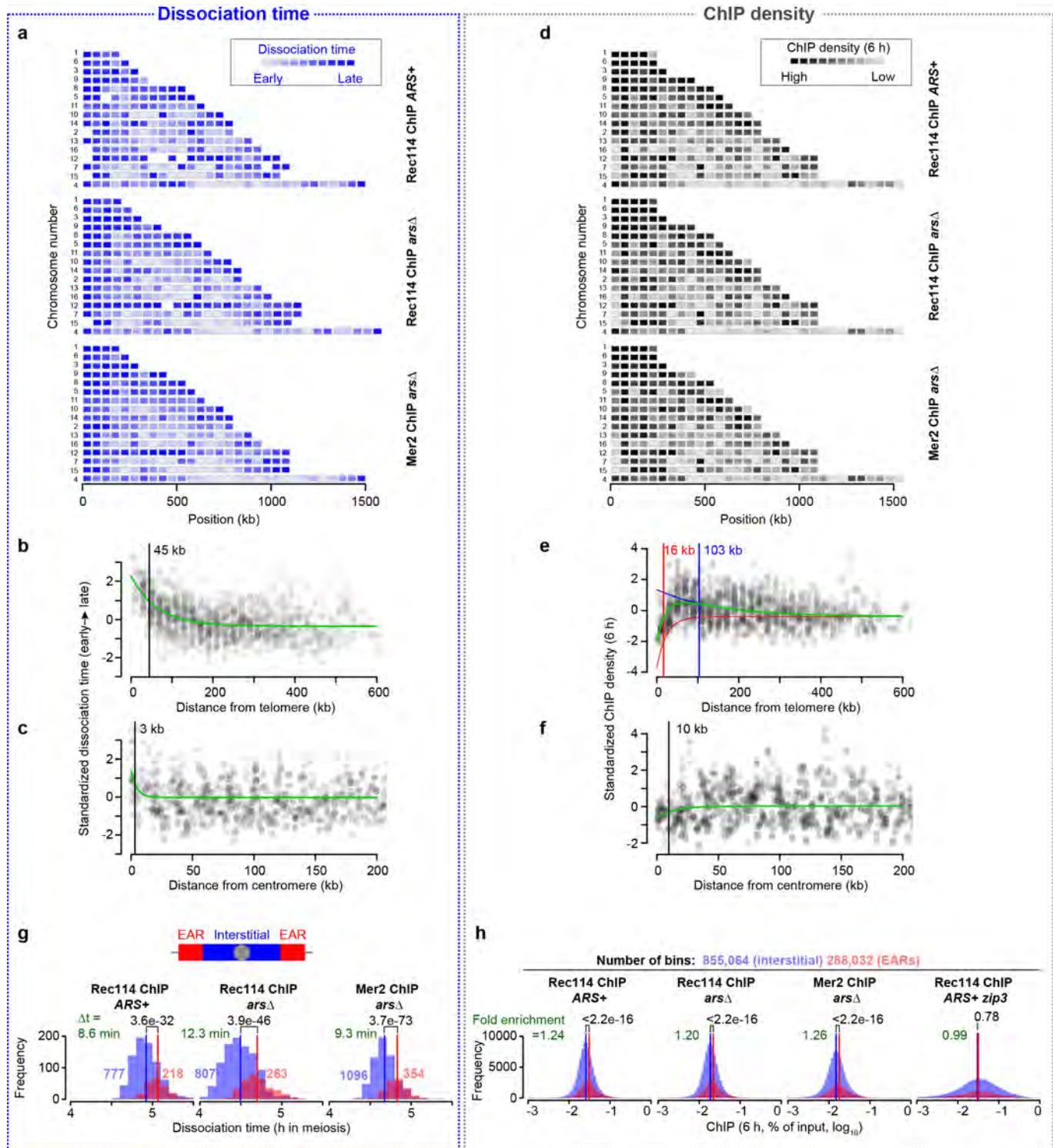


**Extended Data Fig. 7** | See next page for caption.

**Extended Data Fig. 7 | Effect of chromosome axis proteins on Rec114 chromatin binding patterns.** **a**, Long axes on chr3. To assess axis lengths, we used published measurements of synaptonemal complexes on pachytene chromosomes (see supplementary table 1 from ref.<sup>31</sup>). In that study, spread, immunostained synaptonemal complexes were traced from the positions of *lacO* arrays (bound by LacI–GFP) integrated at the right end of chr3, chr4 or chr15 to the left end of each chromosome. We therefore used the SK1 genome assembly coordinates<sup>51</sup> of the *lacO* integration sites to estimate the nucleotide length corresponding to the synaptonemal complex measurement (0.30 Mb for chr3, 1.48 Mb for chr4, and 0.99 Mb for chr15) and calculated the per-Mb axis lengths. Box plots summarize results from eight independent experiments in the wild type, including synaptonemal complex length measured by Zip1 (experiments 1 and 2) or Rec8-HA (experiment 3) staining. Note that the greater variance for chr3 is a consequence of the absolute measurement error (in μm) being a much larger fraction of the total chromosome length compared with the longer chromosomes. Data were pooled by chromosome for application of two-sided Wilcoxon tests. Box plots are as described in Extended Data Fig. 2c. **b**, Distributions of inter-peak distances. To ask whether the preferential binding of DSB proteins on short chromosomes is due to the presence of a higher density of DSB-protein-binding sites, we measured the distribution of the distances between DSB-protein ChIP-seq peaks. Vertical bars indicate medians. Coloured numerals above histograms indicate the number of distance measurements. Black numerals are *P* values from one-sided Wilcoxon tests. There was no significant difference (*P*>0.05) between short and long chromosomes in any dataset, indicating that the density of preferred binding sites does not track with chromosome size. DSB-protein-binding sites correspond to sites where Hop1, Red1 and Rec8 are also enriched—that is, presumptive sites where DNA is most likely embedded in the axis<sup>3,52</sup>. We can then question how to reconcile the similar DSB protein binding site densities between short and long chromosomes with the longer per-Mb axis length on chr3 (a). Notably, the preferred binding sites are defined on a population-average basis. Therefore, one straightforward interpretation is that smaller chromosomes have more of their potential DSB-protein-binding sites axis-associated in each cell, whereas larger chromosomes are more likely to have loops that skip over preferred axis sites. This would yield smaller loop

sizes and correspondingly longer axes on the short chromosomes despite a similar density of preferred axis sites per unit length of DNA. **c**, The dependence of Rec114 binding to chromosomes on the size of the chromosomes is lost in the absence of Hop1 or Red1, but not Rec8. Results are presented as in Fig. 4b. Note that, although correlations with chromosome size remain statistically significant in both the *hop1* and the *red1* mutants, their slopes are negligible compared to that of the wild type (Fig. 1e). *n*=1 time course for each strain. **d**, Spo11-oligonucleotide labelling to compare DSB levels between the wild type and the *hop1;red1* double mutant. Flag-tagged Spo11 was immunoprecipitated from denaturing meiotic extracts, then Spo11-oligonucleotide complexes were end-labelled with terminal deoxynucleotidyl transferase and [ $\alpha$ -<sup>32</sup>P]dCTP. Samples were separated on SDS-PAGE and imaged using a phosphorimager. Spo11-oligonucleotide complexes in yeast run as two prominent bands reflecting the different sizes of attached oligonucleotides<sup>53</sup>. Points and error bars represent mean and s.d. of three independent meiotic cultures. For gel source data, see Supplementary Fig. 1. **e**, Progression of meiosis in axis mutants. Samples from meiotic cultures were fixed and stained with DAPI, then fractions of cells completing meiosis I (MI) or both MI and meiosis II (MII) were counted. Identical wild-type data are presented in both panels to aid comparison. Points and error bars represent mean and s.d. of three independent meiotic cultures. **f**, Correlation matrix of DSB-protein ChIP datasets from wild type and axis mutants. *n*=1 time course for each strain. **g**, Distribution of DSB proteins relative to open reading frames (ORFs). Using an R package provided by the Hochwagen laboratory<sup>17</sup>, ORFs were standardized to a length of 1 kb, then ChIP-seq profiles were averaged over the standardized ORFs plus 0.5 kb of upstream and downstream sequence. The 4-h wild-type pattern (light grey shading) is repeated in each panel to facilitate comparison. **h**, The centromere effect is retained (albeit spreading less far) in *hop1* and *red1* single mutants but is lost in *rec8* mutants. Rec114 ChIP data at 2 h were binned and standardized as in Fig. 2b. The *rec8* mutation was epistatic to *hop1;red1* for loss of the centromere effect. **i**, The per-chromosome duration of DSB-protein binding has an inverse proportional relationship with chromosome size. Duration data from Fig. 1g and Extended Data Fig. 2b were combined, censoring the cold region between *CEN3* and *MAT*.

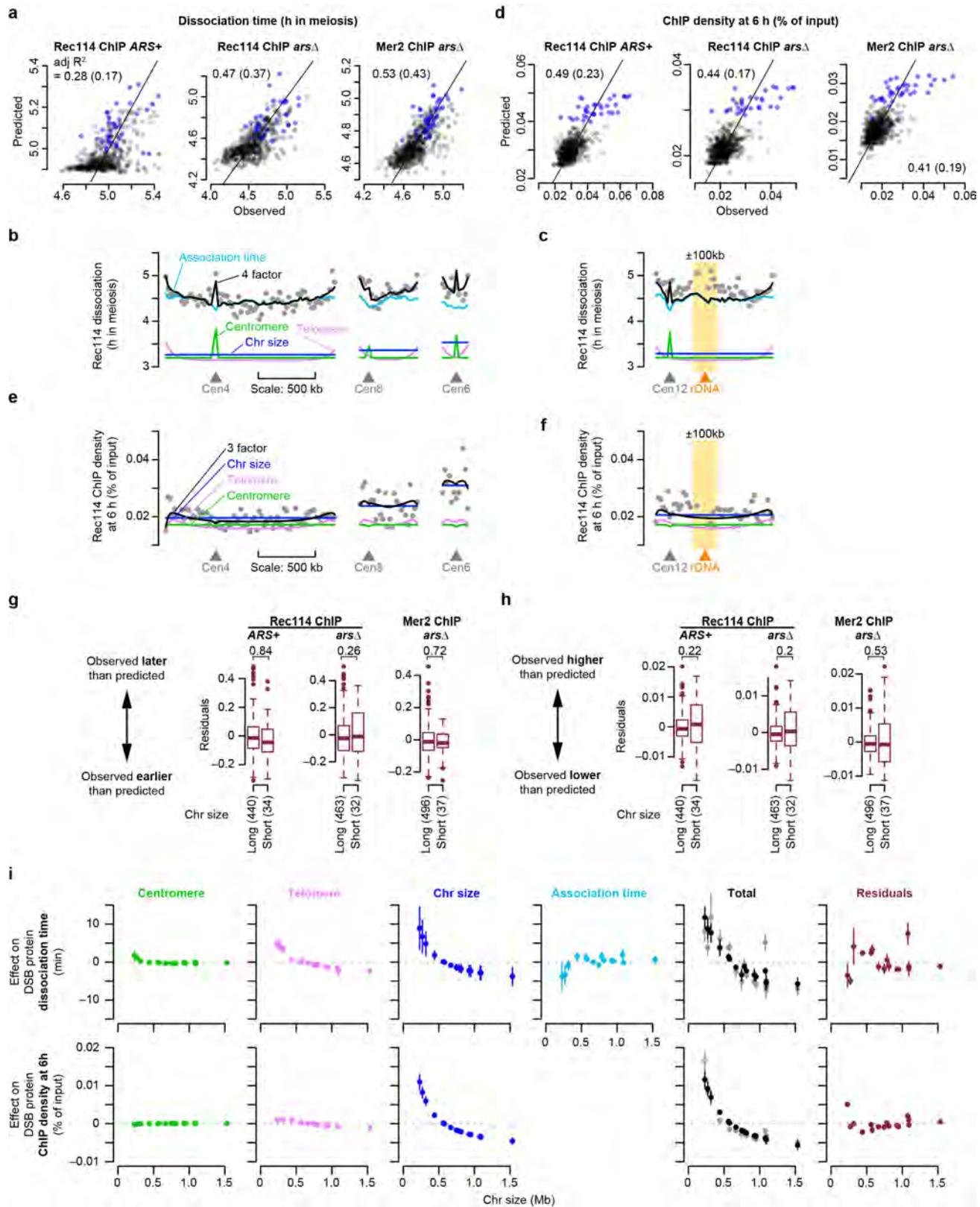
# Article



**Extended Data Fig. 8 | Proximity to telomeres influences the timing and the degree of DSB-protein dissociation from chromosomes.** **a, d,**

Intra-chromosomal distributions of Rec114 and Mer2 dissociation times (**a**) and ChIP density at 6 h (**d**) in the *ARS*<sup>+</sup> and *ars*<sup>Δ</sup> strain. Each block represents a 50-kb bin colour-coded according to the average of the Rec114 or Mer2 dissociation times for peak positions within the bin, or average ChIP density for the bin. Chromosomes are ranked by size, with the left chromosomal end at position zero. **b, c, e, f,** Effects of proximity to the centromere and the telomere on dissociation time (**b, c**) and ChIP density (**e, f**) of DSB proteins. The two Rec114 and one Mer2 ChIP time courses were combined as follows. Dissociation time and ChIP density (6 h) data from three datasets were binned (20-kb windows), standardized, and plotted as in Fig. 2b, c and Extended Data Fig. 4e, f (grey dots). Green lines are fitted exponential models. Vertical bars indicate the

distance at which the effect decays to half of the original value. For the telomere effect modelling on ChIP density at 6 h (**e**), we instead fitted a composite curve (green line) consisting of two exponential models (red and blue lines) to describe repression and enrichment, respectively. **g, h,** EAR (end-adjacent region) effects on DSB-protein dissociation time (**g**) and ChIP density at 6 h (**h**). Histograms show the distribution of dissociation time data at each peak position and ChIP density data binned in 10-bp windows located within EARs (defined as the regions from 20 to 110 kb from each telomere<sup>33</sup>) and interstitial regions (that is, all segments between EARs). ChIP density data were log-transformed to decrease skewness. Numerals in red and blue indicate the numbers of peaks located within EARs and interstitial regions, respectively (**g**). P values are from one-sided Wilcoxon tests (**g**) and one-sided t-tests (**h**).

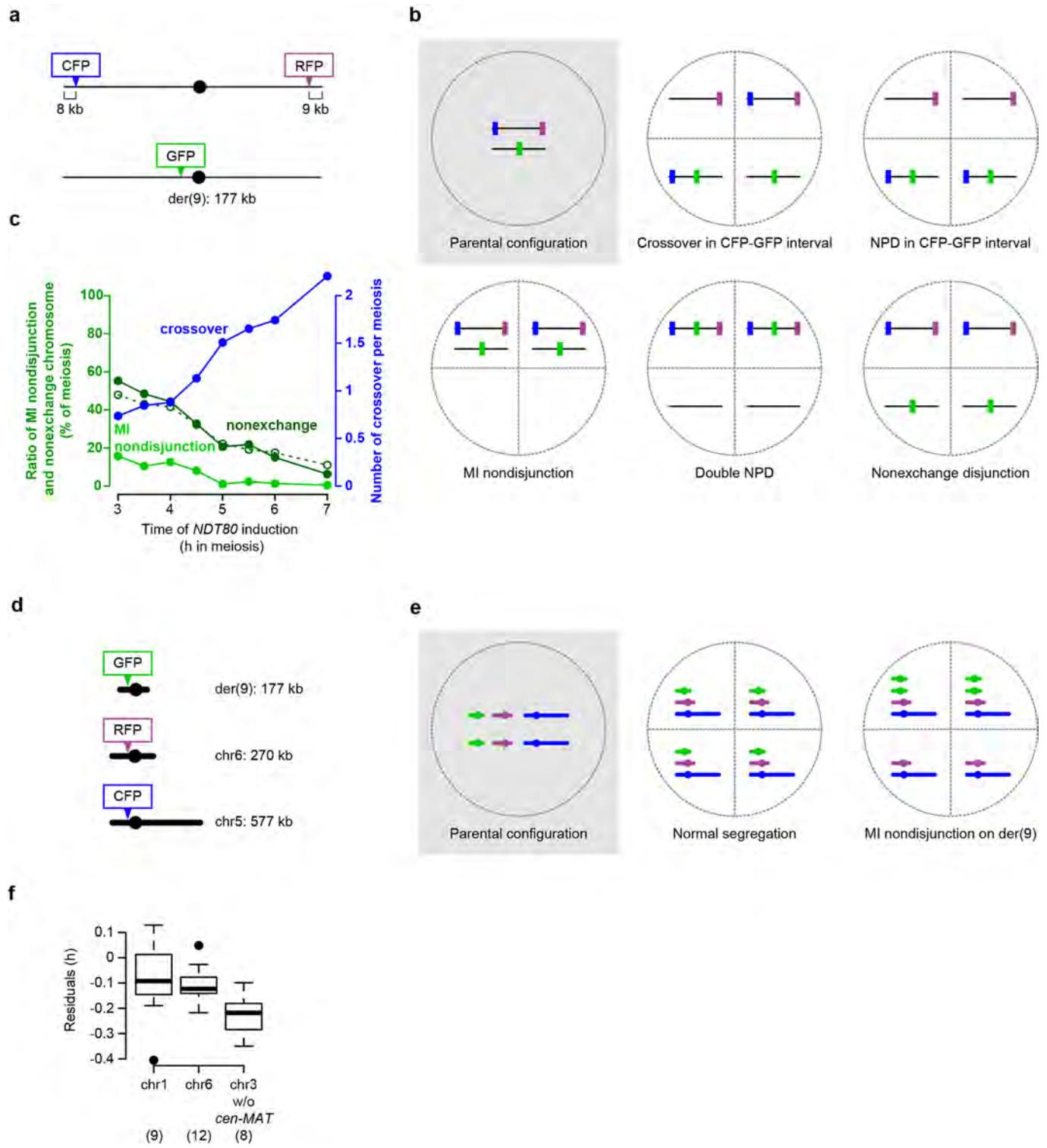


**Extended Data Fig. 9** | See next page for caption.

## Article

**Extended Data Fig. 9 | Chromosome size, separate from EAR effects, is a major determinant of DSB-protein behaviour in late prophase.** **a, d,** A model incorporating association time, chromosome size, and the centromere and telomere effects. Association time, chromosome size, centromere, and telomere effects were binned in 20-kb windows and used as explanatory variables to model DSB-protein dissociation time (**a**) or ChIP density at 6 h (**d**) by multiple linear regression. Association time was excluded in **d**. Each point compares the observed and model-predicted value within a bin. Bins on short chromosomes are blue. Adjusted  $R^2$  values are shown for all data points and (in parentheses) for only the longer chromosomes. Note that the fits for the ChIP-density models are substantially worsened by removing the small chromosomes (**d**). This suggests that these models are mainly capturing between-chromosome differences rather than within-chromosome variation on the long chromosomes.  $n = 505, 539$ , and  $583$  bins for Rec114 ChIP  $ARS^+$ , Rec114 ChIP  $ars\Delta$  and Mer2 ChIP  $ars\Delta$  datasets, respectively. **b, c, e, f,** Examples of within-chromosome patterns predicted by the multiple regression model and each of its component factors. Grey dots are observed Rec114 dissociation times (**b, c**) or Rec114 ChIP density at 6 h (**e, f**) in 20-kb bins from the  $ars\Delta$  strain. Blue, cyan, green and magenta lines are the components chromosome size, association time, centromere effects and telomere effects, respectively, and the black line is the prediction from the four-factor (**b, c**) and three-factor (**e, f**)

regression models. The 200 kb surrounding the rDNA on chr12 is shaded yellow in **c** and **f**. **g, h,** The multiple regression models do not underperform on the small chromosomes for DSB-protein dissociation (**g**) or ChIP density at 6 h (**h**). Residuals from the regression models were grouped for the three shortest chromosomes compared with the remaining chromosomes, except for chr12. Numbers in parentheses indicate the number of bins.  $P$  values are from one-sided Wilcoxon tests. Compare with the substantial underperformance on the smallest chromosomes of regression models applied to early DSB-protein patterns (Extended Data Fig. 5b, d). Box plots are as described in Extended Data Fig. 2c. **i,** Size dependence and integration of each component shaping late DSB-protein behaviour (analogous to Fig. 4e). Net effects for each chromosome (relative to the genome-wide means) were estimated from the three-factor or four-factor models in **a** and **d** (points and vertical bars are means  $\pm$  s.d. of the three datasets). Of particular importance for our purposes, note that a separate parameter capturing chromosome size per se dominates the models even when accounting for the telomere effect (which includes contribution of the EARs). This result suggests that, although EARs contribute to chromosome-size-related differences in DSB-protein dissociation<sup>33</sup>, another size-related process(es) is quantitatively more important. Homologous pairing kinetics might be such a process<sup>2,16,54</sup>. See Supplementary Discussion 6 for further details.



**Extended Data Fig. 10** | See next page for caption.

## Article

**Extended Data Fig. 10 | Small chromosomes are at risk of chromosome missegregation.** **a–c.** Premature exit from prophase I compromises crossing over and chromosome segregation on an artificial short chromosome. **a,** Configuration of spore-autonomous fluorescent markers<sup>39</sup> used to detect crossovers and MI nondisjunction of the artificial short chromosome (der(9), see Extended Data Fig. 6a–f). The strain also carries *NDT80* under the control of a galactose-inducible promoter ( $P_{GAL}\text{-}NDT80$ ) and expresses the Gal4 transcription factor fused to a portion of the mammalian oestrogen receptor (Gal4-ER)<sup>55,56</sup>. These constructs enable induction of Ndt80 expression at defined times by addition of β-oestradiol to the medium. Ndt80 expression is sufficient to drive exit from pachynema and prophase I (see Supplementary Discussion 8 for further detail). Note that this experimental set-up enables us to query a small chromosome that lacks the normal short-chromosome boost and to drive cells out of prophase before the homologue engagement pathway is allowed to operate fully to assure DSB formation. **b,** Schematics of chromosome segregation patterns with and without crossovers. NPD, nonparental ditype. Note that configuration of markers (crossover or nonexchange) cannot be determined in MI nondisjunction tetrads. Note also that double NPD tetrads are indistinguishable from MI nondisjunction; however, double NPDs are expected to be rare. **c,** Frequencies of MI nondisjunction and nonexchange (EO) chromosomes and average number of crossovers detected per meiosis. For EO, filled dark green circles are estimates of total EO by combining observed EO + observed MI nondisjunction; open circles are the predicted frequencies of EO assuming a Poisson distribution with the observed mean number of crossovers per meiosis. Note the good agreement between these independent estimates of EO frequency. β-oestradiol was added at the indicated times, then chromosome segregation and crossing over were scored after the completion of meiosis and sporulation (54 h). Pachytene exit would normally begin to occur after about 5 h in wild-type cells under these culture conditions<sup>46</sup>. Early induction of *NDT80* resulted in a decreased yield of crossovers and an increased frequency of chromosome missegregation, matching our expectation. The crossover and segregation defects were progressively attenuated the later *NDT80* was induced. Thus, delaying pachytene exit is sufficient to improve the likelihood of der(9) achieving a crossover. At least some of this improvement is probably attributed to providing more time in a DSB-permissive state, but other aspects of prophase I chromosome pairing and recombination may also be facilitated. **d, e,** Configuration of spore-autonomous fluorescent markers (**d**) and example segregation patterns (**e**) to detect MI nondisjunction for three different chromosomes. The strain also contained  $P_{GAL}\text{-}NDT80$  and Gal4-ER. Data are presented in Fig. 4f. **f,** Residuals for each member of the shortest chromosome trio (not including the *CEN3*–*MAT* interval) from the three-factor multiple regression model applied to Rec114 association time (*ARS*<sup>+</sup> dataset, Fig. 1f). The model performs poorly for all three chromosomes, but performs especially poorly for the parts of chr3 that must have a high DSB frequency to balance the suppression of DSB formation in the *CEN3*–*MAT* interval. See Supplementary Discussion 9 for further detail. Numbers in parentheses indicate the number of bins. Box plots are as described in Extended Data Fig. 2c.

Corresponding author(s): Scott Keeney and Hajime Murakami

Last updated by author(s): Feb 12, 2020

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

No software was used for data collection

Data analysis

All analyses were conducted in R or GraphPad Prism. Custom code for mapping of Spo11-oligo sequences has been published previously and is available online (as described in Methods).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All sequencing data were deposited at the Gene Expression Omnibus (GEO) with the accession numbers GSE52970 (Rec114 ChIP-seq including tof1), GSE84859 (Spo11 oligos in hop1 and red1), GSE119786 (Mer2 ChIP-seq), GSE119787 (all Rec114 ChIP-seq generated in this study) and GSE119689 (Spo11-oligo maps in wild type at 4 and 6 h).

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences     Behavioural & social sciences     Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](http://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculations were performed. Sample sizes were chosen based on established best practices in the field for experimental methods used.
Data exclusions	No samples were excluded. Upon estimating times of association and dissociation of DSB proteins by curve fitting, called peak positions with badly fitted profiles were removed using criteria described in methods. The criteria were defined in our previous study (Murakami and Keeney, Cell 2014).
Replication	All conclusions described in the paper were based on findings reproduced in biological replicate experiments.
Randomization	Not relevant: All experiments involved comparison of control (wild type) and mutant yeast strains.
Blinding	ChIP sequencing library preparation was performed by a core facility without knowledge of the identity of samples. All molecular genomics results presented in the paper are quantitative and involve internally controlled comparisons. Operator bias is not a relevant consideration for this experimental design, so it does not require blinding.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies	<input type="checkbox"/>	<input checked="" type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines	<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		

## Antibodies

Antibodies used	Anti myc antibody (9E10): BioXcell #BE0238. Anti flag antibody (M2): Sigma-Aldrich #F3165
Validation	Only monoclonal anti-epitope tag antibodies were used. These anti-myc and anti-Flag monoclonals have been in use for many decades and are standard reagents in molecular biology. Their specificity for use in yeast is well established through lack of signal in untagged control strains or, in our study, lack of ChIP signal at time zero (before meiotic entry)

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	Sources of all yeast strains are given in Supplementary Table 2.
Authentication	Yeast strains were verified by PCR and Southern blotting.
Mycoplasma contamination	Not applicable (yeast strains only).
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	Not applicable (yeast strains only).

# ChIP-seq

## Data deposition

Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).

Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

### Data access links

*May remain private before publication.*

All sequencing data were deposited at the Gene Expression Omnibus (GEO) with the accession numbers GSE52970 (Rec114 ChIP-seq including tof1), GSE84859 (Spo11 oligos in hop1 and red1), GSE119786 (Mer2 ChIP-seq), GSE119787 (all Rec114 ChIP-seq generated in this study) and GSE119689 (Spo11-oligo maps in wild type at 4 and 6 h).

### Files in database submission

All raw and processed files are submitted in GEO.

### Genome browser session (e.g. [UCSC](#))

[https://genome.ucsc.edu/cgi-bin/hgTracks?hgS\\_doOtherUser=submit&hgS\\_otherUserName=murakamh&hgS\\_otherUserSessionName=nature190105](https://genome.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=murakamh&hgS_otherUserSessionName=nature190105)

## Methodology

### Replicates

All experiments are performed once.

### Sequencing depth

*Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*

### Antibodies

Anti myc antibody (9E10)  
Anti flag antibody (M2)

### Peak calling parameters

Peak calling was performed using a custom code in R described in Methods.

### Data quality

*Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*

### Software

Base calls performed using HiSeq Control Software v.2.0.10.0 for samples sequenced on Illumina HiSeq 2500. Reads were mapped to the SacCer2 genome assembly and S. mikatae genome assembly from (Scannell et al., 2011) using BWA (0.7) MEM to generate coverage maps for each time point from each strain. Further data analyses are performed using R (3.5.1).

# Structure and mechanism of the mitochondrial Ca<sup>2+</sup> uniporter holocomplex

<https://doi.org/10.1038/s41586-020-2309-6>

Received: 28 August 2019

Accepted: 15 April 2020

Published online: 20 May 2020

Check for updates

Minrui Fan<sup>1,4</sup>, Jinru Zhang<sup>1,4</sup>, Chen-Wei Tsai<sup>2,4</sup>, Benjamin J. Orlando<sup>3</sup>, Madison Rodriguez<sup>2</sup>, Yan Xu<sup>1</sup>, Maofu Liao<sup>3</sup>, Ming-Feng Tsai<sup>2</sup> & Liang Feng<sup>1</sup>

Mitochondria take up Ca<sup>2+</sup> through the mitochondrial calcium uniporter complex to regulate energy production, cytosolic Ca<sup>2+</sup> signalling and cell death<sup>1,2</sup>. In mammals, the uniporter complex (uniplex) contains four core components: the pore-forming MCU protein, the gatekeepers MICU1 and MICU2, and an auxiliary subunit, EMRE, essential for Ca<sup>2+</sup> transport<sup>3–8</sup>. To prevent detrimental Ca<sup>2+</sup> overload, the activity of MCU must be tightly regulated by MICUs, which sense changes in cytosolic Ca<sup>2+</sup> concentrations to switch MCU on and off<sup>9,10</sup>. Here we report cryo-electron microscopic structures of the human mitochondrial calcium uniporter holocomplex in inhibited and Ca<sup>2+</sup>-activated states. These structures define the architecture of this multicomponent Ca<sup>2+</sup>-uptake machinery and reveal the gating mechanism by which MICUs control uniporter activity. Our work provides a framework for understanding regulated Ca<sup>2+</sup> uptake in mitochondria, and could suggest ways of modulating uniporter activity to treat diseases related to mitochondrial Ca<sup>2+</sup> overload.

The uniporter is quiescent in resting cellular conditions and becomes activated only when local Ca<sup>2+</sup> levels rise above approximately 1 μM (refs. <sup>11,12</sup>). This Ca<sup>2+</sup>-dependent activation, mediated by a MICU1–MICU2 heterodimer<sup>13,14</sup>, prevents excessive Ca<sup>2+</sup> influx that can increase mitochondrial oxidative stress<sup>15</sup> and trigger apoptosis<sup>1</sup>. Mutations that perturb MICU1's regulatory function have been linked to debilitating neuromuscular diseases in humans<sup>16,17</sup>. Recent breakthroughs have advanced our understanding of the composition, function and regulation of the uniplex<sup>9,10,18–32</sup>, but the structural basis underlying MICU-mediated control of the uniporter remains unclear. Here we determine the structures of the human uniplex in low- and high-Ca<sup>2+</sup> conditions using single-particle cryo-electron microscopy (cryo-EM). Combined with functional analyses, this work reveals the molecular mechanisms by which intracellular Ca<sup>2+</sup> signals control mitochondrial Ca<sup>2+</sup> uptake.

## Structural determination

The purified human uniplex showed excellent biochemical behaviour, with MICU1 and MICU2 being connected by a disulfide bridge (Extended Data Fig. 1a, b), recapitulating their reported properties<sup>13,14</sup>. Cryo-EM revealed two main particle species in low Ca<sup>2+</sup> (5 mM EGTA, no added Ca<sup>2+</sup>), namely a uniplex monomer and a V-shaped dimer (Extended Data Fig. 2), whose equivalent parts superimpose well onto each other. The monomer yielded a substantially higher-resolution map (3.3 Å) (Extended Data Fig. 3a–f), therefore becoming the focus of subsequent analyses. Consistent with previous studies<sup>24,27</sup>, the amino (N)-terminal domain (NTD) of MCU is well resolved in the uniplex dimer but not the monomer. In high Ca<sup>2+</sup> (2 mM), particles are predominantly dimers (Extended Data Fig. 4), with ~57% containing all four subunits and ~43%

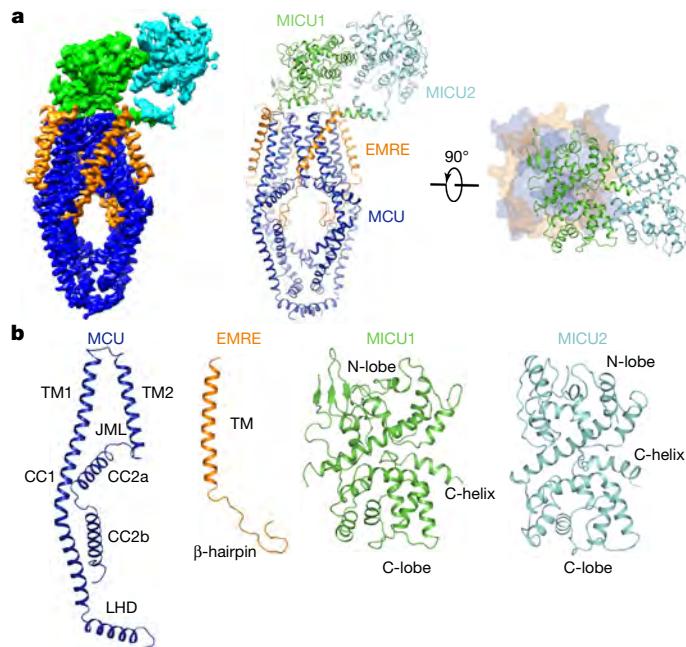
without visible density of MICU1–MICU2, which might dissociate during protein purification or grid preparation. The best classes containing all subunits yielded a 3.6 Å map (Extended Data Fig. 3g–j), corresponding to the uniplex dimer with well resolved NTDs. Attempts to reconstitute purified uniplex into liposomes were hindered by difficulties in removing the lauryl maltose neopentyl glycol (LMNG) detergent. We therefore used well established, HEK cell based mitochondrial Ca<sup>2+</sup>-uptake assays<sup>8,18</sup> to validate the structures.

## Overall structure and pore in low Ca<sup>2+</sup>

The structure determined in low-Ca<sup>2+</sup> conditions reveals a 4:4:1:1 stoichiometry of MCU, EMRE, MICU1 and MICU2 (Fig. 1). MCU tetramerizes to form a Ca<sup>2+</sup>-conducting pore with EMREs attached to its periphery around a central approximate fourfold symmetry axis. MICU1 forms an extensive interacting surface with MCU to seal the pore's intermembrane space (IMS) entrance, while MICU2 binds to MICU1 from the side without contacting MCU (Fig. 1a). Thus, the uniplex has the overall shape of an inverted L.

The uniplex's transmembrane (TM) portion superimposes well onto the published MCU–EMRE subcomplex structure<sup>27</sup> (Extended Data Fig. 5a). EMRE makes three contact sites with MCU's TM1, TM2 and coiled-coil 2 (CC2) domains (Extended Data Fig. 5b). The IMS entrance of the MCU pore is formed by four D261 residues. One helical turn downward is a glutamate ring that mediates high-affinity Ca<sup>2+</sup> binding (Extended Data Fig. 5c). An extra density, possibly a cation, lies at the centre of this ring (Extended Data Fig. 3b). The matrix end of the pore—presumably gated by EMRE via MCU's juxtamembrane loop (JML)<sup>27</sup>—is wide open (Extended Data Fig. 5d). Densities consistent with lipids lie at the fenestration<sup>24</sup> between TM1 and TM2 without

<sup>1</sup>Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, CA, USA. <sup>2</sup>Department of Physiology and Biophysics, University of Colorado Anschutz Medical Campus, Aurora, CO, USA. <sup>3</sup>Department of Cell Biology, Harvard Medical School, Boston, MA, USA. <sup>4</sup>These authors contributed equally: Minrui Fan, Jinru Zhang, Chen-Wei Tsai.  
✉e-mail: ming-feng.tsai@cuanschutz.edu; liangf@stanford.edu



**Fig. 1 | Overall structure of the human uniplex in low-Ca<sup>2+</sup> conditions.**

**a**, Architecture of the uniplex. The cryo-EM map (left) and ribbon representation (middle) are viewed from the membrane. On the right, the uniplex (with MCU–EMRE shown as a surface representation and MICU1–MICU2 as a ribbon) is viewed from the top. **b**, Domain organization of MCU, EMRE, MICU1 and MICU2. LHD, linker helix domain.

blocking the pore (Extended Data Fig. 3f). Notably, the overall pore structure in the full uniplex is virtually identical to that in the MICU-free, MCU–EMRE subcomplex<sup>27</sup> (Extended Data Fig. 5a, d), indicating that MICU-dependent gating does not require substantial conformational changes inside the pore.

### MCU–MICU1 interface in low Ca<sup>2+</sup>

One striking feature of the uniplex is the 4:1 MCU:MICU1 stoichiometry. This symmetry mismatch allows an MCU tetramer to devote its entire IMS surface to bind a single MICU1 (Fig. 2a). The relatively flat bottom of MICU1's N-lobe docks directly onto the flat IMS surface of MCU such that a lysine/arginine ring (K126, R129, R259, R261 and R263) of MCU1 forms a cap to seal the D261 ring at the MCU pore entrance (Fig. 2b). All D261 residues are engaged in electrostatic interactions with these K/R-ring residues (Extended Data Fig. 5e), explaining previous findings that MICU1 binding to MCU is sensitive to ionic strength and can be abolished by the D261A mutation<sup>22</sup>. Several additional contacts appear to stabilize MCU–MICU1 interactions (Fig. 2b): first, Y114 in MICU1 forms a hydrogen bond with D261 in MCU; second, R129 and R263 in MICU1 contact MCU's S259, a residue important for MICU1 binding<sup>22</sup>; third, Y121 in MICU1 forms hydrophobic π-stacking interactions with Y258 and I262 in MCU; and fourth, R261 of MICU1 interacts with the helix dipole of TM1.

We carried out coimmunoprecipitation to validate this MCU–MICU1 interface. Although K126E or R129E mutations alone in MICU1 are insufficient to perturb MCU binding, the K126E–R129E double mutation does strongly destabilize the MCU–MICU1 complex (Extended Data Fig. 1c). The K126A–R129A mutant still binds MCU (Extended Data Fig. 1c), consistent with the presumably less disruptive effects of alanine substitution on electrostatic interactions with MCU's D261. Adding Y114A and Y121A mutations to the K126A–R129A mutant abolishes interactions with MCU (Extended Data Fig. 1c), reflecting their contributions to the MCU–MICU1 interface.

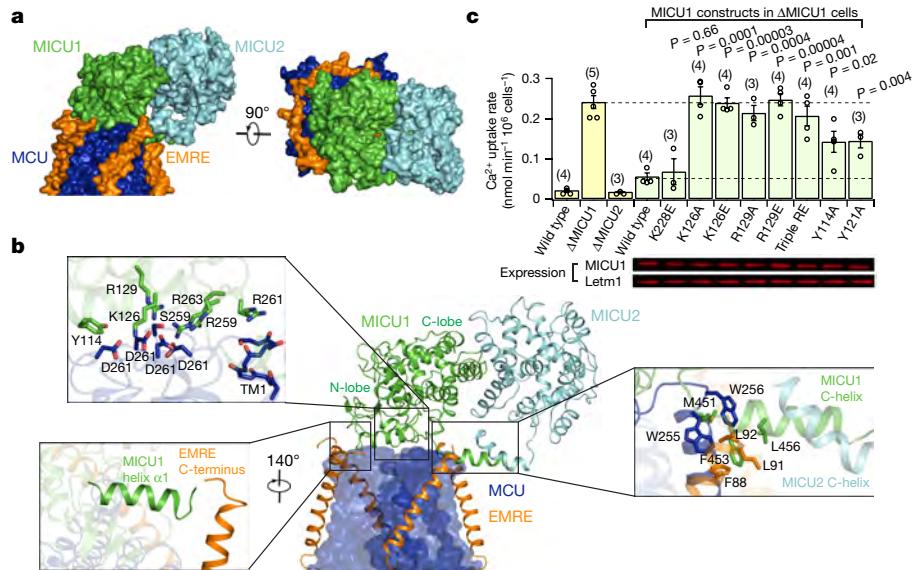
In addition to the major MCU–MICU1 interface, the N-terminal end of MICU1's carboxy (C)-terminal helix interacts directly with MCU and EMRE through hydrophobic interactions (Fig. 2b). As this amphiphilic helix parallels the membrane surface, it might also facilitate interactions with MCU via membrane anchoring. Indeed, C-helix truncation substantially weakens MICU1's interaction with MCU but not MICU2 (Extended Data Fig. 1d). Finally, the N terminus of MICU1's α1 helix contacts the C terminus of another EMRE (Fig. 2b). Thus, their immediate neighbouring KKRR ('polybasic sequence') of MICU1 and EMRE's C-terminal polyaspartate tail, although unresolved, are within range for direct interaction, as proposed previously<sup>18</sup>. This MICU1–EMRE interaction has important roles in uniplex stability, as truncating EMRE's C terminus or mutating MICU1's polybasic sequence facilitates MICU1 dissociation from the MCU–EMRE subcomplex (Extended Data Fig. 1e). Together, these results unveil multiple molecular contacts that facilitate and stabilize the docking of MICU1 on MCU in low-Ca<sup>2+</sup>, resting cellular conditions.

### Molecular mechanisms of Ca<sup>2+</sup> activation

The uniplex structure in low Ca<sup>2+</sup> leads to the suggestion that MICU1 uses its K/R ring to seal the MCU pore. To further understand how Ca<sup>2+</sup> activates the channel, we determined the cryo-EM structure of the uniplex in 2 mM Ca<sup>2+</sup>. The O-shaped complex has identical subunit stoichiometry ratio to the uniplex in low Ca<sup>2+</sup>, containing a V-shaped dimer of MCU–EMRE tetramers with two MICU1–MICU2 heterodimers arching over the top (Fig. 3 and Extended Data Fig. 6a). The V-shaped dimer is similar to that observed in the MCU–EMRE subcomplex structure<sup>27</sup>, except that the two tetramers rotate slightly against MCU's NTD, so that the top of the tetramers moves roughly 7 Å closer (Extended Data Fig. 6b–d). The relative flexibility between the NTD and the remaining parts of MCU is likely to facilitate this rotation. The MCU pore assumes a conformation that is essentially identical to the low-Ca<sup>2+</sup> state, with a putative Ca<sup>2+</sup> density in the glutamate ring and the presumed JML luminal gate fully open.

The most striking feature of the uniplex in high Ca<sup>2+</sup> is that MICU1 no longer covers the pore; instead, it moves to the edge of the MCU–EMRE tetramer, losing most of its interactions with MCU (Fig. 3). This suggests that Ca<sup>2+</sup> activates the uniporter by removing MICU1 from the MCU surface to open the pore. Such a mechanism predicts that mutations in MICU1's K/R ring might disrupt the tight sealing to allow Ca<sup>2+</sup> to leak into the pore, thus reducing MICU1's ability to shut MCU in low-Ca<sup>2+</sup> conditions. We tested this hypothesis using a quantitative <sup>45</sup>Ca<sup>2+</sup> flux assay. Knocking out MICU1, but not MICU2, induces massive Ca<sup>2+</sup> leakage into mitochondria (Fig. 2c), consistent with recent studies<sup>19,20</sup> and with the structural observation that MICU1 but not MICU2 seals the MCU pore. Wild-type MICU1 or a control mutant (K228E, distal from the interface) strongly inhibits the uniporter (Fig. 2c). In contrast, mutations targeting the K/R ring—K126A/E, R129A/E, and R259E–R261E–R263E—eliminate MICU1's inhibitory function (Fig. 2c). Alanine substitutions of Y114 or Y121 compromise MICU1 regulation to a lesser degree, reflecting the indirect roles of these residues in shutting MCU by facilitating MICU1 binding and positioning the K/R ring. These results confirm the critical role of MCU–MICU1 interfacial residues in Ca<sup>2+</sup>-dependent, MICU1-mediated gating. Finally, truncating MICU1's C-helix reduces MICU1 functionality (Extended Data Fig. 1d), consistent with this C-helix stabilizing MCU–MICU1 interactions.

In the high-Ca<sup>2+</sup> structure, MICU1 loses interactions with MCU's IMS surface, relying on limited contacts with EMRE to stay within the uniplex. Specifically, EMRE's C terminus contacts the N-terminal end of MICU1's helix α1, so that EMRE's polyaspartate tail and MICU1's polybasic region before helix α1, although unresolved, can directly interact. The structure also implicates a tangential contact between the C terminus of another EMRE and the N-terminal end of MICU1's helix α10, although limited local resolution precludes resolving specific



**Fig. 2 | Interface between MICU1–MICU2 and MCU–EMRE in low- $\text{Ca}^{2+}$  conditions.** **a**, Surface representation of the uniplex. The CC and LHD domains of MCU and the N-terminal part of EMRE are omitted for clarity. **b**, Uniplex subunit interfaces. MCU is shown as a surface representation (blue). EMRE, MICU1 and MICU2 are shown as ribbons. Boxed panels show zoomed-in views of contact areas. **c**, Effects of MCU–MICU1 interfacial mutations on mitochondrial  $\text{Ca}^{2+}$  uptake in low  $\text{Ca}^{2+}$  (300 nM). Yellow bars compare wild-type, MICU1-knockout ( $\Delta\text{MICU1}$ ) and MICU2-knockout ( $\Delta\text{MICU2}$ ) cells. Green bars

compare various MICU1 constructs expressed in MICU1-knockout cells. MICU1 expression was adjusted to similar levels, as shown in the western images below (Lettm1, loading control; three independent experiments were performed with similar results). Data are presented as means  $\pm$  s.e.m. Numbers in parentheses represent numbers of biologically independent experiments. Triple RE: R259E–R261E–R263E. Two-tailed *t*-test was used to compare the ability of wild-type and mutant MICU1 to gate MCU (*P*-values provided on the bar chart). For gel source data, see Supplementary Fig. 1.

side-chain interactions. Coimmunoprecipitation showed that raising  $\text{Ca}^{2+}$  levels strongly facilitates MICU1 dissociation from the uniplex, but a considerable portion of MICU1 stays bound (Extended Data Fig. 7a), consistent with a much-reduced interacting surface of MICU1 with MCU–EMRE upon  $\text{Ca}^{2+}$  elevation. C-terminal truncation of EMRE or charge reversal of MICU1’s polybasic sequence induces full MICU1 dissociation (Extended Data Fig. 7a, b), corroborating a previous report<sup>18</sup> and the structural implication that EMRE tethers MICU1 in the uniplex during  $\text{Ca}^{2+}$  stimulation.

The O-shaped structure shows a uniplex dimer connected by MCU’s NTD, and two MICU2s making a ‘back-to-back’ contact. Interestingly, size-exclusion chromatography shows that a D123R mutation that disrupts NTD interactions<sup>27</sup> fully monomerizes the uniplex, but that the uniplex dimer remains intact with a K121A mutation known to break MICU2 back-to-back dimers<sup>29</sup> or with R107E–R120E–K121E–D154R quadruple mutations (Extended Data Fig. 7c, d). Thus, it appears that primarily MCU’s NTD is responsible for uniplex dimerization. Consistent with this, the NTD interface is sufficient for MCU–EMRE subcomplex dimerization<sup>27</sup>. The D123R MCU mutant mediates as robust  $\text{Ca}^{2+}$  transport as wild-type MCU in high  $\text{Ca}^{2+}$  (Extended Data Fig. 7e, f), and is properly regulated by MICU1 in low  $\text{Ca}^{2+}$  (Extended Data Fig. 7g). Similar observations were obtained with K121A or R107E–R120E–K121E–D154R

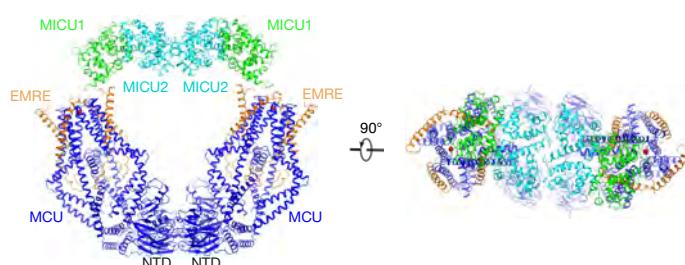
MICU2 mutants (Extended Data Fig. 7e–g). Thus, uniplex dimerization is not necessary for its basic channel function.

As V-shaped dimerization of the  $\text{F}_0\text{F}_1$ -ATPase drives this complex to the curved ridges of the inner mitochondrial membrane<sup>33</sup>, we investigated whether dimerization also affects uniplex localization. Corroborating a previous report<sup>34</sup>, MCU is more enriched in inner- and outer-membrane contact points compared with a crista-membrane protein, COX2 (Extended Data Fig. 7h, i). The dimer-breaking D123R mutation induces a COX2-like distribution of MCU, while neither MICU2-knockout nor MICU2 dimer-interface mutation alters MCU distribution (Extended Data Fig. 7j–l). The results thus suggest that dimerization might be involved in contact-point enrichment of the uniplex.

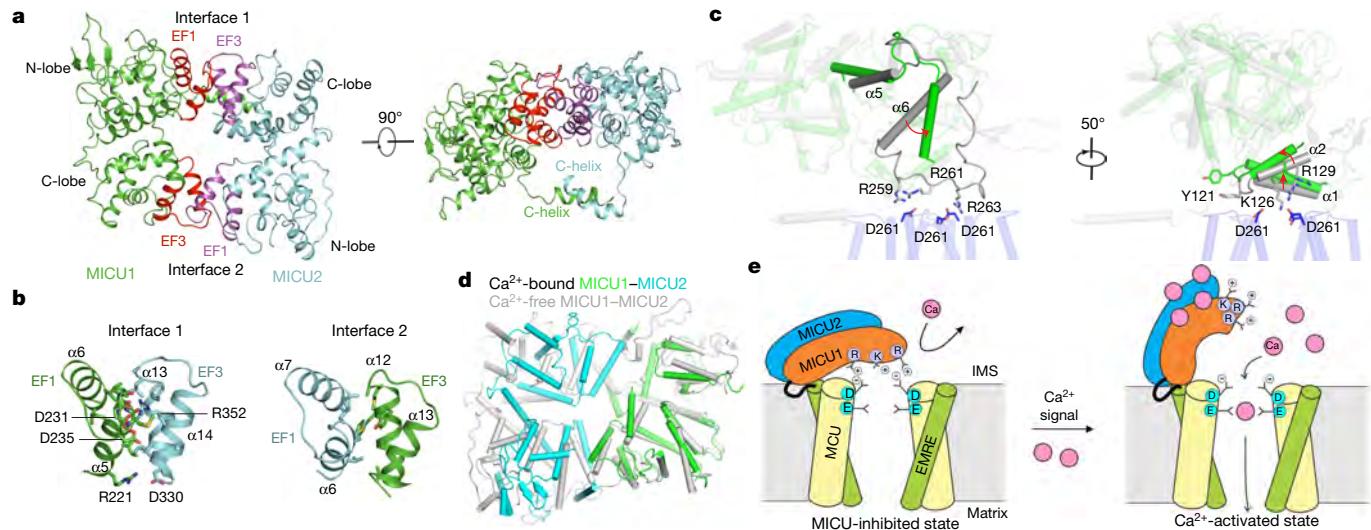
### $\text{Ca}^{2+}$ -induced MICU conformational changes

To understand how  $\text{Ca}^{2+}$  induces conformational changes in MICU to unblock the MCU pore, we examined the structures of the MICU1–MICU2 heterodimer within the uniplex under low- and high- $\text{Ca}^{2+}$  conditions. Overall, MICU1 and MICU2 share similar architectures<sup>28–31</sup>, comprising N-lobes with canonical EF1 and pseudo EF2 EF hands, C-lobes with canonical EF4 and pseudo EF3 EF hands, and C-terminal helical tails (Fig. 1b).

In low  $\text{Ca}^{2+}$ , the conformations of the canonical EF hands (EF1 and EF4) of MICU1 and MICU2 are compatible with no  $\text{Ca}^{2+}$  binding (Extended Data Fig. 8a, b). Moreover, both MICU1 and MICU2 match well with  $\text{Ca}^{2+}$ -free but not  $\text{Ca}^{2+}$ -bound MICUs<sup>28–31</sup>. MICU1 and MICU2 form a ‘face-to-face’ heterodimer whose interface includes two parts (Fig. 4a, b). The first part is between MICU1’s N-lobe EF1 and MICU2’s C-lobe EF3. In particular, R352 of MICU2 interacts electrostatically with D231 and D235 in the EF1 site of MICU1, the main-chain oxygen of residues between them, and the dipole of MICU1’s helix  $\alpha$ 5. R221 of MICU1 forms electrostatic contacts with D330 of MICU2. Multiple hydrophobic interactions also occur across the interface. The second part is reciprocal between MICU1’s C-lobe EF3 and MICU2’s N-lobe EF1,



**Fig. 3 | Uniplex structure in high- $\text{Ca}^{2+}$  conditions.** The overall structure of the high- $\text{Ca}^{2+}$  uniplex is shown in ribbon representation. Red spheres represent  $\text{Ca}^{2+}$ .



**Fig. 4 |**  $\text{Ca}^{2+}$ -induced MICU conformational changes and the mechanism of uniplex activation. **a**, Overall structure of the MICU1–MICU2 heterodimer. The interfacial EF hands of MICU1 and MICU2 are coloured in red and magenta, respectively. **b**, Interfaces between MICU1 and MICU2. **c**,  $\text{Ca}^{2+}$ -induced conformational changes in MICU1 (grey, low  $\text{Ca}^{2+}$ ; green, high  $\text{Ca}^{2+}$ ) near its

interface with MCU. **d**, Superposition of  $\text{Ca}^{2+}$ -bound and  $\text{Ca}^{2+}$ -free MICU1–MICU2 heterodimers. **e**, Molecular model of  $\text{Ca}^{2+}$  activation of the uniplex. Only two copies of MCU–EMRE in the tetramer are presented to reveal the  $\text{Ca}^{2+}$  pathway. Key amino acids are shown in single-letter code.

mediated mainly by hydrophobic interactions. Residues participating in MICU1–MICU2 interactions are mostly conserved in vertebrates (Extended Data Figs. 8e, 9). Confirming R352’s critical roles in MICU1–MICU2 dimerization, coimmunoprecipitation shows that R352E disrupts the MICU1–MICU2 complex (Extended Data Fig. 1f). Besides the main heterodimer interface, MICU1 and MICU2’s C-terminal helices also interact via an antiparallel two-helix bundle, positioning C463 of MICU1 and C413 of MICU2 in range to form a disulfide bridge.

Upon  $\text{Ca}^{2+}$  elevation, the MICU1–MICU2 heterodimer exhibits substantial conformational changes. MICU1 and MICU2 now match well with  $\text{Ca}^{2+}$ -bound but not  $\text{Ca}^{2+}$ -free MICUs<sup>28–31</sup>, and their canonical EF hands (EF1 and EF4) are compatible with  $\text{Ca}^{2+}$ -bound states (Extended Data Fig. 8c, d). Compared with the low- $\text{Ca}^{2+}$  structure, MICU1’s outer helix ( $\alpha_6$ ) of EF1 swings roughly  $30^\circ$  towards the interface edge and its C-terminal end is displaced approximately 9 Å outward (Fig. 4c). This would cause an associated outward movement of the loop after  $\alpha_6$ —which hosts R259, R261 and R263—thus weakening MCU–MICU1 interactions. Moreover, an upward movement of MICU1’s N-terminal helices  $\alpha_1$  and  $\alpha_2$  causes movement of Y121 away from the interface and an upward shift of K126 and R129, further crippling MICU1 interactions with MCU (Fig. 4c). Together,  $\text{Ca}^{2+}$  induces substantial rearrangements of key residues in MICU1 to disrupt their interactions with the relatively static MCU surface, thus facilitating MICU1 withdrawal from the MCU pore to unblock the channel.

Finally, the structures hint at cooperative activation of the uniporter<sup>12</sup>. If  $\text{Ca}^{2+}$  binds to only one MICU in the heterodimer, then the conformational changes could cause a clash between EF1 and EF3 across the heterodimer interface (Fig. 4d). If both MICU subunits bind  $\text{Ca}^{2+}$ , however, reciprocal conformational changes are accommodated across the interface, thus stabilizing the fully  $\text{Ca}^{2+}$ -occupied conformation that favours dislocation of the MICU1–MICU2 heterodimer from the pore.

## Discussion

Tight regulation of the uniporter by cytosolic  $\text{Ca}^{2+}$  signals is crucial for normal physiology<sup>1,2,9,10</sup>. Our work elucidates the structural and functional mechanisms of MICU-mediated  $\text{Ca}^{2+}$  regulation of the uniporter (Fig. 4e). We demonstrate that a single MICU1–MICU2 heterodimer

is sufficient to gate an MCU–EMRE tetramer—an interesting parallel to calmodulin-mediated inhibition of the  $\text{Ca}^{2+}$  channel TRPV5/6 and charybdotoxin-mediated block of  $\text{K}^+$  channels<sup>35–38</sup>. Rather than altering the pore conformation, MICU1 shuts the uniporter in resting conditions (at  $\text{Ca}^{2+}$  levels of less than 1  $\mu\text{M}$ ) by using a five-residue K/R ring to cover the aspartate ring at the MCU pore entrance. Upon  $\text{Ca}^{2+}$  elevation,  $\text{Ca}^{2+}$  binds to the MICU1–MICU2 heterodimer cooperatively, inducing conformational changes that weaken MCU–MICU1 interactions. The MICU heterodimer then moves away from the pore, leading to  $\text{Ca}^{2+}$  activation of the uniporter. It remains to be investigated whether MICU1 also modulates MCU allosterically via EMRE<sup>8</sup>.

In low- $\text{Ca}^{2+}$  conditions, multiple interfaces keep MICU1 tightly docked on MCU to occlude the pore. The major MCU–MICU1 contact site is formed by MICU1’s N-lobe and MCU’s IMS surface. Interfacial residues are highly conserved (Extended Data Figs. 8e, 9), particularly those in MICU1’s K/R ring and the aspartate at MCU’s pore entrance, consistent with the observation that human MICU1 can form complexes with MCUs from distant organisms<sup>22</sup>. Additionally, MICU1’s C-terminal helix helps to stabilize the MCU–MICU1 complex. The robust interactions between MCU and MICU1, which can accommodate K126E or R129E single mutations in the main interface, help to explain why a previous alanine mutagenesis screen in MICU1 failed to implicate the K/R ring<sup>22</sup>. Previous studies<sup>21,22</sup> proposed that MICU1’s R119/R154 and R440/R443 residues, which are buried in our structures, contribute to MCU binding. This is probably because R119/R154 makes critical contacts with helices that position the K/R ring to contact MCU, and R440/R443 helps to maintain the proper MICU1 conformation to interact with MCU. Thus, their mutation indirectly affects interactions with MCU. To shut MCU, MICU1 must tightly seal the pore entrance. Thus, single K/R-ring mutations could be sufficient to compromise the seal to produce  $\text{Ca}^{2+}$  leakage, but multiple mutations are needed to substantially separate MICU1 from MCU.

Animal MCUs require EMRE binding for  $\text{Ca}^{2+}$  transport<sup>8,18</sup>. This EMRE-dependent gating—distinct from the MICU-regulation mechanism described here—might involve a putative luminal gate<sup>27</sup>. Future structural work on MCU in an EMRE-free state is needed to understand how EMRE controls this process. Previous work suggested that EMRE also binds MICU1 via its C-terminal tail, an interaction necessary for

proper  $\text{Ca}^{2+}$  regulation of the uniplex<sup>18</sup>. Although the proposed contact site is unresolved in our structures, it is indeed in range for interaction. We further show that the MICU1–EMRE interaction helps to keep MICU1 in place to seal the MCU pore in low  $\text{Ca}^{2+}$ , and prevents MICU1 from dissociating from the uniplex when  $\text{Ca}^{2+}$  increases. It is conceivable that EMRE’s MICU1-tethering function might allow MICUs to stay close to the pore during  $\text{Ca}^{2+}$  stimulation so as to rapidly terminate flux once the  $\text{Ca}^{2+}$  signal dissipates. Our results show that MICU2 does not directly block MCU. By analogy, the neuron-specific MICU3 (refs.<sup>7,39</sup>) might also indirectly modulate uniporter activation through concerted conformational changes with MICU1. Finally, our structures and a previous MCU–EMRE subcomplex structure<sup>27</sup> show that the uniporter can dimerize. Such dimerization is not required for basic channel function, but may contribute to biased uniporter distribution towards inner- and outer-membrane contact sites, where nearby crista junctions could potentially provide favourable curvature to accommodate the V-shaped dimer. Such spatial arrangement could presumably maximize the uniporter’s exposure to intracellular  $\text{Ca}^{2+}$  signals, thus facilitating effective  $\text{Ca}^{2+}$  transmission from the endoplasmic reticulum to the mitochondrial matrix<sup>40</sup>.

In summary, this work defines the architecture of the major  $\text{Ca}^{2+}$ -signalling hub in mitochondria and reveals the MICU-mediated gating mechanism that underlies uniporter activation in response to intracellular  $\text{Ca}^{2+}$  signals. Our results establish a framework for understanding the principles governing tightly regulated mitochondrial  $\text{Ca}^{2+}$  transport, shed light on  $\text{Ca}^{2+}$  signalling and homeostasis, and provide a starting point to aid the development of interventions to suppress pathological  $\text{Ca}^{2+}$  overload in mitochondria, which is associated with a variety of diseases such as heart failure<sup>41</sup> and ischaemic brain injury<sup>42</sup>.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2309-6>.

- Rizzuto, R., De Stefani, D., Raffaello, A. & Mammucari, C. Mitochondria as sensors and regulators of calcium signalling. *Nat. Rev. Mol. Cell Biol.* **13**, 566–578 (2012).
- Giorgi, C., Marchi, S. & Pinton, P. The machineries, regulation and cellular functions of mitochondrial calcium. *Nat. Rev. Mol. Cell Biol.* **19**, 713–730 (2018).
- Kirichok, Y., Krapivinsky, G. & Clapham, D. E. The mitochondrial calcium uniporter is a highly selective ion channel. *Nature* **427**, 360–364 (2004).
- De Stefani, D., Raffaello, A., Teardo, E., Szabó, I. & Rizzuto, R. A forty-kilodalton protein of the inner membrane is the mitochondrial calcium uniporter. *Nature* **476**, 336–340 (2011).
- Baughman, J. M. et al. Integrative genomics identifies MCU as an essential component of the mitochondrial calcium uniporter. *Nature* **476**, 341–345 (2011).
- Perozzi, F. et al. *MICU1* encodes a mitochondrial EF hand protein required for  $\text{Ca}^{2+}$  uptake. *Nature* **467**, 291–296 (2010).
- Plovanić, M. et al. MICU2, a paralog of MICU1, resides within the mitochondrial uniporter complex to regulate calcium handling. *PLoS ONE* **8**, e55785 (2013).
- Sancak, Y. et al. EMRE is an essential component of the mitochondrial calcium uniporter complex. *Science* **342**, 1379–1382 (2013).
- Kamer, K. J. & Mootha, V. K. The molecular era of the mitochondrial calcium uniporter. *Nat. Rev. Mol. Cell Biol.* **16**, 545–553 (2015).
- De Stefani, D., Rizzuto, R. & Pozzan, T. Enjoy the trip: calcium in mitochondria back and forth. *Annu. Rev. Biochem.* **85**, 161–192 (2016).
- Malliankaraman, K. et al. MICU1 is an essential gatekeeper for MCU-mediated mitochondrial  $\text{Ca}^{2+}$  uptake that regulates cell survival. *Cell* **151**, 630–644 (2012).
- Csordás, G. et al. MICU1 controls both the threshold and cooperative activation of the mitochondrial  $\text{Ca}^{2+}$  uniporter. *Cell Metab.* **17**, 976–987 (2013).
- Patron, M. et al. MICU1 and MICU2 finely tune the mitochondrial  $\text{Ca}^{2+}$  uniporter by exerting opposite effects on MCU activity. *Mol. Cell* **53**, 726–737 (2014).
- Petrungaro, C. et al. The  $\text{Ca}^{2+}$ -dependent release of the Mia40-induced MICU1-MICU2 dimer from MCU regulates mitochondrial  $\text{Ca}^{2+}$  uptake. *Cell Metab.* **22**, 721–733 (2015).
- Peng, T. I. & Jou, M. J. Oxidative stress caused by mitochondrial calcium overload. *Ann. NY Acad. Sci.* **1201**, 183–188 (2010).
- Logan, C. V. et al. Loss-of-function mutations in MICU1 cause a brain and muscle disorder linked to primary alterations in mitochondrial calcium signaling. *Nat. Genet.* **46**, 188–193 (2014).
- Musa, S. et al. A middle eastern founder mutation expands the genotypic and phenotypic spectrum of mitochondrial MICU1 deficiency: a report of 13 patients. *JIMD Rep.* **43**, 79–83 (2018).
- Tsai, M. F. et al. Dual functions of a small regulatory subunit in the mitochondrial calcium uniporter complex. *eLife* **5**, e15545 (2016).
- Kamer, K. J., Grabarek, Z. & Mootha, V. K. High-affinity cooperative  $\text{Ca}^{2+}$  binding by MICU1-MICU2 serves as an on-off switch for the uniporter. *EMBO Rep.* **18**, 1397–1411 (2017).
- Payne, R., Hoff, H., Roskowski, A. & Foskett, J. K. MICU2 restricts spatial crosstalk between InsP3R and MCU channels by regulating threshold and gain of MICU1-mediated inhibition and activation of MCU. *Cell Rep.* **21**, 3141–3154 (2017).
- Paillard, M. et al. MICU1 interacts with the D-ring of the MCU pore to control its  $\text{Ca}^{2+}$  flux and sensitivity to Ru360. *Mol. Cell* **72**, 778–785 (2018).
- Phillips, C. B., Tsai, C. W. & Tsai, M. F. The conserved aspartate ring of MCU mediates MICU1 binding and regulation in the mitochondrial calcium uniporter complex. *eLife* **8**, e41112 (2019).
- Fan, C. et al. X-ray and cryo-EM structures of the mitochondrial calcium uniporter. *Nature* **559**, 575–579 (2018).
- Baradaran, R., Wang, C., Siliciano, A. F. & Long, S. B. Cryo-EM structures of fungal and metazoan mitochondrial calcium uniporters. *Nature* **559**, 580–584 (2018).
- Nguyen, N. X. et al. Cryo-EM structure of a fungal mitochondrial calcium uniporter. *Nature* **559**, 570–574 (2018).
- Yoo, J. et al. Cryo-EM structure of a mitochondrial calcium uniporter. *Science* **361**, 506–511 (2018).
- Wang, Y. et al. Structural mechanism of EMRE-dependent gating of the human mitochondrial calcium uniporter. *Cell* **177**, 1252–1261 (2019).
- Wang, L. et al. Structural and mechanistic insights into MICU1 regulation of mitochondrial calcium uptake. *EMBO J.* **33**, 594–604 (2014).
- Xing, Y. et al. Dimerization of MICU proteins controls  $\text{Ca}^{2+}$  influx through the mitochondrial  $\text{Ca}^{2+}$  uniporter. *Cell Rep.* **26**, 1203–1212 (2019).
- Kamer, K. J., Jiang, W., Kaushik, V. K., Mootha, V. K. & Grabarek, Z. Crystal structure of MICU2 and comparison with MICU1 reveal insights into the uniporter gating mechanism. *Proc. Natl. Acad. Sci. USA* **116**, 3546–3555 (2019).
- Wu, W. et al. The crystal structure of MICU2 provides insight into  $\text{Ca}^{2+}$  binding and MICU1-MICU2 heterodimer formation. *EMBO Rep.* **20**, e47488 (2019).
- Kamer, K. J. & Mootha, V. K. MICU1 and MICU2 play nonredundant roles in the regulation of the mitochondrial calcium uniporter. *EMBO Rep.* **15**, 299–307 (2014).
- Davies, K. M., Anselmi, C., Wittig, I., Faraldo-Gómez, J. D. & Kühlbrandt, W. Structure of the yeast  $\text{F}_{1\text{-ATP}}$  synthase dimer and its role in shaping the mitochondrial cristae. *Proc. Natl. Acad. Sci. USA* **109**, 13602–13607 (2012).
- De La Fuente, S. et al. Strategic positioning and biased activity of the mitochondrial calcium uniporter in cardiac muscle. *J. Biol. Chem.* **291**, 23343–23362 (2016).
- Singh, A. K., McGoldrick, L. L., Twomey, E. C. & Sobolevsky, A. I. Mechanism of calmodulin inactivation of the calcium-selective TRP channel TRPV6. *Sci. Adv.* **4**, eaau6088 (2018).
- Dang, S. et al. Structural insight into TRPV5 channel function and modulation. *Proc. Natl. Acad. Sci. USA* **116**, 8869–8878 (2019).
- Banerjee, A., Lee, A., Campbell, E. & Mackinnon, R. Structure of a pore-blocking toxin in complex with a eukaryotic voltage-dependent  $\text{K}^+$  channel. *eLife* **2**, e00594 (2013).
- Park, C. S. & Miller, C. Interaction of charybdotoxin with permeant ions inside the pore of a  $\text{K}^+$  channel. *Neuron* **9**, 307–313 (1992).
- Patron, M., Granatiere, V., Espino, J., Rizzuto, R. & De Stefani, D. MICU3 is a tissue-specific enhancer of mitochondrial calcium uptake. *Cell Death Differ.* **26**, 179–195 (2019).
- Rizzuto, R. et al. Close contacts with the endoplasmic reticulum as determinants of mitochondrial  $\text{Ca}^{2+}$  responses. *Science* **280**, 1763–1766 (1998).
- Santulli, G., Xie, W., Reiken, S. R. & Marks, A. R. Mitochondrial calcium overload is a key determinant in heart failure. *Proc. Natl. Acad. Sci. USA* **112**, 11389–11394 (2015).
- Starkov, A. A., Chinopoulos, C. & Fiskum, G. Mitochondrial calcium and oxidative stress as mediators of ischemic brain injury. *Cell Calcium* **36**, 257–264 (2004).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Methods

### Expression and purification of the human uniplex

Codon-optimized DNAs encoding human MCU, EMRE (with a C-terminal Strep tag), MICU1 and MICU2 were synthesized and cloned into modified BacMam expression vectors<sup>43</sup>. Recombinant baculoviruses for MCU, EMRE, MICU1 and MICU2 were generated separately using the Bac-to-Bac system as described<sup>43</sup>. HEK293S cells (ATCC catalogue number CRL-3022; no further authentication was carried out and no mycoplasma contamination was tested here) were co-infected by four viruses; 10 mM sodium butyrate was added to the culture medium after 12 h. Cells were further grown at 30 °C and harvested 72 h after infection.

Cell pellets were resuspended and homogenized in buffer A (50 mM Tris pH 7.4, 40 mM NaCl, 5 mM EGTA, 100 mM sorbitol and protease inhibitors). The crude membrane fractions were extracted using 2% lauryl maltose neopentyl glycol (LMNG, Anatrace)/0.2% cholesteryl hemisuccinate (CHS, Anatrace) at 4 °C. After 2 h of solubilization, the cell lysate was spun at 40,000g for 40 min, and the supernatant was incubated with prewashed Strep-Tactin Sepharose resin (IBA) for 2 h at 4 °C. The slurry was then poured out into a gravity-flow column (Bio-Rad). After washing the resin with buffer B (20 mM Tris pH 7.4, 40 mM NaCl, 5 mM EGTA and 0.01% LMNG/0.002% CHS), the protein was eluted with buffer C (20 mM Tris pH 7.4, 40 mM NaCl, 5 mM EGTA and 0.0033% LMNG/0.00066% CHS) containing 10 mM desthiobiotin (Sigma). To analyse the intactness of the purified human uniplex, we used a Superose 6 column (GE Healthcare) for gel filtration on an ÄKTA purifier system (GE Healthcare), using buffer B to run the gel filtration. The peak fractions were analysed by SDS-PAGE. To obtain the high-Ca<sup>2+</sup> uniplex, we eluted the protein with buffer D (20 mM Tris pH 7.4, 50 mM NaCl, 0.1 mM EGTA, 2 mM CaCl<sub>2</sub> and 0.0033% LMNG/0.00066% CHS) plus 10 mM desthiobiotin. Both samples were concentrated to about 9 mg ml<sup>-1</sup> for cryo-EM grid preparation.

### EM sample preparation and data collection

For cryo-EM, we applied 3 µl of purified human uniplex to glow-discharged 300-mesh Quantifoil R2/1 holey carbon grids, blotted them for 2.0 s at 96% humidity on a Leica EM GP2, and plunge-froze them in liquid ethane cooled by liquid nitrogen. Grids of low-Ca<sup>2+</sup> uniplex were imaged on a Titan Krios electron microscope (Thermo Fisher Scientific) operated at 300 kV using a slit width of 20 eV on a GIF Quantum energy filter. Images were collected on a K2 Summit detector (Gatan) in super-resolution counting mode at a magnification of ×130,000, corresponding to a physical pixel size of 1.06 Å. SerialEM<sup>44</sup> was used for data collection, with a set of customized scripts enabling automated low-dose image acquisition. Data were collected using image shift to collect one image per hole by the multiple record method, with a 3 × 3 set of holes per stage movement. Grids of high-Ca<sup>2+</sup> uniplex were imaged on a Talos Arctica electron microscope (Thermo Fisher Scientific) operated at 200 kV. Images were collected on a K3 Summit detector (Gatan) in super-resolution counting mode at a magnification of ×36,000, corresponding to a physical pixel size of 1.1 Å. Data were collected using image shift to collect one image per hole by the multiple record method, with a 2 × 2 set of holes per stage movement.

### Cryo-EM data processing

For the human low-Ca<sup>2+</sup> uniplex, we collected a total of 5,816 movies and subjected them to beam-induced motion correction using the program MotionCor2 (ref. <sup>45</sup>). We calculated the sum of all frames from each movie following a dose-weighting scheme, and used this for all image-processing steps. Contrast transfer function (CTF) parameters for each micrograph were estimated by CTFFIND4 (ref. <sup>46</sup>). Automated particle picking was first performed with cisTEM (ref. <sup>47</sup>) using 1,000 images, and the picked particles were extracted with a box size of 360 pixels and subjected to two-dimensional (2D) classification in cisTEM. The resulting high-quality 2D class averages representing projections in

different orientations were selected and imported to Relion 3.0 (ref. <sup>48</sup>) as templates for automatic particle picking. All particles picked in Relion were extracted with a box size of 360 pixels with the original pixel size of 1.06 Å, and then imported to cryoSPARC (ref. <sup>49</sup>) for 2D classification. Two rounds of 2D classification yielded 531,350 particle images with clear features of the uniplex.

We subjected these particles to ab initio three-dimensional (3D) reconstruction without symmetry, requesting six classes and a maximum resolution of 12 Å. Among all the six classes, one class with 93,926 particles stood out as demonstrating clear overall features of a tetramer. Non-uniform refinement of these particles produced a map at 3.9-Å resolution without symmetry. On the basis of this map, two rounds of Bayesian polishing were performed in Relion 3.0. Further 3D classification and map investigation implied the heterogeneity of MICU1/2. To improve the map quality of the MICU part, we carried out local 3D classification focused on MICU1/2. The signal of the MCU–EMRE part was subtracted from the particles with a mask covering the whole uniplex except for MICU1/2. The modified particle set was subjected to another round of 3D classification without alignment using a mask around MICU1/2 ( $k=6, T=20$ ). After classification, the class with best features of MICU1/2 was selected and the corresponding 64,131 particles were imported back to cryoSPARC for final refinement. Non-uniform refinement in cryoSPARC yielded an improved map with better details at a resolution of 3.3 Å. From the ab initio reconstruction in cryoSPARC, we observed classes representing the dimer of the uniplex, comprising 322,210 particles. Those particles were transferred to Relion for 3D classification ( $k=8, T=4$ ). The class with the highest resolution contained 69,856 particles, and yielded a map at 4.5-Å resolution after cryoSPARC non-uniform refinement without symmetry, and a map at 4.1-Å resolution with C2 symmetry. All resolutions were estimated by applying a soft mask around the protein and using the gold-standard Fourier shell correlation (FSC) = 0.143 criterion. We used BlocRes implemented in cryoSPARC to calculate the local-resolution map.

For the high-Ca<sup>2+</sup> uniplex, all steps before cryoSPARC ab initio reconstruction were the same as for the low-Ca<sup>2+</sup> uniplex. Starting from the cryoSPARC ab initio 3D reconstruction, we selected one class containing 224,504 particles with obvious MICU1/2 map density and imported it to Relion for further 3D classification. The optimized 3D classification yielded two classes with clear MICU1/2 density, containing a total of 128,221 particles. Those particles were refined with a C2 symmetry using non-uniform refinement, yielding a map at 4.1-Å resolution. Further particle polishing based on the map at 4.1-Å resolution and additional 3D classification improved the map quality and resolution to 3.6 Å. Local resolution was calculated using BlocRes implemented in cryoSPARC. The number of particles in each data set and other details related to data processing are summarized in Extended Data Figs. 2, 4 and Extended Data Table 1.

### Model building and refinement

The low-Ca<sup>2+</sup> uniplex model was built into a 3.3-Å cryo-EM map using available structures of uniplex components as templates (Protein Data Bank (<https://www.rcsb.org>) codes: MCU-EMRE, PDB 6O5B; MICU1, PDB 4NSC; MICU2, PDB 6EAZ and 6AGH). The map shows excellent densities in MCU, EMRE, MICU1 and the majority of MICU2. Modelling on the distal part of MICU2 that shows relatively weaker density was facilitated by the available high-resolution crystal structures of MICU2 (refs. <sup>29,30</sup>). Local parts were manually rebuilt in Coot<sup>50</sup>. The high-Ca<sup>2+</sup> uniplex model was built into a 3.6-Å cryo-EM map using available structures of uniplex components as templates (MCU-EMRE, PDB 6O5B; MICU1, PDB 4NSD; MICU2, PDB 6IIH). High-quality cryo-EM maps and the availability of the structures of individual components<sup>27,28,31</sup> facilitated model building. Local parts were manually adjusted in Coot. The models were refined using Phenix real space refine<sup>51</sup>, and the geometry of the models was evaluated by Molprobity<sup>52</sup>. All figures were prepared in PyMol (Schrödinger) or UCSF Chimera<sup>53</sup>.

## Cell culture and molecular biology

Genes encoding uniporter subunits were cloned into the pcDNA3.1(+) expression vector. Site-directed mutagenesis was performed using a QuickChange kit (Agilent) and confirmed with Sanger sequencing. All MCU and MICU1 constructs contain a C-terminal 1D4 tag (TETSQVAPA) and a C-terminal FLAG tag (DYKDDDDK), respectively, for coimmunoprecipitation and western detection. HEK293 cells (ATCC catalogue number CRL-11268, authenticated by ATCC via short tandem repeat (STR) profiling; mycoplasma contamination was ruled out using polymerase chain reaction) were cultured in Dulbecco's modified Eagle's medium (Gibco) supplemented with 10% fetal bovine serum (FBS), and were incubated at 37 °C with 5% CO<sub>2</sub>. CRISPR knockout cell lines have been documented previously<sup>18</sup>. Transient transfection was performed using lipofectamine 3000 (Invitrogen), following the manufacturer's instructions. Cells were harvested for experiments 24–48 h after transfection.

## Coimmunoprecipitation

Coimmunoprecipitation experiments were performed at 4 °C. Transfected HEK 293 cells in a 60-mm dish were lysed in 0.5 ml solubilization buffer (100 mM NaCl, 20 mM Tris, 1 mM EGTA, 5 mM n-dodecyl β-D-maltoside (DDM), adjusted to pH 7.5 with HCl) supplemented with an EDTA-free protease-inhibitor cocktail ('complete ultra', Roche). When high Ca<sup>2+</sup> conditions were needed, 1 mM EGTA in the solubilization buffer was substituted with 1 mM CaCl<sub>2</sub>. The lysate was clarified by spinning down. We removed 50 µl of the supernatant for total protein determination using a bicinchoninic acid (BCA) assay. We used 10–50 µg of protein for whole-cell lysate analysis. We then added 25 µl of FLAG-conjugated (Sigma Aldrich, catalogue number A2220) or 1D4-conjugated (homemade, 50% slurry) beads to the rest of the supernatant for batch binding (30 min). The beads were collected on a spin column, washed five times with 1 ml of solubilization buffer, and then eluted with 0.2 ml of SDS loading buffer. We used 10–60 µl of the elute for SDS-PAGE/western blot.

For western blots, proteins on SDS gels were transferred to low-fluorescence polyvinylidene fluoride (PVDF) membranes (EMD-Millipore), which were blocked in a Tris-buffered saline (TBS)-based intercept blocking buffer (Li-Cor), and then incubated with primary antibodies in TBST (TBS plus 0.075% Tween-20) at 4 °C overnight. After a 1-h incubation with infrared fluorescent secondary antibodies at room temperature, signals were acquired using an Odyssey CLx imaging system (Li-Cor), and quantified using ImageStudio software (Li-Cor version 5.0). MCU and MICU1 were detected using anti-1D4 and anti-FLAG antibodies, respectively. Primary antibodies and dilutions are as follows: anti-FLAG (Sigma Aldrich F1804, 1:10,000), anti-MICU2 (Abcam ab101465, 1:10,000), anti-1D4 (homemade, 50 ng ml<sup>-1</sup>), anti-Letm1 (Abcam ab55434, 1:2,000), anti-COX2 (Abcam ab110258, 100 ng ml<sup>-1</sup>), anti-TIM23 (Santa Cruz sc-514463, 100 ng ml<sup>-1</sup>), anti-mitofilin (Abcam ab110329, 100 ng ml<sup>-1</sup>), anti-VDAC1 (Abcam ab14734, 100 ng ml<sup>-1</sup>), and anti-EMRE (Santa Cruz, 86337, 1:400). Secondary antibodies were goat anti-rabbit IRDye 680RD (Li-Cor, 1:10,000) and goat anti-mouse IRDye 680RD (Li-Cor, 1:15,000).

## Mitochondrial Ca<sup>2+</sup>-flux assays

For the fluorescence-based assay to test uniporter activity in high Ca<sup>2+</sup>, 2 × 10<sup>7</sup> HEK 293 cells were suspended in 10 ml of wash buffer (120 mM KCl, 25 mM HEPES, 2 mM KH<sub>2</sub>PO<sub>4</sub>, 1 mM MgCl<sub>2</sub>, 50 µM EGTA, adjusted to pH 7.2 with KOH), pelleted, and then resuspended in 2.2 ml of recording buffer (120 mM KCl, 25 mM HEPES, 2 mM KH<sub>2</sub>PO<sub>4</sub>, 5 mM succinate, 1 mM MgCl<sub>2</sub>, 5 µM thapsigargin, pH 7.2 with KOH). Then, 2 ml of the cell suspension was placed in a stirred quartz cuvette in a Hitachi F-7100 spectrophotometer (ex: 506 nm; ex-slit: 2.5 nm; em: 532 nm; em-slit: 2 nm; sampling rate: 2 Hz). Reagents were added into the cell suspension in the following order: 0.25 µM calcium green 5N (Thermo, catalogue

number C3737), 30 µM digitonin (Sigma Aldrich, D141), 10 µM CaCl<sub>2</sub> and 75 nM Ru360 (homemade). Quantification was done by linear fit to the fluorescent signal between 5 s and 10 s after adding Ca<sup>2+</sup>. The slope after Ru360 addition was subtracted to yield uniporter-specific uptake.

For the <sup>45</sup>Ca<sup>2+</sup>-based assay to test uniporter activity in low Ca<sup>2+</sup>, 2 × 10<sup>6</sup> cells were suspended in 0.5 mL wash buffer, spun down, and then resuspended in 110 µl wash buffer, supplemented with 5 µM thapsigargin (Sigma Aldrich, T9033) and 30 µM digitonin. To initiate mitochondrial Ca<sup>2+</sup> uptake, 100 µl cell suspension was transferred to 300 µl flux buffer (recording buffer plus 0.69 mM EGTA, 0.5 mM CaCl<sub>2</sub>, 20 µM <sup>45</sup>CaCl<sub>2</sub>, 30 µM digitonin, pH 7.0 KOH). The total free Ca<sup>2+</sup> in flux buffer is roughly 300 nM, as calibrated by Fluo-4. At 2, 4 and 6 min after the reaction started, Ca<sup>2+</sup> uptake was terminated by adding 100 µl of the sample to 5 mL ice-cold wash buffer, and then filtered through 0.45-µm nitrocellulose membranes on a vacuum filtration manifold. The membrane was washed immediately with 5 mL ice-cold wash buffer, and later transferred into scintillation vials for counting. The data were fit with a linear function to obtain the rate of Ca<sup>2+</sup> transport. <sup>45</sup>Ca<sup>2+</sup> radioisotope was obtained from Perkin Elmer, and has a specific activity of 12–18 mCi mg<sup>-1</sup>.

## Isolation of inner/outer-membrane contact sites

HEK293 cells cultured in six 15-cm dishes were harvested, and mitochondria were extracted as described<sup>18</sup>. All of the following procedures were performed at 4 °C. Pelleted mitochondria were resuspended in 25 mL of swelling buffer (0.5 mM EDTA, 20 mM 3-(N-morpholino)propanesulfonic acid (MOPS), pH 7.4 KOH) and incubated for 5 min. Then, 10 mL of 60% sucrose was added. After 5 min of incubation, the sample was sonicated three times for 30 s with a 30 s break in between using a sonic dismembrator (Model 505, Fisher) at a power level of 3. Intact mitochondria were removed by spinning at 20,000g for 20 min. The supernatant was then transferred to an ultracentrifuge tube with 0.3 mL of 60% sucrose added as a cushion. After 2 h of spinning at 100,000g, membrane vesicles were harvested into a 1 mL suspension with 20% sucrose. Subsequently, 0.2 mL of the sample was loaded onto a 30–40–50–60% (0.7 mL per layer) discontinuous sucrose gradient, and spun at 200,000g for 18 h. Afterwards, 14 of 0.2–0.22-mL fractions were taken from the bottom of the tube for western blot analysis.

## Statistics

All functional experiments were repeated with at least three independent measurements, and the data are presented as means ± s.e.m. Statistical analysis was performed with two-tailed Student's *t*-test, with significance defined as *P* < 0.05. No statistical methods were used to predetermine sample size. The 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

The three-dimensional cryo-EM density maps are deposited into the Electron Microscopy Data Bank (<https://www.ebi.ac.uk/pdbe/emdb/>) under accession numbers EMD-21642 and EMD-21643. The coordinates are deposited in the Protein Data Bank (<https://www.rcsb.org>) with accession numbers 6WDN and 6WDO.

43. Goehring, A. et al. Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nat. Protoc.* **9**, 2574–2585 (2014).
44. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
45. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

## Article

46. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
47. Grant, T., Rohou, A. & Grigorieff, N. cisTEM, user friendly software for single-particle image processing. *eLife* **7**, e35383 (2018).
48. Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
49. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
50. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
51. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
52. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
53. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
54. Ohlendieck, K., Riesinger, I., Adams, V., Krause, J. & Brdiczka, D. Enrichment and biochemical characterization of boundary membrane contact sites from rat-liver mitochondria. *Biochim. Biophys. Acta* **860**, 672–689 (1986).

**Acknowledgements** We thank L. Montabana and D.-H. Chen at Stanford–SLAC cryo-EM facilities; C. Xu, K. Lee and K. Song at the University of Massachusetts cryo-EM facility for help

with collecting EM data; A. Van Keuren for technical assistance; and S. R. Levinson for assistance in establishing the submitochondrial fractionation assay. This work was made possible by support from Stanford University and the Harold and Leila Y. Mathers Charitable Foundation to L.F.; an American Heart Association postdoctoral fellowship to M.F.; and a Dean's fellowship to J.Z. M.-F.T., C.-W.T. and M.R. are supported by the National Institutes of Health (NIH) grant R01-GM129345.

**Author contributions** M.F. and J.Z. carried out biochemical experiments and cryo-EM studies. C.-W.T. carried out functional studies. B.J.O. contributed to EM studies. M.R. contributed to functional studies. Y.X. contributed to biochemical and functional characterizations. M.L. advised the EM studies. M.-F.T. oversaw the functional studies. L.F. directed the overall project. M.F., M.-F.T. and L.F. wrote the manuscript with input and support from J.Z., B.J.O., C.-W.T. and M.L.

**Competing interests** The authors declare no competing interests.

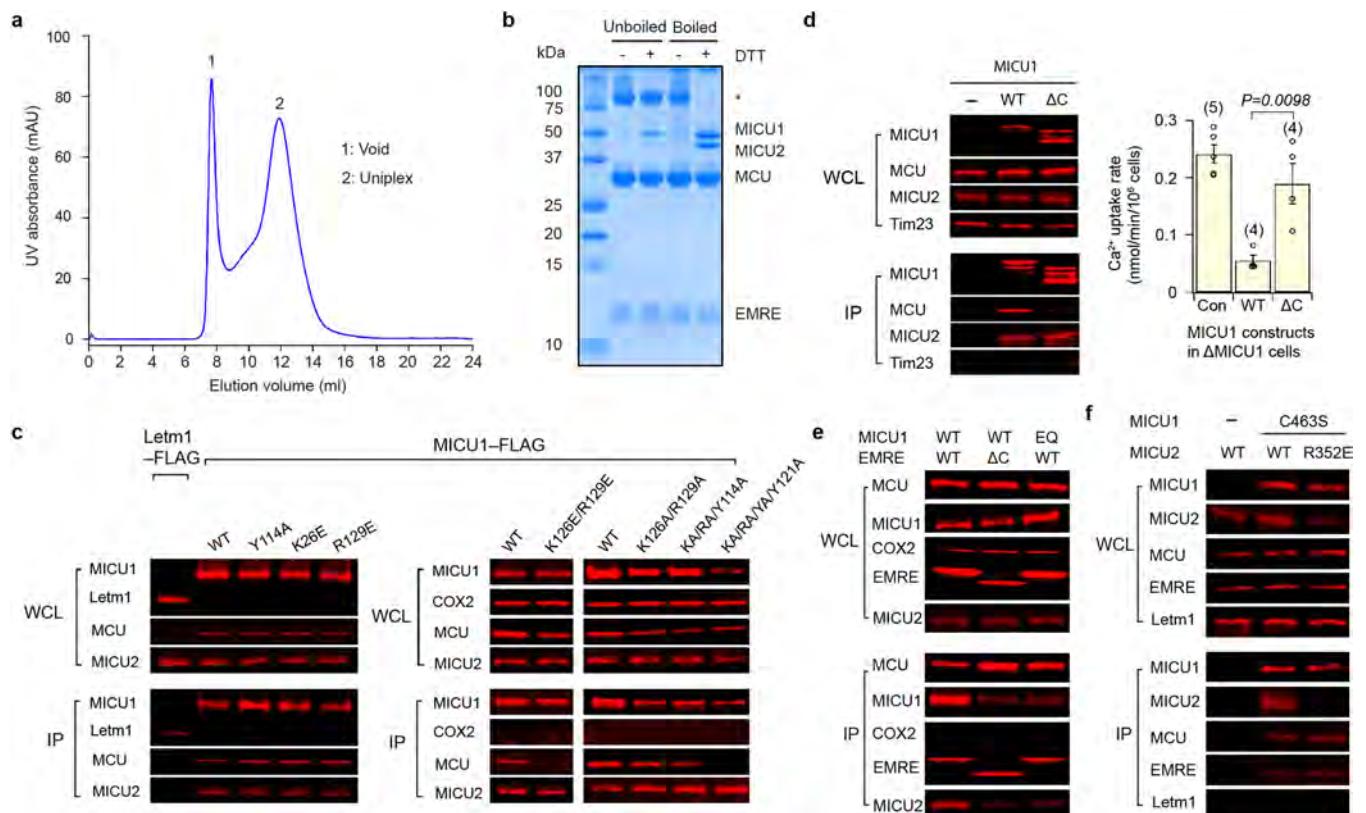
### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2309-6>.

**Correspondence and requests for materials** should be addressed to M.-F.T. or L.F.

**Peer review information** *Nature* thanks Murali Prakriya, Alexander I. Sobolevsky and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

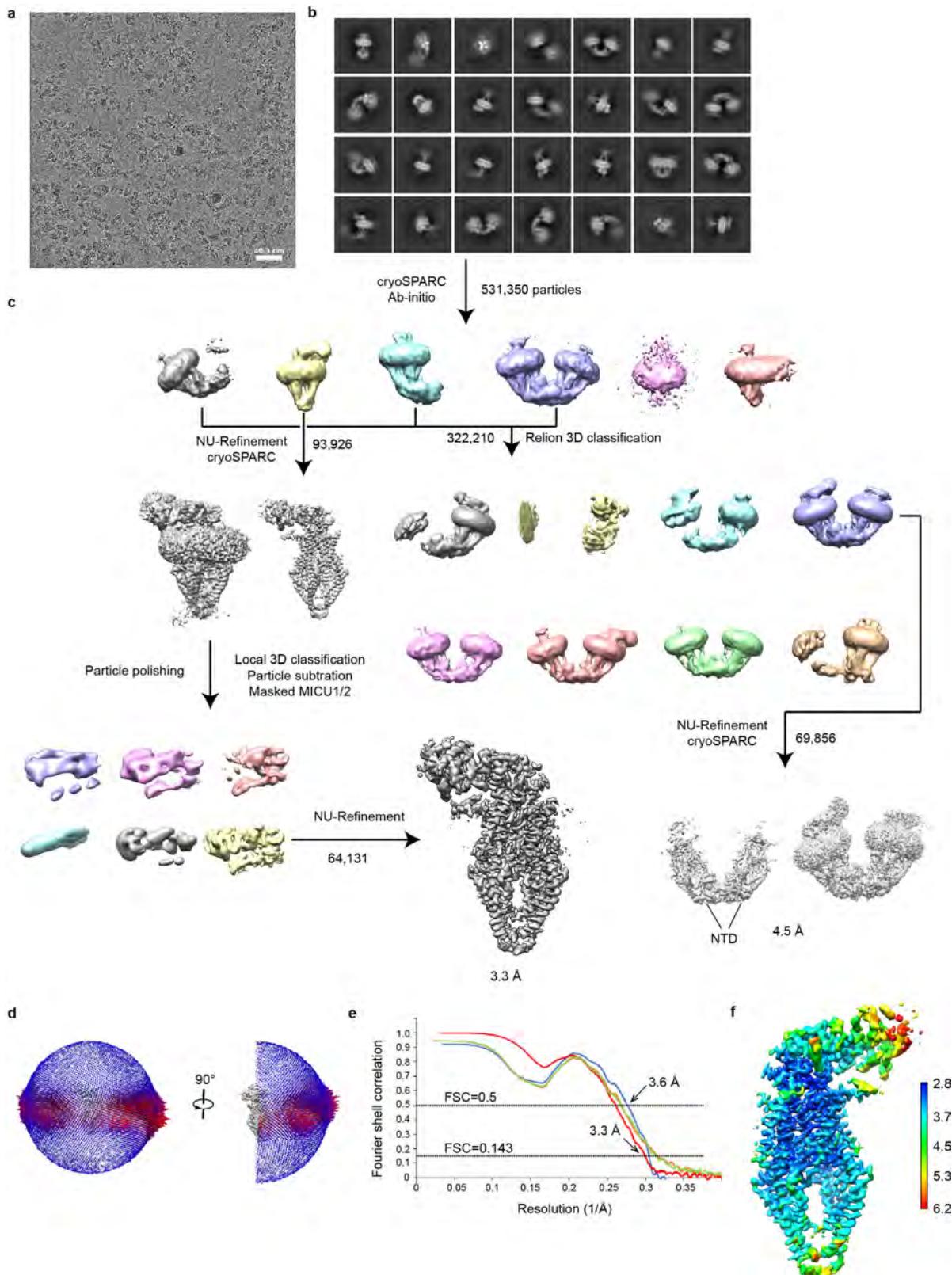


#### Extended Data Fig. 1 | Biochemical characterization of the purified human uniplex and validation of the interfaces of the low-Ca<sup>2+</sup> uniplex.

**a**, Size-exclusion chromatography profile of the purified human uniplex. **b**, SDS-PAGE analysis of the purified human uniplex. The disulfide-linked MICU1–MICU2 heterodimer is labelled with an asterisk. Data in **a**, **b** are representative of five independent experiments with similar results. **c**, Effects of MCU–MICU1 interfacial mutations on complex stability. FLAG-tagged MICU1 was immobilized to pull down ID4-tagged wild-type (WT) MCU co-expressed in MCU/EMRE/MICU1-knockout cells. IP, immunoprecipitate (eluted protein); WCL, whole-cell lysate. KA, RA and YA stand for K126A, R129A and Y114A mutations, respectively. In a separate control coimmunoprecipitation experiment, Letm1-FLAG was expressed alone in MCU/EMRE/MICU1-knockout cells, which were solubilized and incubated with FLAG beads. The eluent was then analysed. Letm1 in WCL and immunoprecipitate was detected using anti-FLAG and anti-Letm1 antibodies, respectively. Cytochrome c oxidase subunit 2 (COX2) serves as a control, showing that MICU1 does not interact nonspecifically with other mitochondrial inner-membrane proteins. MICU2 signals were obtained by targeting native proteins. An EMRE blot was not performed, as the EMRE gene was deleted in these cells. MICU1 mutants were properly folded: they still formed a complex with MICU2. **d**, Functional roles of MICU1's C-terminal helix. In coimmunoprecipitation, wild-type MCU, wild-type MICU2 and FLAG-tagged MICU1 constructs were expressed in MCU/EMRE/MICU1-knockout cells, with MICU1 used to pull down other subunits. C-terminal truncation (ΔC, residues 445–476 deleted) of MICU1 greatly weakens its interaction with MCU without affecting MICU2 binding. Tim23,

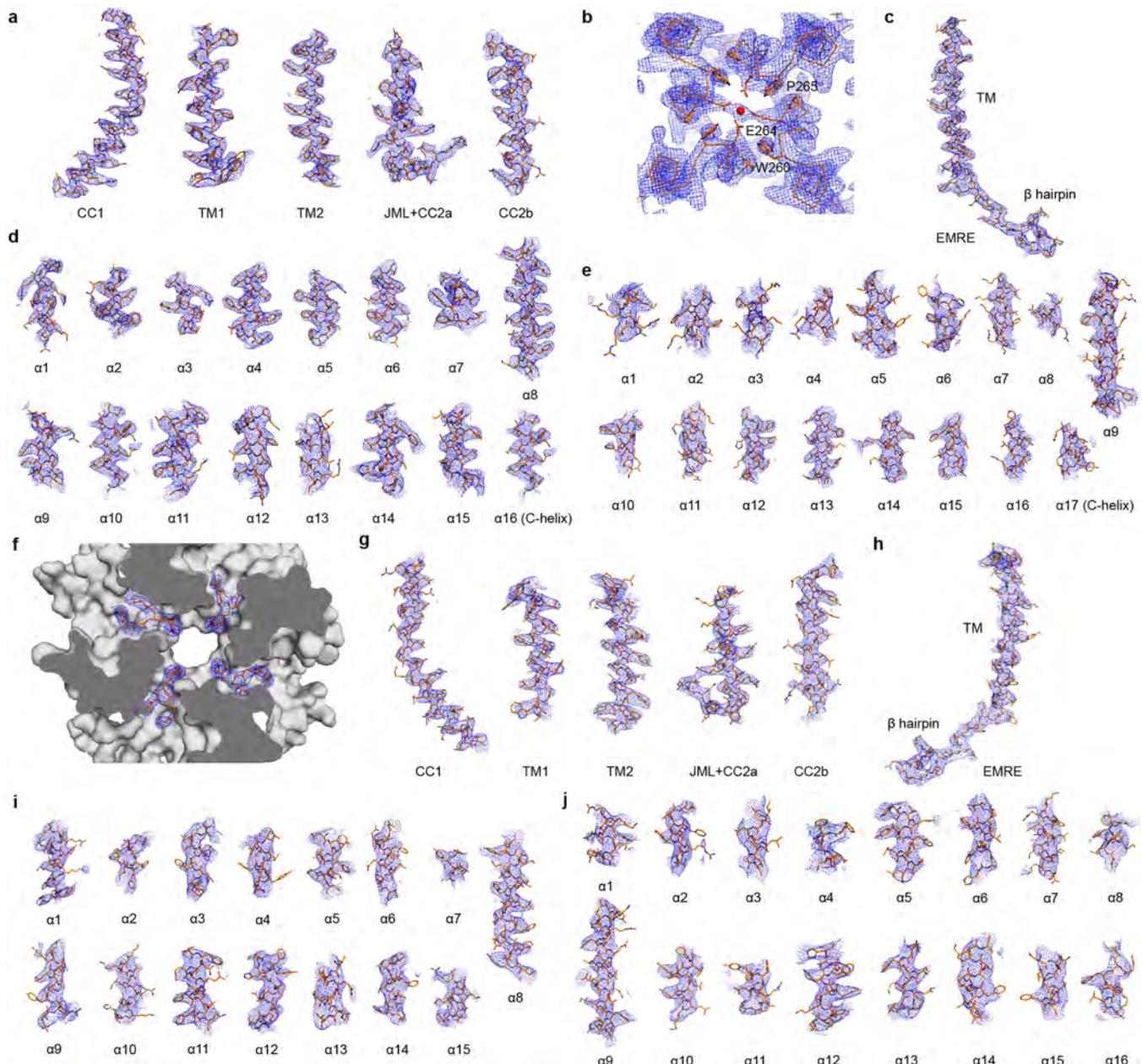
is a membrane-embedded component of the mitochondrial translocase of the inner membrane, was used to rule out nonspecific binding. The bar chart summarizes the effect of MICU1 C-truncation on the gatekeeping function. Wild-type or ΔC MICU1 was expressed in MICU1-knockout cells, and mitochondrial Ca<sup>2+</sup> uptake in low-Ca<sup>2+</sup> conditions (300 nM) was quantified using <sup>45</sup>Ca<sup>2+</sup> flux. Results are shown as means ± s.e.m. Numbers of independent repeats are provided in parentheses. ΔC MICU1 has a much weaker ability to gate MCU than wild-type MICU1, as determined by two-tailed *t*-test (*P* = 0.0098). Con, untransfected MICU1-knockout cells. **e**, Roles of the MICU1–EMRE interaction in uniplex stability. The experiment assessed the complex stability of wild-type MCU and the indicated MICU1 constructs in the presence of wild-type or C-truncated (residues 96–107 deleted) EMRE in low-Ca<sup>2+</sup> conditions. These three subunits were coexpressed in MCU/MICU1/EMRE-knockout cells. C-truncation of EMRE or charge-reversal of MICU1's KKRR sequence to QEQQEQ (EQ) greatly weakens MICU1's association with MCU. **f**, Contribution of residue R352 to formation of the MICU1–MICU2 heterodimer. Complex formation between C463S MICU1, which cannot form a disulfide MICU dimer, and wild-type or R352E MICU2 was examined in MICU1/MICU2-knockout cells. The R352E mutation in MICU2 strongly perturbs dimerization with MICU1. Letm1, detected using anti-Letm1 antibody, serves as control for nonspecific interactions. MCU and EMRE signals reflect native proteins. All coimmunoprecipitation experiments (**c**–**f**) were performed four times with similar results using independent biological samples. For gel source data, see Supplementary Fig. 1.

# Article



**Extended Data Fig. 2 | Single-particle cryo-EM analysis of the uniplex in low  $\text{Ca}^{2+}$  conditions.** **a**, Representative cryo-EM image of the purified uniplex in low  $\text{Ca}^{2+}$  conditions. **b**, 2D class averages of the uniplex. **c**, Workflow for classification and refinement. NU-refinement, non-uniform refinement. **d**, Angle distributions of the particles for the final reconstruction. **e**, Fourier

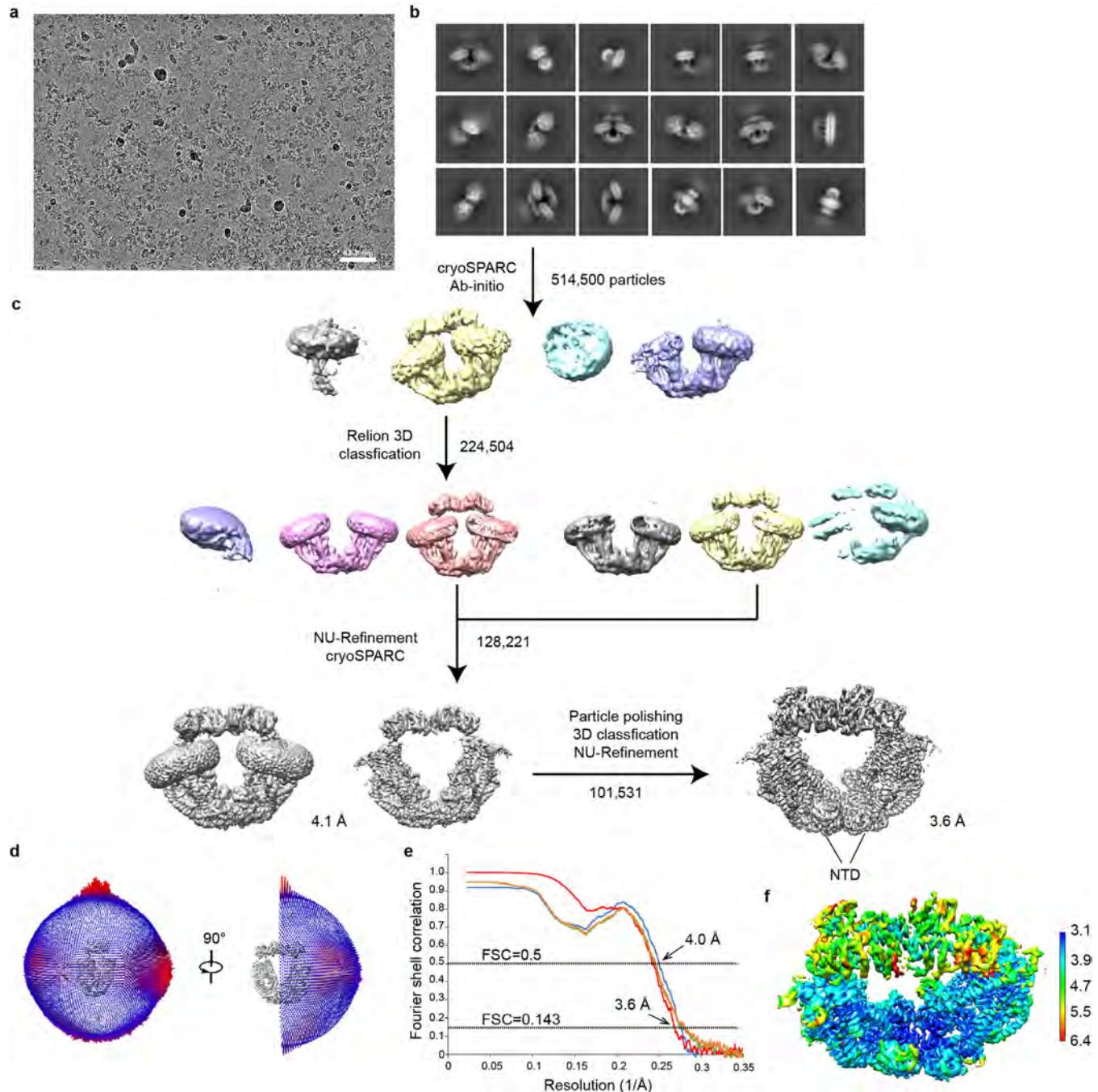
shell correlation (FSC) of the final reconstruction as a function of resolution. Red, gold-standard FSC curve,  $\text{FSC} = 0.143$ ; blue,  $\text{FSC} = 0.5$ ; orange, FSC curve between the final model and half-map 1; green, FSC curve between the final model and half-map 2. **f**, Local resolution of the map calculated by BlocRes.



**Extended Data Fig. 3 | Representative cryo-EM density maps of the uniplex in low-Ca<sup>2+</sup> and high-Ca<sup>2+</sup> conditions.** **a, b**, Cryo-EM density maps of MCU (**a**) and its selectivity filter (**b**) for the low-Ca<sup>2+</sup> uniplex. The putative cation is shown as a red sphere in **b**. **c**, Cryo-EM density map of EMRE of the low-Ca<sup>2+</sup>

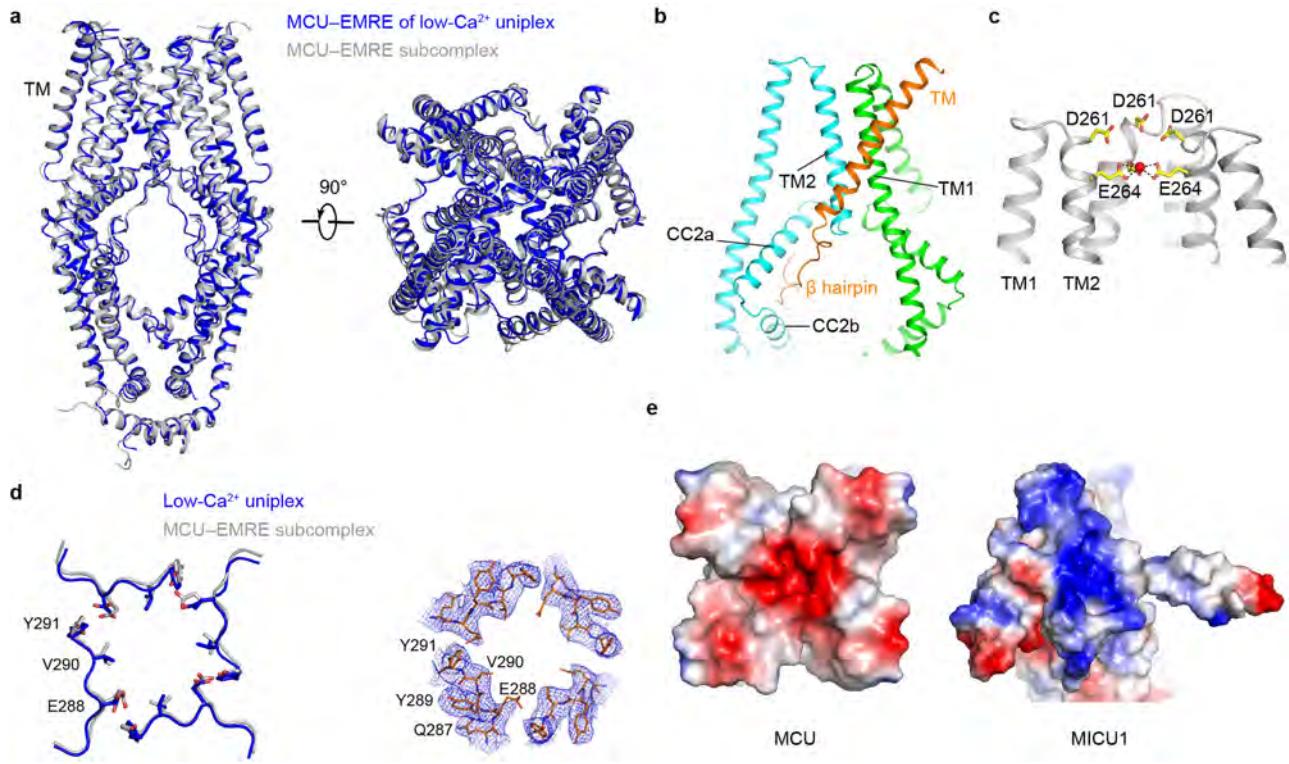
uniplex. **d, e**, Cryo-EM density of the α-helices in MICU1 (**d**) and MICU2 (**e**) from the low-Ca<sup>2+</sup> uniplex. **f**, Cryo-EM density maps of lipids bound to the MCUTM region of the low-Ca<sup>2+</sup> uniplex. **g–j**, Cryo-EM density maps of MCU (**g**), EMRE (**h**), MICU1 (**i**), and MICU2 (**j**) of the high-Ca<sup>2+</sup> uniplex.

# Article

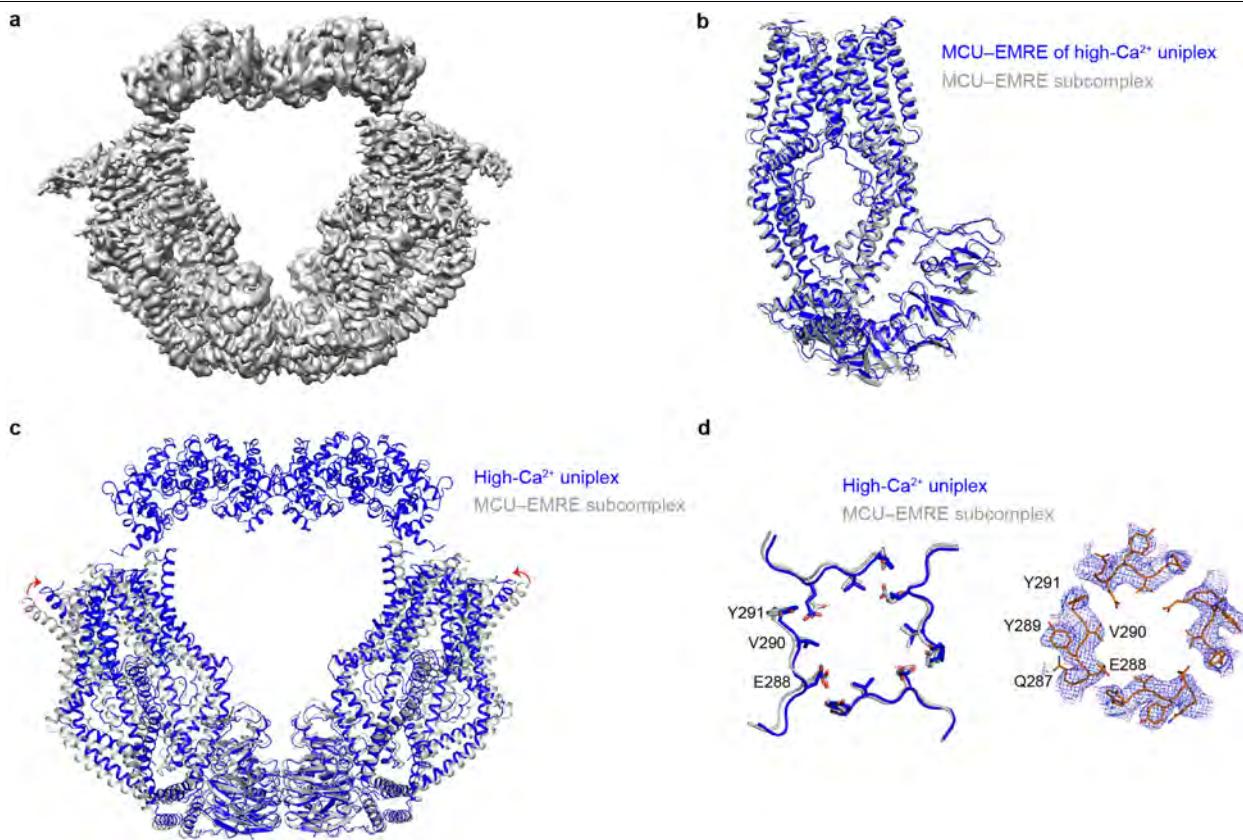


**Extended Data Fig. 4 | Single-particle cryo-EM analysis of the uniplex in high-Ca<sup>2+</sup> conditions.** **a**, Representative cryo-EM image of the uniplex in high-Ca<sup>2+</sup> conditions. **b**, 2D class averages of the uniplex. **c**, Workflow of classification and refinement. **d**, Angle distributions of the particles for the final reconstruction. **e**, The final reconstruction FSC as a function of

resolution. Red, gold-standard FSC curve, FSC = 0.143; blue, FSC = 0.5; orange, FSC curve between the final model and half-map 1; green, FSC curve between the final model and half-map 2. **f**, Local resolution of the map calculated using BlocRes.

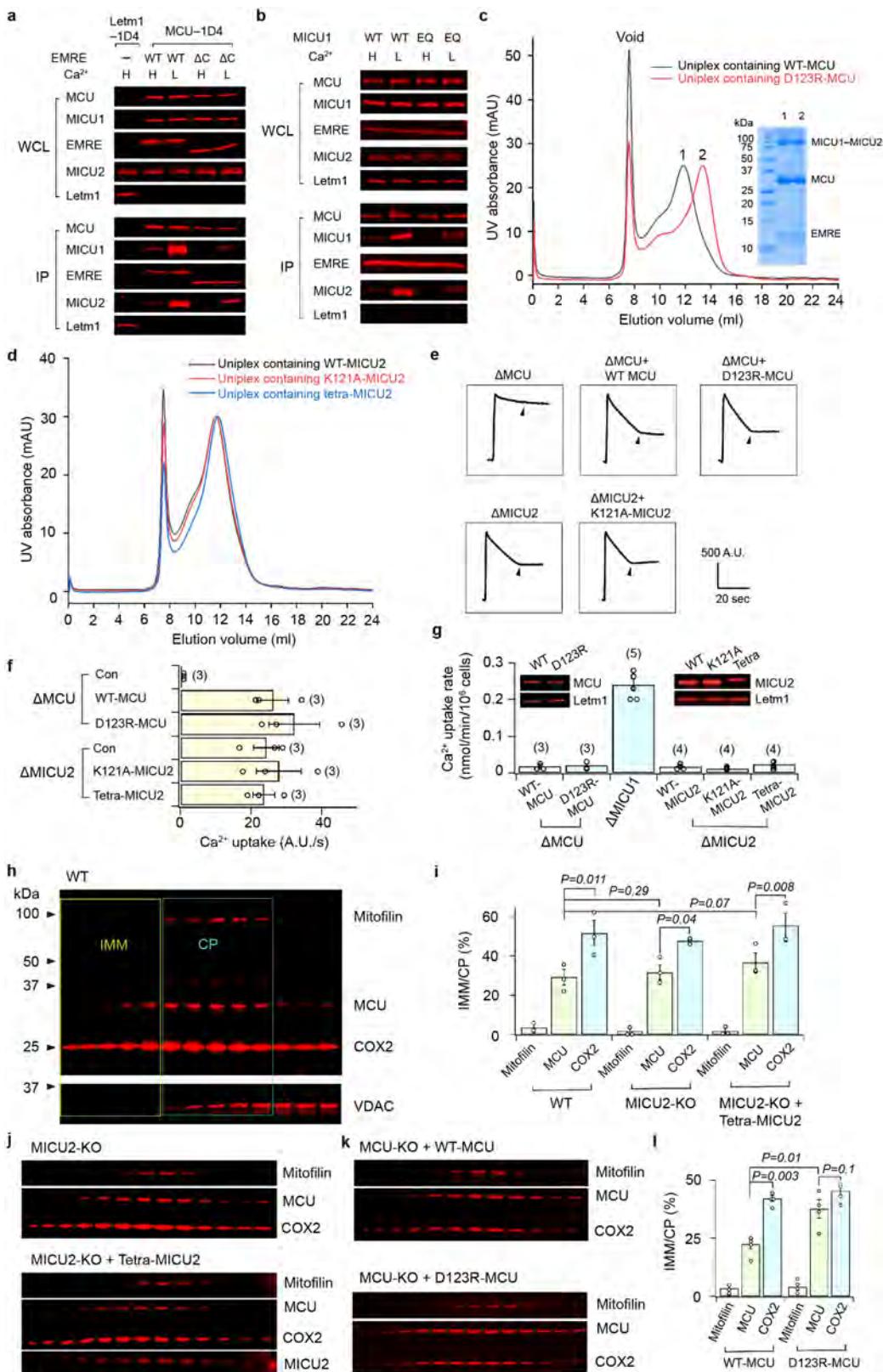


**Extended Data Fig. 5 | Structural comparison of low-Ca<sup>2+</sup> uniplex and the MCU-EMRE subcomplex.** **a**, Structural superposition of the MCU-EMRE part of low-Ca<sup>2+</sup> uniplex (blue) and the MCU-EMRE subcomplex (grey). **b**, Interactions between MCU and EMRE in the uniplex. Two MCU subunits are coloured green and cyan, and one EMRE is coloured orange. **c**, The selectivity filter of MCU in the uniplex. The side chains of D261 and E264 are shown as sticks. The putative cation is shown as a red sphere. **d**, Comparison of the luminal gate of MCU in the uniplex (blue) and the MCU-EMRE subcomplex (grey). The cryo-EM density of the uniplex luminal gate is shown on the right. **e**, Surface representation of the MCU-MICU1 interface, coloured according to electrostatic potential (red, negative; blue, positive).



**Extended Data Fig. 6 | Structural comparison of the high-Ca<sup>2+</sup> uniplex and MCU-EMRE subcomplex.** **a**, Cryo-EM map of the high-Ca<sup>2+</sup> uniplex. **b**, Superposition of one copy of the high-Ca<sup>2+</sup> uniplex (blue) and the MCU-EMRE subcomplex (grey). The MICU1 and MICU2 parts of the uniplex are

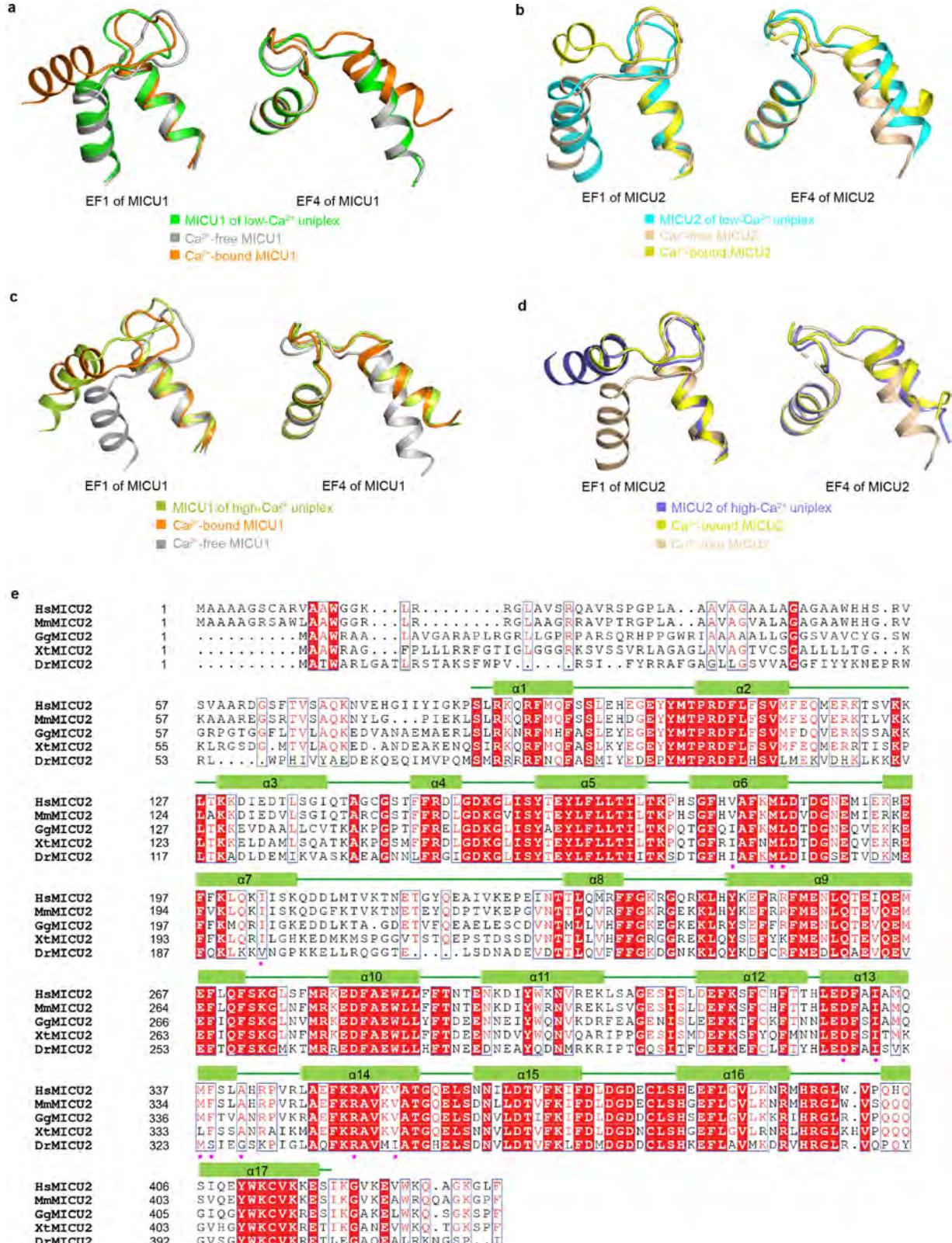
omitted for clarity. **c**, Superposition of the dimeric high-Ca<sup>2+</sup> uniplex and MCU-EMRE subcomplex. **d**, Comparison of the luminal gate of MCU in the high-Ca<sup>2+</sup> uniplex (blue) and the MCU-EMRE subcomplex (grey). The cryo-EM density of the uniplex luminal gate is shown on the right.



**Extended Data Fig. 7** | See next page for caption.

# Article

**Extended Data Fig. 7 | Validation of the interface of the high-Ca<sup>2+</sup> uniplex and functional roles of uniplex dimer interfaces.** **a**, The effect of Ca<sup>2+</sup> elevation and EMRE C-truncation on MICU1's association with the uniplex. 1D4-tagged wild-type (WT) MCU was used to precipitate wild-type MICU1 and the indicated EMRE constructs in high- or low-Ca<sup>2+</sup> conditions (H and L, respectively). All three subunits were expressed in MCU/MICU1/EMRE-knockout cells. The Letm1 control was performed as in Extended Data Fig. 1c, but with an 1D4-tagged, rather than FLAG-tagged, version of Letm1. Letm1 in WCL or immunoprecipitate (IP) was detected with anti-1D4 or anti-Letm1 antibodies, respectively. Four independent experiments were performed, yielding similar results. **b**, Roles of MICU1's polybasic sequence in MICU1 binding to the MCU–EMRE tetramer. The image compares the stability of wild-type MCU complexed with wild-type MICU1 or the KKKKR-to-QEQEQ mutant (EQ) of MICU1 in low or high Ca<sup>2+</sup>. Wild-type MCU/EMRE and MICU1 constructs were expressed in MCU/MICU1/EMRE-knockout cells. Four independent repeats were performed, leading to similar results. **c**, Size-exclusion chromatography profiles of the purified human uniplex containing wild-type or D123R MCU. The inset shows SDS-PAGE gel analysis of the uniplex. Data are representative of three independent experiments with similar results. **d**, Size-exclusion chromatography profiles of the purified human uniplex containing wild-type or mutant MICU2. The uniplex was expressed in MICU2-knockout HEK293 cells to eliminate the effect of endogenous MICU2. The experiment was performed twice independently with similar results. Tetra, R107E–R120E–K121E–D154R mutant MICU2. **e–g**, Functional roles of the uniporter's dimer interfaces. We analysed the D123R MCU mutant expressed in MCU-knockout cells, or K121A or R107E–R120E–K121E–D154R (tetra) MICU2 mutants expressed in MICU2-knockout cells, using a standard fluorophore-based mitochondrial Ca<sup>2+</sup>-uptake assay in 10 μM Ca<sup>2+</sup> (**e**, **f**) or by <sup>45</sup>Ca<sup>2+</sup> flux in 300 nM Ca<sup>2+</sup> (**g**). Numbers in parentheses indicate numbers of independent repeats. Arrowheads in **e** indicate the addition of Ru360. Con, untransfected cells. In <sup>45</sup>Ca<sup>2+</sup> flux experiments, wild-type MICU1 was coexpressed with wild-type or D123R MCU in MCU-knockout cells to ensure sufficient copies of MICU1 to gate MCU (1 μg MCU and 2 μg MICU1 DNA per well in six-well plates). The tetra-MICU2 construct has lower expression levels despite using threefold more DNA in transient expression. **h–j**, Localization of the uniporter in the mitochondrial inner membrane of wild-type (**h**) or MICU2-knockout (KO; **j**) cells. Mitochondrial membrane fractions enriched in outer membrane, inner/outer-membrane contact points (CP) or inner membrane (IMM) were separated in a sucrose gradient as described<sup>54</sup>. COX2, mitofillin and VDAC were used as the markers for inner membrane, inner/outer-membrane contact points and outer membrane, respectively. We found MCU to be more enriched in contact points (**h**, **i**). This feature was not affected by knocking out MICU2 or expressing tetra-MICU2 in MICU2-knockout cells (**i**, **j**). The sucrose gradient goes from 60% to 30% from left to right. The bar chart in **i** presents the ratio of total western signals in IMM (yellow box in **h**) over the signals in contact points (cyan box in **h**). Three biologically independent experiments were carried out, with similar results (**h**, **j**), as summarized in the bar chart (**i**). Two-tailed *t*-tests were performed, with *P*-values labelled on the bar chart. **k–l**, Effect of the D123R mutation on uniporter distribution. D123R or wild-type MCU was expressed in MCU-knockout cells, and MCU localization was analysed. D123R reduces the biased distribution of MCU in contact points. Four biologically independent experiments were performed, with similar results (**k**), as summarized in the bar chart (**l**). Statistical analyses were carried out through two-tailed *t*-tests. All bar charts (**f**, **g**, **i**, **l**) present data as means ± s.e.m. For gel source data, see Supplementary Fig. 1.



**Extended Data Fig. 8 | MICU1–MICU2 heterodimer in low-Ca<sup>2+</sup> and high-Ca<sup>2+</sup> conditions and sequence alignment of MICU2.** **a**, Structural comparison of canonical EF hands 1 and 4 of MICU1 in the low-Ca<sup>2+</sup> uniplex (green) with those known in the Ca<sup>2+</sup>-free (grey) and Ca<sup>2+</sup>-bound (orange) states. **b**, Structural comparison of canonical EF hands 1 and 4 of MICU2 in the low-Ca<sup>2+</sup> uniplex (cyan) versus those known in the Ca<sup>2+</sup>-free (wheat) and Ca<sup>2+</sup>-bound (yellow) states. **c**, Structural comparison of canonical EF hands 1 and 4 of MICU1 in the high-Ca<sup>2+</sup> uniplex (lemon) with those known in the Ca<sup>2+</sup>-bound (orange) and Ca<sup>2+</sup>-free (grey) states. **d**, Structural comparison of canonical EF hands 1 and 4 of MICU2 in the high-Ca<sup>2+</sup> uniplex (blue) versus those known in the Ca<sup>2+</sup>-bound (yellow) and Ca<sup>2+</sup>-free (wheat) states. **e**, Sequence alignment of MICU2 homologues from *Homo sapiens* (Hs), *Mus musculus* (Mm), *Gallus gallus* (Gg), *Xenopus tropicalis* (Xt) and *Danio rerio* (Dr). The residues participating in MICU1–MICU2 interactions are indicated with magenta circles.

high-Ca<sup>2+</sup> uniplex (lemon) with those known in the Ca<sup>2+</sup>-bound (orange) and Ca<sup>2+</sup>-free (grey) states. **d**, Structural comparison of canonical EF hands 1 and 4 of MICU2 in the high-Ca<sup>2+</sup> uniplex (blue) versus those known in the Ca<sup>2+</sup>-bound (yellow) and Ca<sup>2+</sup>-free (wheat) states. **e**, Sequence alignment of MICU2 homologues from *Homo sapiens* (Hs), *Mus musculus* (Mm), *Gallus gallus* (Gg), *Xenopus tropicalis* (Xt) and *Danio rerio* (Dr). The residues participating in MICU1–MICU2 interactions are indicated with magenta circles.

# Article



**Extended Data Fig. 9 | Sequence alignment of MICU1 from different species.**

Sequence alignment of MICU1 homologues from *Homo sapiens* (Hs), *Gallus gallus* (Gg), *Danio rerio* (Dr), *Ciona intestinalis* (Ci), *Drosophila melanogaster*

(Dm), *Caenorhabditis elegans* (Ce), *Arabidopsis thaliana* (At) and *Dictyostelium discoideum* (Dd). The residues participating in MICU1-MICU2 and MICU1-MCU interactions are indicated with magenta and blue circles, respectively.

**Extended Data Table 1 | Cryo-EM data collection, refinement and validation statistics**

	Low-Ca uniplex (EMDB-21642) (PDB 6WDN)	High-Ca uniplex (EMDB-21643) (PDB 6WDO)
<b>Data collection and processing</b>		
Magnification	130,000	36,000
Voltage (kV)	300	200
Electron exposure (e-/Å <sup>2</sup> )	63	30
Defocus range (μm)	1.5-3.0	1.5-2.5
Pixel size (Å)	1.06	1.1
Symmetry imposed	C1	C2
Initial particle images (no.)	900,008	2,572,617
Final particle images (no.)	64,131	101,531
Map resolution (Å)	3.3	3.6
FSC threshold	0.143	0.143
Map resolution range (Å)	2.8-6.2	3.1-6.4
<b>Refinement</b>		
Initial model used (PDB code)	6O5B, 4NSC, 6AGH	6O58, 4NSD, 6IIH
Model resolution (Å)	3.6	4.0
FSC threshold	0.5	0.5
Model resolution range (Å)	3.15-381	3.43-381
Map sharpening <i>B</i> factor (Å <sup>2</sup> )	-54	-98
Model composition		
Non-hydrogen atoms	12,612	29,867
Protein residues	1,540	3,645
Ligands	0	2
<i>B</i> factors (Å <sup>2</sup> )		
Protein	135.59	130.26
Ligand	-	96.81
R.m.s. deviations		
Bond lengths (Å)	0.005	0.005
Bond angles (°)	0.729	0.716
Validation		
MolProbity score	2.12	2.37
Clashscore	14.61	21.32
Poor rotamers (%)	0.52	0.77
Ramachandran plot		
Favored (%)	92.9	90.2
Allowed (%)	7.1	9.8
Disallowed (%)	0	0

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

## Software and code

Policy information about [availability of computer code](#)

Data collection

SerialEM (v 3.7)

Data analysis

Phenix (v 1.16-3549); Coot (v 0.8.9); MolProbity (v 4.4); Pymol (v 1.8); MotionCor2 (v 1.1.0); CTFFIND4 (v 4.1.5); RELION 3.0; Chimera (v 1.13); cisTEM (1.0.0-beta); cryoSPARC (2.9.0)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The three-dimensional cryo-EM density maps are deposited into the Electron Microscopy Data Bank under accession numbers EMD-21642 and EMD-21643. The coordinates are deposited into the Protein Data Bank with accession numbers 6WDN and 6WDO.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences

Behavioural & social sciences

Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](http://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. Sufficient cryo-EM data were collected to achieve adequate map reconstructions.
Data exclusions	No data were excluded from analyses.
Replication	Biochemical and functional experiments were repeated at least three times (except twice in Extended Data Fig. 7d) in independent experiments. Experimental findings were reproduced reliably. Cryo-EM data processing was not repeated. Following the same procedures, the data processing would be reproducible.
Randomization	This is not relevant to our study, because no grouping was needed.
Blinding	Investigators were not blinded to group allocation, because no grouping was needed for this study.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies	<input type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines	<input type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input type="checkbox"/>	<input checked="" type="checkbox"/> MRI-based neuroimaging
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		

## Antibodies

Antibodies used	Anti-Flag (Sigma F1804, Lot #SLBX2256, 1:10,000) Anti-EMRE (Santa Cruz 86337, Lot #K0215, 1:400) Anti-LetM1 (Abcam ab55434, Lot #GR148824-4, 1:2,000) Anti-Tim23 (Santa Cruz sc-514463, Lot #R1219, 100 ng/mL) Anti-MICU2 (Abcam ab101465, Lot #GR3212268-15, 1:10,000) Anti-COX2 (Abcam ab110258, Lot #GR3239988-5, 100 ng/mL) Anti-mitoflin (Abcam ab110329, Lot #GR3299197-4, 100 ng/mL) Anti-VDAC1 (Abcam ab14734, Lot #GR3296736-10, 100 ng/mL) Anti-1D4 (home-made, 100 ng/mL)
Validation	<ol style="list-style-type: none"> <li>1. Anti-Flag antibody is validated by the manufacturer (Sigma): the antibody detects a single target band on a Western blot from mammalian crude cell lysate.</li> <li>2. Anti-LetM1 antibody is validated by the manufacturer (Abcam): the antibody detects a single band at ~75 kDa on a Western blot from HeLa cell lysate.</li> <li>3. Anti-Tim23 antibody is validated by the manufacturer (Santa Cruz): the antibody detects a single band at ~23 kDa on a Western blot from the lysate of MCF7 cells.</li> <li>4. Anti-MICU2 antibody is validated by the manufacturer (Abcam): the antibody detects a single band at ~45 kDa on a Western blot from U87-MG whole cell lysate.</li> <li>5. Anti-COX2 antibody is validated by the manufacturer (Abcam): the antibody detects a single band at ~25 kDa on a Western blot from mitochondrial lysate from human heart tissue.</li> <li>6. Anti-mitoflin antibody is validated by the manufacturer (Abcam): the antibody detects a single band at ~100 kDa on a Western blot from human cerebellum tissue.</li> <li>7. Anti-VDAC1 antibody is validated by the manufacturer (Abcam): the antibody detects a single band at ~35 kDa on a Western</li> </ol>

blot from mitochondrial lysate from human heart tissue.

8. Anti-1D4 antibody was validated in-house: the antibody detects a single band at ~35 kDa on a Western blot from the crude lysate of HEK293 cells transiently expressing an 1D4-tagged MCU protein. The Western signal can be blocked by a synthesized 1D4 peptide (TETSQVAPPA). Detailed results were documented in Tsai et al., (eLife 2016).

9. Anti-EMRE antibody was validated in-house: the antibody detects a single and at ~15 kDa on a Western blot from HEK293 cell lysate. Moreover, the Western signal was eliminated by EMRE-KO. Results were documented in Tsai et al., (eLife 2016).

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

sf9 (Expression Systems, #94-001S); HEK293s (ATCC, #CRL-3022); HEK 293T (ATCC, #CRL-11268)

Authentication

HEK 293T cells for functional experiments were authenticated by ATCC using short tandem repeat (STR) profiling when purchased (Jan, 2012), and was re-authenticated by ATCC using STR profiling in May 2016. sf9 and HEK293S cells were only used for producing proteins for structural determination. These cells were not further authenticated after being purchased.

Mycoplasma contamination

HEK293T cells used for functional studies were checked for mycoplasma contamination semi-annually using PCR (no contamination has been detected in our last test in Jan 2020). sf9 and HEK293S cells for structural studies have not been tested for mycoplasma contamination.

Commonly misidentified lines  
(See [ICLAC](#) register)

None of the cell lines used are listed in the ICLAC database.

## Corrections & amendments

# Retraction Note: miR-34a blocks osteoporosis and bone metastasis by inhibiting osteoclastogenesis and Tgif2

---

<https://doi.org/10.1038/s41586-020-2273-1>

Retraction to: *Nature* <https://doi.org/10.1038/nature13375>

Published online 25 June 2014

 Check for updates

---

Jing Y. Krzeszinski, Wei Wei, HoangDinh Huynh, Zixue Jin,  
Xunde Wang, Tsung-Cheng Chang, Xian-Jin Xie, Lin He,  
Lingegowda S. Mangala, Gabriel Lopez-Berestein, Anil K. Sood,  
Joshua T. Mendell & Yihong Wan

---

Upon re-examination of the bone histomorphometry data in Extended Data Figs. 1i, 2d, 3h, 4h, 5n, 6e, 9g and 10f of this Letter, anomalies were found that call into question the integrity of these data. These concerns undermine the confidence in the study and the authors thus wish to retract the Letter in its entirety. The authors regret this situation and apologize to the scientific community. All authors agree with the Retraction, but author Xunde Wang did not respond.

# Work

Your  
story

Send your careers story  
to: naturecareerseditor  
@nature.com



Meetings can be accessed while working from home.

## LEARNING TO LOVE VIRTUAL CONFERENCES

COVID-19 has provided an opportunity to rethink the scientific conference. How do researchers make the most of online meetings? **By Chris Woolston**

Lynn Cominsky, an astrophysicist at Sonoma State University in California, had planned to spend part of late March and early April in Johannesburg, South Africa, for the Ninth Fermi International Symposium, an event that would have gathered hundreds of astronomers and astrophysicists with an interest in γ-ray surveys conducted by the Large Area Telescope (LAT) and the Gamma-ray Burst Monitor. When that plan, like so many others, was scuttled by the

COVID-19 pandemic, she ended up attending a virtual meeting with fellow members of the LAT collaboration – no plane ticket or hotel room required.

For her, it was a disappointing, but eventually fruitful, fall-back plan. “I enjoy going to conferences,” she says. “Some of the best proposals and projects that I’ve thought about have come from random interactions with people at conferences.”

Online meetings might lack many of the

benefits of an in-person conference: conversations over dinner; face-to-face networking; fresh perspectives that can come from simply leaving one’s home ground. But as Cominsky and others have found, there’s a lot to like about virtual conferencing. And as more meetings move online – a trend likely to continue even after the pandemic fades – researchers are increasingly trying to accept the new virtual reality of scientific gatherings. “I was able to listen to a bunch of talks without leaving

## Online Course on Effective Collaboration in Research

Learn how to  
participate in and  
lead successful  
collaborative projects.

→ Find out more at  
**masterclasses.  
nature.com**



Bite-size design for  
busy researchers

Subscribe as a lab  
or institution

## Work / Careers

home," Cominsky says.

Prompted by *Nature*'s coverage of the virtual conference of the American Physical Society (APS) in April, researchers shared their personal, and generally positive, experiences of online conferences in a *Nature* online poll. Of 485 respondents, 82% said that they would be willing to attend an online conference in future. *Nature* reached out to a sample of those respondents for more on what they had learnt and their advice for future attendees.

### Embrace the awkward

Cominsky says that scientists shouldn't expect online conferences to run perfectly, especially when the conversion to online was unexpected and hastily planned. She attended a few of the virtual presentations from the APS meeting, which was put together very quickly after it became clear that a physical gathering would be impossible. The event was widely considered a success, drawing more than 7,200 registrants compared with an average of 1,600 to 1,800, but there were some glitches and awkward moments along the way. "It's OK to acknowledge that this is a weird time and that you probably aren't going to get as much out of the situation as you were hoping," she says. "If remote conferences start with that kind of acknowledgement, people can feel like they're not alone."

For attendees, the awkwardness may include moments of less-than-optimal sound and occasional technical glitches. Presenters, for their part, could find it challenging to deliver a talk to a monitor instead of to a live audience. "It's hard to know what kind of reaction you're getting if you can't see people's faces," Cominsky says. "It's hard to tell if people understood what you just said."

### Future organizers should mix it up

Chris Wood, a microscopist at the National Autonomous University of Mexico in Mexico City, says that a lack of travel funding had ended any hope of attending the annual Focus on Microscopy meeting, which had been scheduled for April in Osaka, Japan. For Wood and many other scientists – especially those in developing countries – regular travel to far-off conferences is simply impossible. "Grants are quite thin here," he says. "You have to trim your budget to get a grant approved, and travel is the first thing to go."

But when the COVID-19 pandemic forced that conference to become virtual, Wood was able to view several presentations that were moved online, an opportunity that he hopes will repeat itself for future conferences. "It democratizes access in developing countries," he says. "I don't see why we should be excluding tens of thousands of students who have just as much right to attend."

Wood thinks that more conferences in future should adopt a 'hybrid' approach, with

both physical and virtual attendees. The online attendees could pay a registration fee to support the conference, but attendance should be affordable for everyone. "I would be happy to pay for a conference that's especially relevant to me," he says.

### Take advantage of the availability

As more conferences go online, some researchers have moved into binge mode. Paul DeStefano, a graduate student at Portland State University in Oregon, says he has watched about 15 sessions so far from the APS conference, a feat that would have been impossible if he had been there in person. In future, he says, a little more discretion might be in order. "I really enjoyed some of the sessions, but maybe I didn't need to see every last one," he says. "That's something I need to adjust."

Tom Brown, an energy-systems modeller at the Karlsruhe Institute of Technology in

**"I would be happy to pay for a conference that's especially relevant to me."**

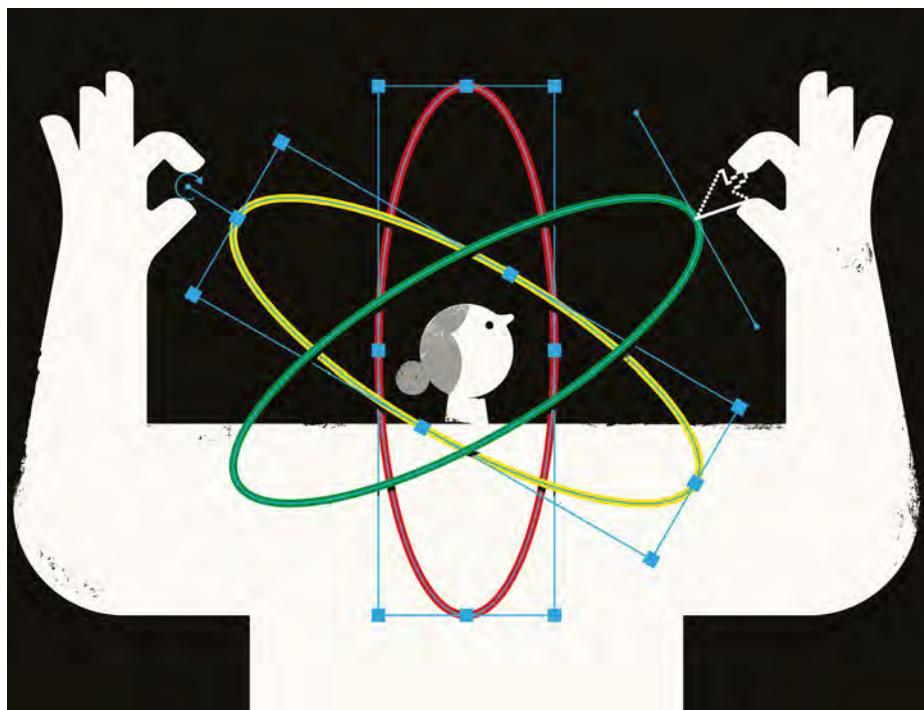
Germany, is similarly dedicating a significant amount of his time to watching virtual conference sessions. "The availability has massively increased," he says. "I've attended workshops in the morning in Europe, and in the evening in the States." Brown, who says he had already stopped flying to conferences to help save energy and protect the climate, has a system to manage his virtual attendance. "I save conferences for Tuesdays and Thursdays," he says. "Mondays, Wednesdays and Fridays are free for actual work."

### Stay interactive

Virtual conferences might lack the intimacy of a physical gathering, but it's still possible for attendees to connect with each other. The APS conference, for example, provided a chat function for real-time feedback. Brown says that he recently sent a follow-up e-mail to a presenter, the virtual equivalent of approaching a speaker after a talk. "I got a response a couple of days later," he says. "If you want to get something going, you have to give them a reason to respond." Some students and other early-career researchers might even find digital communication with eminent scientists to be less intimidating than a face-to-face encounter.

Cominsky notes that junior researchers often use conferences as a way to meet future collaborators. That's harder to do in the virtual world, but it's not impossible. "I would check out the programme and reach out to people ahead of time to form a relationship," she says.

**Chris Woolston** is a freelance writer in Billings, Montana.



# GRAPHIC CONTENT: PICTURING SCIENCE

The web-based tool BioRender has become a staple of biomedical research drawings.

By Jeffrey M. Perkel

**L**ike many researchers, Rita Sattler is trained to conduct science, not to draw it. But papers still need writing – and illustrating. So, when the time came to document what her team had done for a journal article, Sattler did what her colleagues frequently do: she fired up PowerPoint.

Microsoft's tool for creating and displaying slideshow presentations is one of two popular options for creating scientific illustrations; the other is Adobe Illustrator. But Illustrator is a professional-grade tool with a notoriously steep learning curve, and PowerPoint's strength is slide shows, not graphic design. "I'm not gonna give the adjective that I really want to say, but they were just really not nice," Sattler says of the resulting images.

In 2019, a graduate student working in Sattler's laboratory suggested an alternative. BioRender, a web-based tool designed for life-science and medical illustration, is like a feature-lite version of Illustrator. Developed in Toronto, Canada, by a start-up of the same name, and launched in 2017, it features an extensive library of scalable 'icons' from across the life sciences and medicine; researchers can

drop these onto a canvas and manipulate them as if they were circles or polygons.

For Sattler, a neuroscientist at the Barrow Neurological Institute in Phoenix, Arizona, that made it possible to quickly illustrate cell-differentiation pathways without having to draw each cell from scratch. "Everybody was like, oh my God, this is awesome," Sattler says. "We jumped on this very, very quickly, and we've been using really nothing else since then."

### Show, not tell

Scientific success requires an effective level of communication, whether with peer reviewers, funding agencies or colleagues. In a field in which the communication channels that count – journal articles and grant applications – often have page limits, an illustration can be worth far more than a thousand words. "If you want to communicate your findings, or in the case of a grant application if you just want to explain what you want to do, it's really important to use graphics," says Wilfried Rossoll, a cell biologist at Mayo Clinic in Jacksonville, Florida.

Unlike data figures that detail primary

research findings, these graphics are typically illustrated explainers of proposed models, experimental methods or biochemical pathways. More and more journals allow researchers to include graphical abstracts, for instance – illustrations intended to summarize the key conclusions of a paper.

Researchers typically create those illustrations using PowerPoint or Illustrator, or analogues thereof. But none of these was designed specifically for scientists, so they can be challenging to use. "I would have banged my head against the wall" trying to use Illustrator had it not been for BioRender, says Signe Elisabeth Åsberg, a sepsis researcher at the Norwegian University of Science and Technology in Trondheim, as she recalls the graphics she made for her thesis. Akiko Iwasaki, an immunobiologist at Yale University in New Haven, Connecticut, echoes that sentiment. Using Illustrator, she says, "it would take me days to create figures of tissues or cell types or vasculature. But with BioRender, within minutes I can draw what I need."

### Fully equipped

That's thanks mostly to BioRender's library of around 30,000 life-science icons, which includes anatomical drawings and depictions of everything from SARS-CoV-2 virus particles to fruit flies. Users can resize, rotate and change the colour of those icons. But they cannot change their fundamental appearance, for instance to add or remove a protein domain. The library also includes icons for specific pieces of laboratory kit, making it possible to illustrate protocols with images of the actual equipment used.

Among the items depicted are the Orbitrap Fusion Tribrid mass spectrometer, made by Thermo Fisher Scientific in Waltham, Massachusetts, and the MinION DNA sequencer, from Oxford Nanopore Technologies in the United Kingdom. Beth Kenkel, an associate scientist at Bristol Myers Squibb in Seattle, says that this feature helps in her presentations to her team. "I can quickly make a graphic representation of how I plan to do an experiment. And then I can solicit feedback from my co-workers: is this how I should design it, or should I change something?"

Researchers can also create icons that represent specific structures in the Protein Data Bank, an open-access digital archive providing access to 3D structure data for proteins and nucleic acids. In March the company rolled out an enhanced Protein Data Bank interface that enables BioRender to produce "more painterly" renderings than its earlier iteration, says Shiz Aoki, a medical illustrator who co-founded BioRender. Users can also request custom icons from the BioRender team. Rossoll, for example, has ordered icons that represent two pieces of equipment used in tissue processing for immunohistochemical

## Work / Technology & tools

analysis. Created at no charge, those icons are now available to all BioRender users. "You're welcome," he quips.

### All lined up

Fleshing out BioRender's toolbar are drawing tools specific to life-science research. There are pre-made lines and arrow styles for biochemical pathways, for instance; an arc tool for plasmid maps (graphical representations of small circular pieces of DNA commonly used in molecular biology research); and a new series of brushes for drawing plasmid membranes, vasculature and nucleic acids as if they were lines.

Andrew Deans, a protein chemist at St Vincent's Institute in Melbourne, Australia, published a paper (*J. O'Rourke et al. Nucleus* **10**, 221–230; 2019) in October that features four graphics created in BioRender. Deans is comfortable using Illustrator, but notes that it can be difficult to train new people to

use it. As a result, he says, his students 'love' BioRender. But they still create data figures with Illustrator and Adobe Photoshop, he notes. And one member of his team, who is comfortable drawing freehand, creates illustrations using the Tayasui Sketches iPad app and an Apple Pencil.

And there are other options, although none is as bio-focused as BioRender. Kenkel, who also blogs for the reagent distribution service Addgene, says that she illustrates her posts mostly using Vectr, a general-purpose vector-graphics tool, and Canva, an online tool for creating social-media images. Both tools are free, web-based and easy to use, she says, and "basically tick all the boxes for me".

Ken Hughes, an oceanography postdoc at Oregon State University in Corvallis, uses Inkscape, an open-source Illustrator analogue whose version 1.0 release came out on 1 May. And Zen Faulkes, who studies crustacean biology at the University of Texas Rio Grande

Valley in Edinburg, favours CorelDRAW, a commercial tool that he supplements with images from the Noun Project, a collection of more than 2 million freely available icons. (Another option is the Servier Medical Art collection from the pharmaceutical company Servier in Suresnes, France; this includes some 3,000 free-to-use images.)

According to Aoki, BioRender now has some 200,000 users. It's free to use, but the resulting images are watermarked, and output-file resolution and format are limited. Figure publication requires a licence (US\$35 per month, or \$99 per month for 5 users). Users can start working from a blank canvas, but the company also provides hundreds of predesigned templates, including several related to SARS-CoV-2. One of these, detailing the coronavirus replication cycle, is currently the most-used template on the site, Aoki says.

Jeffrey M. Perkel is *Nature*'s technology editor.

# AFRICAN SCIENTISTS LEVERAGE OPEN HARDWARE

A growing emphasis on do-it-yourself science is helping researchers to equip labs in resource-limited areas. By Abdullahi Tsanni

**A** 2018 article in the journal *HardwareX* details "an open source hardware setup to measure locomotor activity in rodents". It has a simple design. But for developer Victor Kumbol, then a neuroscience master's student at the Kwame Nkrumah University of Science and Technology in Kumasi, Ghana, that device, called Actifield, has had an outsized impact.

Actifield is an actimeter, a device that quantifies animal activity (V. Wumbor-Apin Kumbol *et al. HardwareX* <http://doi.org/ggb8hw>; 2018). "I needed the actimeter to test for potential drug compounds that could modify behaviour in mice. But my department had only one actimeter available, and it was outdated," Kumbol says. "So I decided to build one for myself."

With colleagues Elikplim Ampofo and Mary Twumasi, Kumbol travelled to Cape Town, South Africa, to attend a two-week workshop run by the non-profit organization Teaching and Research in Natural Sciences for Development (TReND) in Africa, which promotes research and education on the African continent.

Founded in 2011 by Lucia Prieto-Godino, now at the Francis Crick Institute in London, Sadiq Yusuf at the Uganda Technology and

Management University in Kampala and Tom Baden at the University of Sussex in Brighton, UK, TReND in Africa encourages do-it-yourself research with a focus on low-cost, open-source science. Courses cover such topics as fly genetics, neuroscience and hardware development.

In one example, TReND in Africa instructor and University of Sussex bioengineer André Maia Chagas joined a team including Prieto-Godino and Baden in 2017 to design a microscope. It was built using off-the-shelf and 3D-printed components, and dubbed the €100 lab (A. Maia Chagas *et al. PLoS Biol.* **15**, e2002702; 2017).

"Open-science hardware is not only important in Africa but all over the world," Chagas says. "If you have the blueprint for a piece of equipment, you can understand how it works. You can repair your equipment if it breaks down, and, even more importantly, adapt it to your local needs."

Kumbol's Actifield device is an array of infrared light emitters and detectors inside a box, run using a microcontroller made by the open-source developer Arduino. As rodents move around in the box, they disrupt the beams and the actimeter counts those events.

At US\$122.91, the device is a fraction of the cost of commercial systems, which can amount to \$6,000 or more. The resulting paper in *HardwareX* helped Kumbol to secure funding from the Mozilla Foundation to organize a follow-up workshop last July at the University of Health and Allied Sciences in Ho, Ghana. He has demonstrated Actifield at the Kwame Nkrumah University of Science and Technology, and has built actimeters for the science department there.

Now a PhD student at the Einstein Center for Neurosciences in Berlin, Kumbol has joined the editorial board of *HardwareX*. In February, he was invited to speak about using open-science tools to tackle equipment challenges in Africa at a VIB Core Management Workshop in Leuven, Belgium.

"It was a great experience for me," Kumbol says. "I received a lot of positive feedback on the potential for open hardware, which has really motivated me to start working on the next steps."

Says Chagas, "This is what we want: a new generation of African scientists that will train others on the continent."

**Abdullahi Tsanni** is a science writer based in Abuja, Nigeria.



### Where I work Christian Happi

I study microbial diseases such as COVID-19, Lassa fever and Ebola. My research team at Redeemer's University in Ede, Nigeria, focuses on detecting and characterizing pathogens, so that we can identify the aetiology of previously inexplicable fevers in west Africa. The coronavirus pandemic fits neatly into the scope of my work, and our priority has shifted to COVID-19.

Since Nigeria confirmed its first case of the disease on 27 February, my team and I have been working in staggered shifts in my laboratory. We receive human-tissue samples from states in Nigeria, and then test them for traces of SARS-CoV-2, the coronavirus that causes the disease. We use a technique called polymerase chain reaction, which enables us to detect specific genetic material from the coronavirus if it is present.

On 1 March, I drove 90 kilometres from my lab to the airport in Ibadan, to receive clinical specimens of the coronavirus from the first person in Nigeria known to have contracted COVID-19. I confirmed the presence of the

SARS-CoV-2 virus in those specimens, and my team and I sequenced the virus's genome within 72 hours – an incredible feat. Here, I'm standing beside one of the two sequencing platforms in my lab.

We are now sequencing more coronavirus genomes and developing a rapid molecular diagnostic test that uses CRISPR gene-editing technology to detect SARS-CoV-2. We hope to move this work to remote areas in Africa soon. The government considers our work essential, and my team and I all have passes to get into the lab during the lockdown.

People might have thought that this work was impossible in Africa, but we are demonstrating that the continent's scientists can generate crucial data in the global fight against COVID-19 – as well as contributing to the field of genomics. I feel inspired to be a part of this new narrative.

**Christian Happi** is a molecular biologist at the African Centre of Excellence for Genomics of Infectious Diseases at Redeemer's University in Ede, Nigeria. **Interview by Abdullahi Tsanni**.

Photographed by the African Centre of Excellence for Genomics of Infectious Diseases.

## Corrections & amendments

# Author Correction: Axionic charge-density wave in the Weyl semimetal $(\text{TaSe}_4)_2\text{I}$

---

<https://doi.org/10.1038/s41586-020-2303-z>

Correction to: *Nature* <https://doi.org/10.1038/s41586-019-1630-4>

Published online 07 October 2019



J. Gooth, B. Bradlyn, S. Honnali, C. Schindler, N. Kumar, J. Noky,  
Y. Qi, C. Shekhar, Y. Sun, Z. Wang, B. A. Bernevig & C. Felser

---

In the Acknowledgements section of this Article, the statement 'B.A.B. was supported by the US Department of Energy through grant number DE-SC0016239, the US National Science Foundation EAGER through grant number NOA-AWD1004957, Simons Investigator grants ONR-N00014-14-1-0330 and NSF-MRSEC DMR- 1420541, the Packard Foundation and the Schmidt Fund for Innovative Research.' should be replaced by '**First-principles calculations were funded** by the US Department of Energy through grant number DE-SC0016239. **B.A.B. acknowledges additional support from** the US National Science Foundation EAGER through grant number NOA-AWD1004957, Simons Investigator grants ONR-N00014-14-1-0330 and NSF-MRSEC DMR-1420541, the Packard Foundation and the Schmidt Fund for Innovative Research.'. The original Article has been corrected online.